Comparative Learning

Word Count: 12309

Abstract

In this paper, I'm going to develop a new perspective on the norms of diachronic rationality. Unlike existing approaches, which focus on how we should change either our qualitative beliefs or numerical credences over time, this approach is articulated in terms of what I call 'comparative confidence judgements', i.e. judgements of the form 'I am at least as confident in p as I am in q'. To date, there has been no attempt to systematically address the question of how rational agents should change their comparative confidence judgements over time as they gather new evidence. This paper fills this lacuna by identifying, characterising and evaluating an intuitively compelling learning rule that specifies how agents should revise their comparative confidence judgements in the face of novel evidence.

1 Introduction

We humans are prone to believing things, like when I believe that Trollis is on the sofa. We are also prone to lending credence to things, like when I lend a credence of around 0.1 to there being rain in Windhoek tomorrow. Finally, we are also prone to making comparative confidence judgments, like when I am more confident that Trollis is on the sofa than I am that it will rain in Windhoek tomorrow. While epistemic attitudes of the first two kinds (qualitative belief and numerically graded credence) are widely taken to play a crucial role in framing the fundamental norms by which the rationality of an agent's epistemic states are to be assessed, comparative confidence judgements have attracted much less attention in the contemporary philosophical literature. This is somewhat surprising, given that several eminent figures in the history of inductive inference – e.g. Keynes (1921), de Finetti (1937, 1951), Koopman (1940) and Fine (1973) – have contended that comparative confidence judgements are the most fundamental, intuitive and psychologically basic of all our epistemic attitudes.[1]

[1] Thus, we read, for example,

Over the years, numerous authors have attempted to identify synchronic rationality requirements for comparative confidence orderings (see e.g. Halpern (2003) for a thorough overview). However, the philosophical foundations of this project have, until recently, been largely neglected,[2] and there is still little consensus regarding what kinds of comparative confidence structures are characteristic of rational agents. Happily, this situation is beginning to improve. Icard (2016) has shown that 'money-pump' style arguments can be used to provide a prospective pragmatic justification of the requirement that a rational agent's comparative confidence judgments should always be representable by a probability function. Meanwhile, Fitelson and McCarthy (unpublished) have shown that accuracy dominance arguments can be used to provide epistemic justifications for some significantly weaker synchronic coherence requirements (such as the principle that a rational agent's comparative confidence judgments should always be representable by a Dempster-Schafer belief function).

But despite recent progress in identifying the synchronic coherence constraints that govern the comparative confidence judgments of rational agents at a time, practically nothing has been written on the question of how rational agents should change their comparative confidence judgments *over time* as they gather new evidence. This is the problem with which I'll be concerned in this paper.

Before moving on, it is worth briefly clarifying an important point. My aim in this paper is *not* to justify the claim that the epistemic states of ideally or boundedly rational agents should be conceived of in terms of comparative confidence judgements rather than qualitative beliefs or (precise or imprecise) numerical credences. Rather, I assume in the background that there are at least some scenarios in which such a conception is principled, and then consider the question of how epistemic states, thus conceived, should evolve over time. After all, it is surely at least possible to conceive of a creature whose epistemic state is characterised purely by comparative confidence judgements, and it is surely philosophically interesting to ask what kinds of epistemic norms would determine the rationality of such a creature's reasoning. As I mentioned above, the comparative conceptualisation of epistemic states has numerous illustrious champions, and I mainly take it for granted that the reader will agree that a proper understanding of the dynamics of rational comparative confidence is a worthy philosophical goal.

The fundamental viewpoint of the present work is that the primal intuition of probability expresses itself in a (partial) ordering of eventualities: A certain individual at a certain moment considers the propositions a, b, h, k,...Then the phrase 'a on the presumption that h is true is equally or less probable than b on the presumption that k is true' conveys a precise meaning to his intuition... This is, as we see it, a first essential in the thesis of intuitive probability, and contains the ultimate answer to the question of the meaning of the notion of probability. (Koopman, 1940: 270)

[2]See Fine (1973) for a critical assessment of the philosophical motivations behind several synchronic rationality requirements from the literature.

The structure of the paper is as follows. In section 2, I introduce the standard formalism for analysing comparative confidence judgments and provide a concise summary of some of the most important synchronic coherence constraints from the literature, before briefly reviewing some recent work regarding the normative status of these constraints. In section 3 I turn to the central question of the paper: 'how should a rational agent revise their comparative confidence judgements over time as they acquire new evidence?'. I address this question by studying the way in which a Bayesian agent's comparative confidence judgements change when they conditionalize on new evidence. I then show that the resulting revision rule (which I call 'comparative conditionalisation' (CC)) is intuitively compelling, even outside of the context of probabilistic Bayesian epistemology, and establish some basic properties of CC, before going on to illustrate two important senses in which the comparative rule requires less epistemic structure for its application than its numerical counterpart. The aim of the subsequent sections is to investigate whether the normative arguments that are normally given as justifications for Bayesian conditionalisation can be generalised to provide normative justifications for CC. In section 4, I provide an evidentialist motivation for CC and argue that this motivation is on far sounder footing than an analogous argument that is commonly given for Bayesian conditionalisation. In section 5 I construct and evaluate a prospective pragmatic diachronic Dutch book argument for CC, and catalogue the decision rules that are needed to get such an argument off the ground. In section 6, I show that perhaps the most influential *epistemic* justification of Bayesian conditionalisation, namely the argument from expected inaccuracy, cannot be straightforwardly generalised to the comparative setting (since none of the extant formalisations of inaccuracy for comparative confidence judgements allow for the meaningful articulation of an analogous justification for CC). Section 6 draws some morals, provides a prospectus for future work, and concludes.

2 Coherence Conditions for Confidence Orderings

2.1 Preliminaries

I begin with some technical preliminaries. Firstly, I assume that agents always make comparative confidence judgements about 'propositions' drawn from the Boolean algebra $\mathfrak{B}$ of equivalence classes of logically equivalent sentences of some language $\mathcal{L}$.[3] Intuitively, an agent A can make two kinds of

[3]For simplicity, I assume that $\mathfrak{B}$ and L are always finite. The assumption that the relata of comparative confidence judgements are logical equivalence classes rather than simple sentences can be seen as a logical omniscience assumption, i.e. that the agent is always aware of all logical equivalences.

comparative confidence judgement about propositions in $\mathfrak{B}$. Firstly, they can be strictly more confident in the truth of p than they are in the truth of q. I denote this kind of judgement with the notation $p \succ q$. Alternatively, A can be equally confident in the truth of p and q. I denote this second kind of judgement with the notation $p \sim q$.[4] Together, the set of all A's comparative confidence judgements define a *confidence ordering*, $\succsim$, over some subset of the propositions in $\mathfrak{B}$. I write $p \succsim q$ to indicate the disjunction '$p \succ q$ or $p \sim q$'. I turn now to briefly outlining some of the most important basic structural properties that authors typically assume are satisfied by $\succsim$. Firstly, I follow orthodoxy in assuming that $\succ$ always satisfies the following conditions.

Irreflexivity of $\succ$: For every $p \in \mathfrak{B}$, A does not make the judgement $p \succ p$, i.e. $p \not\succ p$.

Transivity of $\succ$: For every $p, q, r \in \mathfrak{B}$, if $p \succ q$ and $q \succ r$, then $p \succ r$.

Secondly, I assume that $\sim$ is an equivalence relation, i.e.

Reflexivity of $\sim$: For every $p \in \mathfrak{B}$, $p \sim p$.

Transivity of $\sim$: For every $p, q, r \in \mathfrak{B}$, if $p \sim q$ and $q \sim r$, then $p \sim r$.

Symmetry of $\sim$: For every $p, q \in \mathfrak{B}$, if $p \sim q$, then $q \sim p$.

When all of these assumptions are satisfied, I say that the ordering $\succsim$ is a 'partial preorder' over $\mathfrak{B}$. For the remainder of the article, I will assume that the confidence orderings being considered are partial preorders over $\mathfrak{B}$, unless otherwise stated. Many authors also additionally assume the following constraint on $\succsim$

Opinionation: For any $p, q \in \mathfrak{B}$, A makes exactly one of the judgements $p \succ q$, $q \succ p$, $p \sim q$.

Combined with the assumption that $\succsim$ is a partial preorder, Opinionation implies that $\succsim$ is a 'total preorder' over $\mathfrak{B}$.[5] Intuitively, this means that there are 'no gaps' in A's confidence judgements, i.e. A makes a comparative confidence judgement about every pair of propositions in $\mathfrak{B}$. This assumption, though controversial, is standard in the extant literature on comparative confidence orderings.[6] I will not generally assume Opinionation for the rest of this paper, and the learning rule I introduce in Section 3 (as well as its evidential justification in Section 4) is perfectly applicable in non-opinonated

[4]To be clear, $p \sim q$ denotes the judgement that p and q are equally plausible. Depending on one's view of epistemic indifference, this may or may not be distinct from simply being epistemically indifferent between p and q. See Eva (2019) for a discussion of comparative conceptions of epistemic indifference.

[5]Fitelson and McCarthy (unpublished) work in a more general setting than that described here. Specifically, they consider an agent's comparative confidence over arbitrary (possibly proper) subsets of $\mathfrak{B}$, which they call 'agendas'. They then assume only that $\succsim$ is opinionated with respect to the given agenda.

[6]For philosophical critiques of the Opinionation assumption, see e.g. Keynes (1921) and Forrest (1989). One might plausibly contend that one of the primary advantages of conceiving of an agent's epistemic states in terms of comparative confidence judgements rather than numerical credences or qualitative beliefs is that it allows us to study the epistemological consequences of failures of Opinionation.

settings. However, the Dutch book argument outlined in Section 5 is limited to the special case in which the learning rule is applied to prior confidence orderings that satisfy Opinionation. In cases where Opinionation fails and there exist $p, q \in \mathfrak{B}$ such that $\neg(p \succsim q)$ and $\neg(q \succsim p)$, I will write '$p \odot q$' and say that p and q are 'incomparable' in the agent's confidence ordering. I emphasise that this does not constitute an additional category of comparative confidence judgement, but rather the absence of any comparative confidence judgement whatsoever. Finally, the following additional constraint is also sometimes assumed (see e.g. Fitelson and Mccarthy (unpublished)):

Regularity of $\succsim$: For any contingent $p \in \mathfrak{B}$, $\top \succ p \succ \bot$.

Regularity requires that A is always strictly more confident in the tautology than they are in any contingent proposition, and that they are always strictly less confident in the contradiction than they are in any contingent proposition. This is a generalisation of the controversial Regularity condition from Bayesian epistemology.[7] Now, the purpose of this paper is to analyse how a rational agent should revise their comparative confidence judgements upon learning, *with certainty*, the truth of some evidential proposition e. Of course, the very possibility of this kind of learning is ruled out a-priori by the Regularity assumption, so I do not assume Regularity in what follows.[8]

2.2 Synchronic Norms

Before attempting to explicate the diachronic rationality norms that govern the way in which an agent A should revise their comparative confidence judgements over time, it behooves us to review some of the most influential substantive synchronic rationality constraints that have been forwarded in the literature.[9] In comparison to the basic structural properties described in the previous subsection, which are commonly treated as implicit assumptions that are partially constitutive of the very notion of rational comparative confidence, the constraints described in this subsection are both more substantive and more controversial.

The first and most fundamental synchronic rationality constraints for comparative confidence orderings can be stated both qualitatively and in terms of possible numeric representations. I present the qualitative versions first.

[7]See e.g. Lewis (1980) and Skyrms (1980) for philosophical justifications of the regularity condition.

[8]Naturally, the familiar debate surrounding whether one can ever learn a proposition with certainty can be rerun in the comparative setting, but it seems to me that there probably isn't much to gain by doing so. Here, I am of course assuming that it is possible to learn a proposition with certainty, and hence that the kind of learning experience described by standard Bayesian conditionalisation can actually occur.

[9]Note that the literature is replete with possible synchronic coherence constraints for confidence orderings, and it would be impossible to provide an exhaustive survey here (the interested reader should consult e.g. Halpern (2003), Wong *et al* 1991). I review only those synchronic constraints that play a crucial role in what follows.

$$(A1)\ \top \succ \bot.$$

$$(A2)\ \text{For any } p, q \in \mathfrak{B}, \text{ if } p \vdash q \text{ then } q \succsim p.$$

$A1$ requires that rational agents always be strictly more confident in the tautology than they are in the contradiction, and $A2$ is a general monotonicity requirement, which stipulates that agents should never be strictly more confident in p than they are in the logical consequences of p. As well as being intuitively compelling, these rationality constraints have been given a range of pragmatic justifications (see e.g. Fishburn (1986), Halpern (2003)).

Given a comparative confidence ordering $\succsim$ over $\mathfrak{B}$ and a set S of functions $\mu : \mathfrak{B} \to [0, 1]$, say that $\succsim$ is 'fully represnted' by S if and only if for every $p, q \in \mathfrak{B}$, (i) $p \succsim q \Leftrightarrow (\forall \mu \in S)(\mu(p) \succsim \mu(q))$, (ii) $p \odot q \Leftrightarrow (\exists \mu_1, \mu_2 \in S)((\mu_1(p) > \mu_1 q) \wedge (m_2(q) > \mu_2(p)))$. If $\succsim$ is fully represented by the set $S = \{\mu\}$, say that S is fully represented by the function μ. It is easy to see that $\succsim$ is opinionated (satisfies Opinionation) if and only if there exists a function μ such that $\succsim$ is fully represented by μ.

Call a function $\mu : \mathfrak{B} \to [0, 1]$ a 'plausibility function' if it satisfies the following two conditions,

$$(PL1)\ \mu(\top) = 1 \text{ and } \mu(\bot) = 0.$$

$$(PL2)\ \text{For any } p, q \in \mathfrak{B}, \text{ if } p \vdash q \text{ then } \mu(p) \leq \mu(q).$$

It is easy to see that if $\succsim$ is a partial preorder over $\mathfrak{B}$, then $\succsim$ will satisfy $A1$ and $A2$ if and only if $\succsim$ is fully representable by a set S of plausibility functions on $\mathfrak{B}$. Thus, an important prospective synchronic coherence requirement for comparative confidence judgements is

($\mathfrak{C}\mathfrak{1}$) $\succsim$ should be fully representable by a set of plausibility functions, or equivalently, $\succsim$ should satisfy $A1$ and $A2$.[10]

Another prospective qualitative rationality constraint on $\succsim$ is

$(A3)$ For any $p, q, r \in \mathfrak{B}$, if $p \vdash q$ and $< q, r >$ are logically incompatible,[11] then

$$q \succ p \Rightarrow q \vee r \succ p \vee r$$

A3 can be thought of as a weak additivity condition. It is best understood in terms of its implications for the representability of $\succsim$ by numerical functions on $\mathfrak{B}$. Call a function $\mu : \mathfrak{B} \to [0, 1]$ a 'mass function' if

[10]In the presence of the Opinionation assumption, $\mathfrak{C}\mathfrak{1}$ is equivalent to $\succsim$ being representable by a single plausibility function.

[11]i.e. $\vdash q \wedge r \equiv \bot$.

$$(M1)\ \mu(\bot) = 0$$

$$(M2)\ \sum_{p \in \mathfrak{B}} \mu(p) = 1$$

Any mass function $\mu : \mathfrak{B} \to [0,1]$ defines a corresponding 'Dempster-Schafer belief function' $b_\mu : \mathfrak{B} \to [0,1]$ defined by $b_\mu(q) = \sum_{\{p \in \mathfrak{B} | p \vdash q\}} \mu(p)$. It is not hard to see that the set of all belief functions on $\mathfrak{B}$ is generally a proper subset of the set of all plausibility functions on $\mathfrak{B}$ and a proper superset of the set of all probability functions on $\mathfrak{B}$, where probability functions are defined as belief functions whose corresponding mass functions assign all of their mass to the possible worlds in $\mathfrak{B}$.[12] It turns out (see Wong *et al.* (1991)) that $\succsim$ satisfies all of $A1$, $A2$ and $A3$ if and only if $\succsim$ is fully representable by a set of Dempster Schafer belief functions. Thus, the second prospective synchronic rationality requirement for comparative confidence judgements is

($\mathfrak{C}2$) $\succsim$ should be fully representable by a set of Dempster-Schafer belief functions, or equivalently, $\succsim$ should satisfy $A1$, $A2$ and $A3$.[13]

Finally, the strictly strongest prospective synchronic rationality constraint that I will consider here is

($\mathfrak{C}3$) $\succsim$ should be fully representable by a set of probability functions.[14]

It is easy to show that a confidence ordering which satisfies $\mathfrak{C}3$ automatically satisfies all the other synchronic constraints listed here.

We are now ready to ask whether and how these synchronic rationality constraints can be justified. More specifically, we can ask 'what goes wrong when an agent's comparative confidence judgements fail to satisfy these representability requirements?'. Generally, there are two ways to justify prospective epistemic norms such as these. Firstly, one can provide a pragmatic justification, i.e. one can show that agents who violate the norm will be *instrumentally irrational* in the sense that they will act in ways which fail to produce the best outcomes by their own lights. An example of such a justification is the synchronic Dutch Book argument for probabilism. Alternatively, one could provide an epistemic

[12]And 'possible worlds' are just the atoms of the Boolean algebra $\mathfrak{B}$, i.e. the maximal consistent conjunctions of sentences of L.

[13]In the presence of the Opinionation assumption, $\mathfrak{C}2$ is equivalent to the requirement that $\succsim$ be fully representable by a single Dempster-Schafer belief function.

[14]The qualitative statements of $\mathfrak{C}3$ (standardly referred to as 'cancellation axioms') are rather technical, so I omit them here (but see e.g. Harrison-Tailor *et al* (2016), Konek (2019), Scott (1964)). In the presence of the Opinionation assumption, $\mathfrak{C}3$ is equivalent to the requirement that $\succsim$ be fully representable by a single probability function.

justification, i.e. one could attempt to show that agents who violate the norm will end up with epistemic attitudes that are in some way defective. Of course, this typically involves appealing to some more fundamental epistemic norms, since the question of whether an agent's epistemic attitudes are defective is itself inherently normative. An example of such an epistemic justification is the accuracy dominance argument for probabilism (see e.g. Joyce (1998), Pettigrew (2016)), which itself appeals to the fundamental norm that an agent should aim to have 'accurate' numerical credences.

For current purposes, I will not take a substantive stand regarding the synchronic norms of comparative confidence judgements. My aim will rather be to systematically catalogue which synchronic norms one needs to assume in order to obtain normative justifications of the diachronic norm given by comparative conditionalisation. Towards this end, I will briefly pause now to recall the two most prominent extant normative justifications of synchronic norms from the literature.

Firstly, Icard (2016) has given a pragmatic 'money-pump' style argument in favour of $\mathfrak{C}3$.[15] Since $\mathfrak{C}3$ is strictly stronger than all the other synchronic norms presented above, this argument, if successful, can be seen as providing a pragmatic justification of all the synchronic rationality constraints considered here. Secondly, Fitelson and McCarthy (unpublished) have developed a formal framework for assessing the accuracy of comparative confidence orderings.[16] This framework yields epistemic accuracy dominance arguments in support of $\mathfrak{C}1$ and $\mathfrak{C}2$. Importantly, these arguments do not extend to a justifications of $\mathfrak{C}3$, which suggests that $\mathfrak{C}2$ may be the strongest interesting norm that can be justified by accuracy dominance arguments in Fitelson and McCarthy's framework. It should also be noted that Fitelson and McCarthy's framework relies on both the Regularity and Opinionation assumptions described above (more on this in Section 6).[17]

3 Comparative Conditionalisation

We are now ready to address the central question of this article: how should a rational agent revise their comparative confidence judgements after learning the truth of some evidential proposition e? Before going further, it is worth spelling out a couple of important background assumptions.

Firstly, I assume here that the evidential proposition e is learned *with certainty*. Thus, the kind of learning I am interested in is the same as that described by standard Bayesian conditionalisation, where

[15] Actually, Icard's argument aims to justify the strictly stronger condition that $\succsim$ should always be fully representable by a single probability function.

[16] This framework is introduced and explicated in section 6, where it is used to evaluate the possibility of an expected inaccuracy argument for comparative conditionalisation.

[17] Again, this means that the accuracy dominance argument for $\mathfrak{C}2$ actually aims to justify the condition that $\succsim$ should always be fully representable by a single Dempster Schafer belief function.

the agent assigns a posterior probability of 1 to the learned proposition. In the context of comparative confidence judgements, the analogous requirement is that after the learning experience, the agent makes the judgement $e \sim \top$, i.e. that they become equally confident in the truth of the learned proposition and the tautology.

Secondly, I assume that upon learning e, the agent needs to reorganise their comparative confidence judgements in a way that (i) ensures that they become certain in the truth of e, and (ii) defines a confidence ordering that preserves all the relevant synchronic rationality requirements that were satisfied by their initial ordering. So, for example, if we assume $\mathfrak{C}1, \mathfrak{C}2$ and $\mathfrak{C}3$ as synchronic rationality requirements and the agent's initial ordering satisfies all these requirements, then their ordering should still satisfy those requirements after they have revised their comparative confidence judgements to accommodate the new evidence. Whatever the synchronic rationality norms are, learning new evidence should not lead one to violate them.

It is clear that there are generally many ways that an agent can revise their confidence orderings whilst satisfying these basic requirements (for any fixed specification of the synchronic norms). How to choose between them? It is instructive here to take inspiration from a key structural property of Bayesian conditionalisation. Specifically, given a probability distribution P, let $\succsim_P$ be the confidence ordering defined by $q \succ_P p$ if and only if $P(q) > P(p)$ and $p \sim_P q$ if and only if $P(p) = P(q)$. By definition, P fully represents $\succsim_P$, and we can think of $\succsim_P$ as encoding the comparative confidence judgements of a Bayesian agent whose credal state is given by the probability function P. Now, we can ask 'what is the relationship between $\succsim_P$ and $\succsim_{P(-|e)}$?', where $P(-|e)$ is the probability function obtained by conditionalising P on e. Less formally: 'how does conditionalising on e change the comparative confidence judgments implicit in P?'. Happily, this question has a simple answer:

$$\begin{aligned} q \succsim_{P(-|e)} p &\Leftrightarrow P(q|e) \geq P(p|e) \\ &\Leftrightarrow P(q|e)P(e) \geq P(p|e)P(e) \\ &\Leftrightarrow P(e \wedge q) \geq P(e \wedge p) \\ &\Leftrightarrow e \wedge q \succsim_P e \wedge p \end{aligned}$$

Thus, if we let $\succsim_e$ denote the ordering that results from revising the initial ordering $\succsim$ after learning e, a Bayesian agent will always revise their confidence orderings according to the rule

(CC:) $q \succ_e p \Leftrightarrow e \wedge q \succ e \wedge p$, and $q \sim_e p \Leftrightarrow e \wedge q \sim e \wedge p$

Where 'CC' stands for 'comparative conditionalisation'.[18] The question now is whether there is anything special about CC as opposed to other revision rules for comparative confidence judgements. One might be tempted here to simply invoke the observation that there are numerous philosophical justifications for viewing Bayesian conditionalisation as the uniquely rational rule for updating numerical credences, and to conclude that the revision rule defined by Bayesian conditionalisation must therefore be the correct one. However, this kind of justification is clearly flawed. For, it assumes at the outset that an agent's comparative confidence judgements are defined by a specific credal state, and that the way in which an agent revises those judgements will be entirely determined by the rule they use to update that credal state. But, as I noted in the introduction, there is a significant minority of authors who contend that comparative confidence judgements are philosophically and psychologically more fundamental than assignments of numerical credence, and so will reject the implicit assumption that an agent's comparative confidence judgments are always determined by some specific credal state. It may be that the content of an agent's epistemic state is exhausted by their confidence ordering, and that they simply have no well defined credal state.[19] Again, it's at least coherent to conceive of such an agent. And in this context, rejecting CC in favour of another revision rule does not bring one into conflict with Bayesian conditionalisation. For, the way in which an agent revises their comparative confidence judgments will have *no* implications regarding the way in which they update their credences if they have no well defined credences in the first place.[20]

If one hopes to justify CC, then one must do so within the context of the epistemology of comparative confidence judgements. My aim in the rest of this paper is to explore the possibility of systematically justifying CC within the context of a comparativist epistemology. But before doing so, it is worth emphasising the intuitive rationality of CC as opposed to alternative revision rules. Towards this end, consider the following example:

> Mufasa is sitting in a soundproof room with no windows, and he has no idea what the weather is like outside. The room is equipped with a speaker which will occasionally announce some partial information about the weather outside. Based on past experience, he judges that it

[18]Note that CC is also implicitly assumed by alternative quantitative models of inductive learning, including for example Rank conditionalisation (Spohn, 1988) and Possibility Conditionalisation (Zadeh, 1978). Interestingly, it turns out that CC does *not* cohere with Dempster's rule for updating Dempster Schafer belief functions. However, it is easy to observe that it does cohere with Fagin and Halpern's (1991) rule for updating belief functions. Thus, the arguments presented here in favour of CC have significant implications for evaluating competing updating rules for Dempster Schafer belief functions.

[19]Note that this is true even if their confidence ordering is fully representable by a probability function, since there will generally be infinitely many different probability functions that can be used to represent the ordering.

[20]It is worth stressing again here that my aim in this paper is not to defend the position that we should conceive of an agent's epistemic state purely in terms of comparative confidence judgements. Rather, I start from the assumption that there are at least some situations in which such a conception is desirable, and then address the question of how agents in situations of this sort should revise their epistemic states over time in light of this assumption.

is more likely to be raining and thundering outside than it is to be sunny and thundering outside, i.e. he makes the judgements $(r \wedge t) \succ (s \wedge t)$. The speaker then announces that it's thundering outside. Mufasa subsequently revises his comparative confidence judgements in a way that leads him to judge that it is more likely to be sunny outside than it is to be rainy outside, i.e. he makes the judgements $s \succ_t r$.

I take it that there is something intuitively bizarre about the dynamics of Mufasa's confidence judgements here. The question is whether this bizarreness is indicative of diachronic irationality. As a first step, I will now survey some basic formal properties of the CC updating rule.

3.1 Formal Properties of CC

I require that any revision procedure for comparative confidence judgements should satisfy two key properties: (i) after learning e, the agent should make the judgment $e \sim_e \top$, and (ii) the posterior ordering $\succsim_e$ should preserve all relevant synchronic rationality constraints (for some fixed specification of the synchronic rationality constraints). I begin by noting that CC trivially satisfies (i), since

$e \sim_e \top \Leftrightarrow (e \wedge e) \sim (e \wedge \top) \Leftrightarrow e \sim e$, and all partial preorders on $\mathfrak{B}$ satisfy $e \sim e$ for all $e \in \mathfrak{B}$.[21]

Next, we need to establish (ii). A first step is achieved with the following proposition (all proofs in appendix).

Proposition 1 *Let $\succsim$ satisfy $\mathfrak{C}1$ Then $\succsim_e$ satisfies $\mathfrak{C}1$ if and only if $e > \bot$.*

In what follows, I always assume that the agent is initially strictly more confident in the learned evidential proposition e than they are in the contradiction, i.e. $e \succ \bot$.[22] Given this assumption, we can show that CC preserves all the synchronic rationality constraints described in section 2.[23]

Proposition 2 *If $\succsim$ satisfies $\mathfrak{C}2$, then $\succsim_e$ satisfies $\mathfrak{C}2$.*

Proposition 3 *If $\succsim$ satisfies $\mathfrak{C}3$, then $\succsim_e$ satisfies $\mathfrak{C}3$.*

Thus, we know that regardless of which synchronic rationality constraints one imposes on the prior confidence ordering, revising by CC will never lead an agent to replace a coherent confidence ordering

[21] Similarly, $\neg e \sim_e \bot \Leftrightarrow (e \wedge \neg e) \sim (e \wedge \bot) \Leftrightarrow (e \wedge \neg e) \sim \bot$, which again is satisfied by all partial preorders on $\mathfrak{B}$.

[22] This assumption is of course reminiscent of the fact that a Bayesian agent can never condition on a probability 0 event. Critics of Bayesian epistemology typically take this feature to be problematic and unmotivated. I don't address this issue here, but it is certainly worth noting that this aspect of Bayesian inference generalises so naturally to the comparative setting (and so can't be straightforwardly attributed to the ratio definition of conditional probabilities, as is often suggested).

[23] In fact, it is also straightforward to show that CC preserves another synchronic rationality constraint from the literature that is strictly stronger than $\mathfrak{C}2$ but strictly weaker than $\mathfrak{C}3$, namely that $\succsim$ should be a 'comparative probability relation' (see e.g. Kraft *et al.* (1959)).

with an incoherent one. So CC satisfies the second condition on revision procedures for comparative confidence judgements, regardless of which of these synchronic coherence constraints one adopts as epistemic norms.[24]

3.2 Opinionation Failures and Conditional Judgement

I turn now to briefly describing two important points regarding the scope of CC's applicability. Firstly, it is important to note that the definition of CC given above does *not* assume that the prior ordering satisfies Opinionation, even though it was inspired by standard Bayesian conditioning, an updating rule that does implicitly assume Opinionation.[25] To see this, note that CC can be equivalently formulated as follows (where $p \odot_e q$ denotes the case in which the agent makes no judgement regarding the pair (p, q) after learning e).

(CC*:) $q \succ_e p \Leftrightarrow e \wedge q \succ e \wedge p$, $q \sim_e p \Leftrightarrow e \wedge q \sim e \wedge p$, and $q \odot_e p \Leftrightarrow e \wedge q \odot e \wedge p$

The third biconditional (absent from the initial definition) is of course implied by the first two, and is irrelevant when Opinionation is assumed. At this stage, it is instructive to consider the theory of imprecise credences, where it is often assumed that an agent's credences are represented by a *set* of precise probabilistic credence functions, referred to as the agent's 'representor' (see i.e. Joyce (2010), Weatherson (2007)). On this view, an agent's comparative confidence ordering can be derived through the following supervaluationist semantics. Firstly, the agent makes the judgement $p \succsim q$ if and only if every function in their representor assigns p a credence which is at least as high as what it assigns to q. Secondly, if there are two functions P_1, P_2 in the agent's representor such that $P_1(p) > P_1(q)$ and $P_2(q) > P_2(p)$, then the agent makes no comparative confidence judgement regarding p and q, i.e. their confidence ordering satisfies $p \odot q$. By definition, the ordering identified by this semantics always satisfies $\mathfrak{C}_3$. Typically, these imprecise models assume that upon learning a proposition e, a rational agent will replace their prior representor $\mathcal{P}$ by the set $\mathcal{P}(-|e) = \{P(-|e)|P \in \mathcal{P}\}$, i.e. that they will simply condition every function in their prior representor on e and take the set of updated

[24] It is also worth noting that, perhaps unsurprisingly, CC shares many of the key structural properties of Bayesian conditionalisation. For example, CC defines a commutative revision procedure, i.e. the order in which the agent receives novel evidence makes no difference to the comparative confidence judgements that they end up with at the end of the learning process. To see this, let $\succsim_{e_1,e_2}$ be the result of revising $\succsim$ sequentially by e_1 and then e_2. Then

$$p \succ_{e_1,e_2} q \Leftrightarrow e_2 \wedge p \succ_{e_1} e_2 \wedge q \Leftrightarrow e_1 \wedge e_2 \wedge p \succ e_1 \wedge e_2 \wedge q \Leftrightarrow p \succ_{e_1 \wedge e_2} q.$$

The commutativity of CC is of course of fundamental importance, since it ensures that there is always a well defined and intuitively rational way to iterate the revision procedure in sequential learning scenarios.

[25] Opinionation's status as an epistemic norm is contested by many authors (see e.g. Forest (1989), Keynes (1921), Eva (2019))

functions as their new representor. Now, it's easy to see that if $\succsim$ is fully represented by the agent's representor, then the posterior ordering obtained by applying CC* will always be fully represented by the agent's posterior representor (see Proposition 4). So just as CC coheres perfectly with standard conditionalisation, CC* coheres perfectly with its imprecise counterpart. Thus, (since CC and CC* are equivalent) CC can be straightforwardly and naturally applied to the non-opinionated setting, and is in fact directly entailed by the most influential extant attempt to codify the norms of inductive inference in the absence of the Opinionation assumption.

The second important point to note regarding the scope of CC's applicability concerns the rule's relation to supposition and conditional judgement. Here, it is significant that the definition of standard Bayesian conditionalisation relies on the availability of *conditional degrees* of belief. In order to calculate my new credence in q after conditionalising on p, I need to know my prior conditional degree of credence in q given p, $P(q|p)$, which is standardly interpreted as representing my credence in q *under the (indicative) supposition* that p is true. In the comparative context, Koopman (1940) forwarded a set of axioms whose satisfaction allowed for the definition of an analogous notion of *comparative conditional confidence*.[26] It is significant that the definition of CC does not involve reference to any such notion. The rule can be straightforwardly and intuitively applied without any appeal to representations of conditional or suppositional judgement. This suggests that the close relationship between learning, supposition and conditional judgement that is familiar from Bayesian epistemology is likely to be fundamentally different in the comparative setting.

Enough with the preliminaries. As I noted above, CC looks like an intuitively compelling approach to revising comparative confidence judgements on the basis of novel evidence. Indeed, there is something intuitively inconsistent about agents who revise their judgements in a way that conflicts with CC. But I have not yet given any definite argument in support of the claim that agents who revise their confidence orderings by a rule other than CC are irrational. In what follows, I explore the most salient avenues for the pursuit of such an argument.

4 An Argument From Evidential Relevance

The first argument I will present in support of CC is evidentialist, in the sense that it relies on epistemic norms that stipulate how an agent's doxastic attitudes should relate to the evidence they have at their disposal, and assumes that agents should always aim to base their judgements on relevant evidence.

[26]Importantly, Koopman's axioms encode a substantive array of synchronic norms that are not assumed here.

I begin with the following basic norm, which makes a compelling stipulation about the relationship between an agent's evidence on the one hand and their comparative confidence judgements on the other.

Evidential Norm (EN): An agent A should make the initial judgement $q \succsim p$ if and only if they would retain that judgement upon learning (only) the truth of a proposition e that is entailed by both p and q. Formally, for any $p, q, e \in \mathfrak{B}$ with $p \vdash e$, $q \vdash e$

$$q \succsim p \Leftrightarrow q \succsim_e p$$

To illustrate the intuition behind EN, consider the following example. Let p and q be the propositions 'Alice is in Falmouth' and 'Alice is in Redruth', respectively, both of which entail the proposition $e =$ 'Alice is in Cornwall'.[27] Suppose that we are initially at least as confident that Alice is in Falmouth as we are that she is in Redruth. If we subsequently learn only that Alice is in Cornwall, then we have not learned anything about *where* in Cornwall she is. If we were now to change (or simply abandon) our comparative confidence judgement regarding whether she is more likely to be in Falmouth or Redruth, we would, by definition, be changing our judgements in a way that is not warranted by any relevant evidence. So if we assume that one should only change one's judgements when one has relevant evidence that explicitly warrants the change, then violating EN in this way will never be permissible. Generalising, learning that the actual world $w_@$ is in some region e of possible world space does not give one any evidence concerning *where* in e $w_@$ is. In particular, it does not say anything about whether $w_@$ is more likely to be in any subregion p of e than it is to be in any other subregion q of e. So any change in how one compares the plausibility of different subregions of e upon learning only that e is true would be arbitrary and evidentially unjustified. And such changes are exactly what is prohibited by EN. Importantly, we have the following result.

Proposition 4 *Assuming the synchronic coherence constraint $\mathfrak{C}2$, CC is the only updating rule for comparative confidence judgements that satisfies EN in full generality.*

Proposition 4 establishes a clear sense in which agents who want their judgements to be guided by evidence should always abide by CC. If one deviates from CC, then one is committed to sometimes changing one's judgements in the absence of any relevant evidence that actually warrants the change. And this result relies only on the synchronic constraint $\mathfrak{C}2$, which does *not* entail either Opinionation or any form of probabilistic representability.

[27] Redruth and Falmouth are both towns in Cornwall.

At this stage, it is instructive to compare the evidentialist argument for CC outlined above with an analogous argument that is often given in favour of Bayesian conditionalisation. Specifically, it is often observed that Bayesian conditionalisation is the unique updating rule for probabilistic credences with the properties that (i) $P^*(e) = 1$, and (ii) $\frac{P^*(p)}{P^*(q)} = \frac{P(e \wedge p)}{P(e \wedge q)}$, $\forall p, q, e \in \mathfrak{B}$ (where P^* denotes the posterior credence function obtained after learning e). In virtue of property (ii), many authors refer to Bayesian conditionalisation as the unique rule that 'preserves probability ratios'. And there is an intuitive sense in which preserving probability ratios captures the requirement that agents should not change their credences in ways that are not licensed by the evidence. To see this, note that failing to preserve probability ratios means (assuming that P^* is probabilistic) that there exist p, q, e such that $\frac{P^*(e \wedge p)}{P^*(e \wedge q)} \neq \frac{P(e \wedge p)}{P(e \wedge q)}$. Again, since learning that the actual world $w_@$ is in e doesn't tell us anything about where in e $w_@$ is likely to lie, changing the ratio $\frac{P(e \wedge p)}{P(e \wedge q)}$ after learning only e seems unwarranted. If the ratio gets bigger, we seem to be favouring $e \wedge p$ in a way that is not warranted by the evidence. If it gets smaller, we are likewise favouring $e \wedge q$ in a way that is not warranted by the evidence. So like CC, Bayesian conditionalisation is also supported by an intuitively compelling argument from evidential relevance.

However, upon closer inspection, it is easy to see that the evidentialist argument for CC is significantly stronger than the analogous argument for Bayesian conditionalisation. Note first that the evidential arguments for CC and Bayesian conditionalsiation both start from the premise that upon learning e, we should not favour any subregions of e when we update our epistemic state, since the evidence tells us nothing about where in e the actual world is. Specifically, for any p, q, we should not favour $e \wedge p$ against $e \wedge q$ or vice-versa when we learn e. In the comparative setting, this requirement has an obvious unique interpretation, namely that learning e shouldn't change the comparative confidence judgement we initially made about the pair $(e \wedge p, e \wedge q)$. But in the numerical credence setting, it's not at all obvious that there is a single correct interpretation of the requirement that neither $e \wedge p$ nor $e \wedge q$ should be favoured against its counterpart. When we interpret this 'not favouring' requirement in terms of preserving the ratio $\frac{P(e \wedge p)}{P(e \wedge q)}$, we obtain an argument for conditionalisation. But there are other equally natural interpretations of the requirement that do *not* lead to Bayesian conditionalisation. To see this, consider the following example, where $\mathfrak{B}$ is the algebra generated by the three worlds w_1, w_2, w_3. Let $P(w_1) = \frac{1}{2}, P(w_2) = \frac{1}{3}, P(w_3) = \frac{1}{6}$. Upon learning $\neg w_1$, we should not favour either w_2 or w_3, since both entail the evidence. Bayesian conditionalising respects this constraint by preserving the fact that w_2 is viewed as twice as probable as w_3. Specifically, it yields the new posterior function $P^*(w_1|\neg w_1) = 0, P^*(w_2|\neg w_1) = \frac{2}{3}, P^*(w_3|\neg w_1) = \frac{1}{3}$. But there is also a clear sense in which this response to the evidence *does* favour w_2 over w_3, since the increase in w_2's probability ($\frac{1}{3}$) is greater than the increase in w_3's probability ($\frac{1}{6}$). And it seems perfectly reasonable to interpret the requirement that neither w_2 nor w_3 should be favoured upon learning $\neg w_1$ as requiring that they should both increase in probability to the same degree. This interpretation identifies the following posterior function as the correct response to the evidence, $P^*(w_1) = 0, P^*(w_2) = \frac{7}{12}, P^*(w_3) = \frac{5}{12}$. Importantly, this response to the evidence yields the same posterior confidence ordering over $\mathfrak{B}$ as Bayesian conditionalisation does, and therefore coheres perfectly with CC and satisfies EN. So the simple evidentialist requirement that upon learning e, one should not favour any sub region of e over any other is enough to justify CC, but it is *not* enough to justify Bayesian conditionalisation. For, in the comparative setting, this requirement has a single obvious interpretation (EN). In the context of numerical credences, in can be plausibly interpreted in multiple ways, some of which single out Bayesian conditionalisation as a privileged updating rule, and some of which are in direct conflict with Bayesian conditionalisation (despite yielding the same posterior confidence orderings as Bayesian conditionalisation).

Of course, the preceding analysis does not show that the standard evidentialist arguments for Bayesian conditionalisation are fundamentally flawed in any sense. What it shows is that these arguments require significantly stronger premises than the evidentialist justification for CC presented above, which relies only on the premise that upon learning only e, one should not favour any subregion of e over any other. To the extent that an argument's strength is inversely proportional to the strength of its premises, this shows that the evidentialist justification of CC is meaningfully stronger than analogous justifications of Bayesian conditionalisation. It should also be noted that extant evidentialist justifications for Bayesian conditionalisation that are framed in terms of e.g. entropy maximisation (Williams (1980), Skyrms (1985)) rely crucially on the assumption that a rational agent's credences should always be probabilistic. So if one really wants to derive Bayesian conditionalisation from evidentialist norms alone, then one probably needs to begin with an evidentialist justification of probabilism (the thesis that rational credence is always probabilistic). But there is a significant minority of authors who argue that aligning one's credences with the available evidence sometimes *precludes* the possibility of probabilistic credences altogether (see e.g. Schafer (1976), Spohn (2012)). In this context, it is also salient to note that the only synchronic norm presupposed by the evidentialist justification of CC given here is that a rational agent's comparative confidence judgements should always be representable by a set of Dempster Schafer belief functions. This requirement is compatible both with failures of probabilistic representability and failures of Opinionation. Thus, the evidentialist justification for CC given here is also noteworthy insofar as it dispenses with many of the controversial synchronic norms that are presupposed by extant evidentialist arguments for Bayesian conditionalisation.

5 A Comparative Diachronic Dutch Book Argument

Another of the most influential normative justifications for standard Bayesian conditionalisation is the diachronic Dutch book argument, first reported in Teller (1973). The diachronic Dutch book argument purports to provide a purely pragmatic justification of conditionalisation, in the sense that it (allegedly) demonstrates that all violations of conditionalisation are liable to yield behaviours which are instrumentally irrational. Specifically, diachronic Dutch book arguments aim to show that agents who update their credences via a strategy other than Bayesian conditionalisation can be 'Dutch booked' by bookies who exploit the agent's credences to trade them a series of bets which generate a sure loss. In this section, I'll present a natural generalisation of the diachronic Dutch book argument to the comparative setting (§4.1), catalogue which assumptions are needed to get the argument going, and evaluate the extent to which this generalised argument constitutes a genuine normative vindication of CC (§4.2). Note that in the remainder of the paper, $\succsim_e$ will always denote the posterior ordering produced by CC, and alternative posterior ordering will always be denoted by $\succsim^*$.

5.1 Comparative Betting Principles

The first step in constructing the generalised argument is to identify principles linking an agent's comparative confidence judgements to their betting behaviour. In most presentations of the standard diachronic Dutch book argument for Bayesian conditionalisation, it is simply assumed that agents accept all and only bets wth positive expected utility (see i.e. Pettigrew (forthcoming), Lewis (1999)). In the absence of numerical credences, we need more general principles that do not rely on probabilistic expectations for their articulation. To facilitate the identification of such principles, I will assume Opinionation for the remainder of this section, and postpone the generalisation of the Dutch book argument for CC to the non opinionated setting to future work. Once Opinionation is assumed, we will need one betting principle for equal confidence judgements of the form $p \sim q$, and one for judgements of unequal confidence, of the form $p \succ q$. Before presenting the specific principles to be used in the Dutch book argument for CC, consider first the following intuitively compelling principle (which concerns judgements of unequal confidence).

Comparative Betting Principle 1 (CBP1): Let A be an agent who makes the comparative confidence judgement $p \succ q$. Then for any price α, it is rational for A to pay α to buy a bet with the following payoff matrix.

Outcome	**Payoff**
$p \wedge \neg q$	2α
$\neg p \wedge q$	0
$p \equiv q$	α

To get a feeling for the content of CBP1, suppose first that A's comparative confidence judgements are fully represented by some probability function P, and that A makes the judgement $p \succ q$. Then it's easy to see that the bet outlined above will always have positive expected utility according to P. So if we assume that agents should always accept bets which have positive expected utility according to all probabilistic representations of their comparative confidence judgements, then CBP1 is entailed as a practical norm whenever $\succsim$ is probabilistically representable (whenever $\mathfrak{C}3$ is assumed as a synchronic norm). Importantly though, CBP1 is *far* weaker than the joint assumptions of $\mathfrak{C}3$ and the requirement that agents should accept any prospective bet that has positive expected utility according to every probabilistic representation of their confidence ordering. It is intuitively compelling even in the absence of these assumptions, and a Dutch book argument based on something like CBP1 is therefore stronger than an analogous argument based on the requirement that agents accept all and only bets with positive expected utility, in so far as CBP1 is strictly weaker and more generally applicable than the latter requirement. Unfortunately though, CBP1 on its own isn't quite enough to get the Dutch book argument going. We require the following, slightly more controversial, principle.

Comparative Betting Principle 2 (CBP2): Let A be a rational agent who makes the comparative confidence judgment $p \succ q$, and let $\alpha > 0$ be any price. Then there will exist β_1, β_2 such that (i) $\alpha > \beta_1, \beta_2$, (ii) $\beta_1, \beta_2 > 0$, and (iii) A will pay α to buy a bet with the following structure.

outcome	payoff
$p \wedge \neg q$	$2\alpha - \beta_1$
$\neg p \wedge q$	0
$p \equiv q$	$\alpha - \beta_2$

CBP2 is strictly stronger than CBP1 in so far as it positively identifies a broader range of bets that practically rational agents are compelled to buy. Specifically, it requires that when A makes the judgement $p \succ q$, there should be *some* bet which is acceptable to the agent, and which has worse expected utility than the corresponding bet specified by CBP1 in the sense that the payoffs in the $p \wedge \neg q$ and $p \equiv q$ cells are smaller than the corresponding payoffs in the CBP1 bet (by a degree of β_1 and β_2, respectively). The difference between this bet and the bet in CBP1 can be arbitrarily small. The idea is that if one is willing to accept the CBP1 bet, then one should also be willing to accept *some* bet with an infinitesimally less appealing payoff structure. Since the bet in CBP1 is equivalent to the bet in CBP2 in the limit where β_1, β_2 approach zero, and β_1, β_2 can be arbitrarily small, there will always be some values of β_1, β_2 for which the bet in CBP2 becomes acceptable - or so the story goes. Again, it is easy to see that in the special case where $\succsim$ is representable by some probability function P, there will always exist $\beta_1, \beta_2 > 0$ for which the bet in CBP2 has positive expected utility according to P. So like CBP1, CBP2 is directly entailed by the joint assumptions of $\mathfrak{C}3$ and the requirement that agents should accept any prospective bet that has positive expected utility according to every probabilistic representation of their confidence ordering. Like CBP1, CBP2 is in fact far weaker than these assumptions, which means that a Dutch book argument based on CBP2 is still assuming much less than about the norms of rational decision making than standard Dutch book arguments for Bayesian conditionalisation.

Finally, before presenting the argument, I will also need a principle that connects judgements of equal confidence to rational betting behaviour, namely

Comparative Betting Principle 3 (CBP3): Let A be a rational agent who makes the comparative confidence judgment $p \sim q$, and let $\alpha > 0$ be any price. Then for any β_1, β_2 satisfying $\beta_1 + \beta_2 > 2\alpha$,

A will buy bets of the following form.

outcome	**payoff**
$p \wedge \neg q$	β_1
$\neg p \wedge q$	β_2
$p \equiv q$	α

As with CBP1/2, it is worth observing that whenever $\succsim$ is probabilistically representable, all representations of $\succsim$ will expect the bet in CBP3 to have positive utility whenever $p \sim q$ obtains. So again, CBP3 is entailed by the conjunction of $\mathfrak{C}_3$ with the requirement that agents should buy all and only those bets that have positive expected utility on every probabilistic representation of their confidence ordering. But it is also far weaker than those assumptions, and is intuitively compelling and naturally applicable in settings where agents have no numerical credences or probabilistically representable confidence orderings. In the following, I will assume CBP2 and CBP3 as constraints governing the betting behaviour of rational agents.

5.2 The Argument

I am now ready to present the main assumptions which, combined with CBP2 and CBP3, yield a diachronic Dutch book argument for CC. First, note that, as in the standard Dutch book arguments for Bayesian conditionalisation, I assume that (i) the agent A has a fixed updating strategy, i.e. a plan for how they will revise their comparative confidence judgements given any possible piece of evidence e, and (ii) that this updating strategy is known to both A and the bookie.[28] Secondly, I assume that the updating strategy is such that for some proposition e, learning e would lead A to revise their comparative confidence judgements in a way that conflicts with CC. This implies that there exist propositions $x, y \in \mathfrak{B}$ such that one of the following four cases obtains, where $\succsim$ denotes the agent's prior ordering and $\succsim^*$ denotes their posterior ordering after learning e

[28]Both of these assumptions have been the targets for recent critiques of Dutch book arguments for diachronic rationality norms (see i.e. Pettigrew (forthcoming)).

(i) $(e \wedge p) \succ (e \wedge q)$, $q \succ^* p$.

(ii) $(e \wedge p) \succ (e \wedge q)$, $q \sim^* p$.

(iii) $(e \wedge p) \sim (e \wedge q)$, $q \succ^* p$.

(iv) $(e \wedge p) \sim (e \wedge q)$, $p \succ^* q$.

These cases enumerate the four possible types of CC violation. The argument aims to demonstrate that violations of type (i) and (ii) are liable to yield the acceptance of series of bets which generate a sure loss. Analogous arguments can also be made to illustrate the practical irrationality of violation types (iii) and (iv), but these would require significantly longer and more technical argumentation, which I defer to future work in the interest of concision. For now, I am content to show that violations of types (i) and (ii) generate sure losses, and hence that any 'major' departures from CC are instrumentally irrational.

Thirdly, I also assume (importantly) that the bookie knows nothing that the agent A doesn't know, and hence that the truth values of the propositions e, p and q are unknown to the bookie. Finally, I assume (as is standard in Dutch book arguments) that there exists some future time t_e at which the truth value of e will be simultaneously revealed to both the agent and the bookie.[29]

The argument runs as follows, where we are concerned with showing that CC violations of types (i) and (ii) can lead to the acceptance of sequences of bets that generate a sure loss.

1: Suppose that A makes the judgement $(e \wedge p) \succ (e \wedge q)$ at time t_0. Let $\alpha > 0$ be any price. By CBP2, there will exist numbers $\beta_1, \beta_2 > 0$ for which A is willing to pay α for the following bet.

outcome	**payoff**
$e \wedge p \wedge \neg q$	$2\alpha - \beta_1$
$e \wedge \neg p \wedge q$	0
$(e \wedge p) \equiv (e \wedge q)$	$\alpha - \beta_2$

So at time t_0, the bookie sells this bet to A for α.

[29]Where it is also assumed that A does not learn anything else in the intervening time before t_e.

2: The bookie now waits for time t_e, at which point the truth value of e is revealed to A and the bookie. If e is false, then A makes a loss of $\beta_2 > 0$ on the first bet, and the bookie can quit while she's ahead. If e is true, then A will follow their updating strategy to revise their prior confidence judgements. By assumption, they will now make one of the two judgements $q \succ^* p$ or $p \sim^* q$. Suppose first that they make the judgement $q \succ^* p$. By CBP2 there will exist β_3, β_4 for which they are now willing to buy the following bet for price α.

outcome	**payoff**
$\neg p \wedge q$	$2\alpha - \beta_3$
$p \wedge \neg q$	0
$p \equiv q$	$\alpha - \beta_4$

So at time t_e, the bookie sells this bet to A for α. There are now three possibilities. If $p \wedge \neg q$ is true, then the agent will have won the first bet and lost the second, yielding an overall loss of β_1. If $\neg p \wedge q$ is true, the agent will have lost the first bet and won the second, yielding an overall loss of β_3. If $p \equiv q$ is true, the agent will accrue an overall loss of $\beta_2 + \beta_4$. So in all eventualities, A will lose money.

It remains to address the case in which A makes the judgement $p \sim^* q$ at t_e. In this case, CBP3 entails that A will be willing to buy the following bet for the same price α they paid for their first bet, where β_3, β_4 are such that $\beta_3 + \beta_4 > 2\alpha$, $\beta_4 < 2\alpha$ and $\beta_1 > \beta_3$.[30]

outcome	**payoff**
$p \wedge \neg q$	β_3
$\neg p \wedge q$	β_4
$p \equiv q$	α

[30]It's easy to check that there will always exist suitable β_3, β_4. For example, set $\beta_4 = 2\alpha - \frac{\beta_1}{2}$, $\beta_3 = \frac{3\beta_1}{4}$.

So at time t_e, the bookie will sell this bet to the agent for the price α. There are now three possibilities. If $p \wedge \neg q$ is true, then the agent will have won the first bet and 'lost' the second, yielding an overall loss of $\beta_1 - \beta_3 > 0$. If $\neg p \wedge q$ is true, the agent will have lost the first bet and 'won' the second, yielding an overall loss of $2\alpha - \beta_4 > 0$. If $p \equiv q$ is true, the agent will accrue an overall loss of β_2. So in all eventualities, A will lose money.

In sum then, we have considered two of the four ways that the agent can diverge from CC at time t_e. Either they make the judgement $p \sim^* q$ or they make the judgement $q \succ^* p$. Either deviation will leave A susceptible to buying a sequence of bets which will yield a sure loss.

5.3 Discussion

To what extent does the preceding argument constitute a genuine pragmatic vindication of CC?[31] First of all, it is worth reiterating that the argument only counts against two specific types of CC violation, namely the types that occur when an agent makes the prior judgement $e \wedge p \succ e \wedge q$ but fails to make the posterior judgement $p \succ^* q$. It does not apply to the kinds of violations that occur when the agent makes the prior judgement $e \wedge p \sim e \wedge q$ but fails to make the posterior judgement $p \sim^* q$. To make this precise, say that a confidence ordering $\succsim_2$ is an 'extension' of an ordering $\succsim_1$ if and only if $p \succ_1 q$ implies $p \succ_2 q$. Note that $\succsim_2$'s being an extension of $\succsim_1$ leaves open the possibility that $\succsim_1$ and $\succsim_2$ can differ in the sense that there can exist p, q such that $p \sim_1 q$ and $p \succ_2 q$. The comparative diachronic Dutch book argument presented here purports to show that an agent who learns e and adopts a posterior ordering that is not an extension of the posterior ordering $\succsim_e$ produced by CC will be disposed to accept a series of bets that generate a sure loss. Thus, the argument is not a fully fledged justification of CC as a practical norm, but rather a justification of the rule that one should always update in a way which produces confidence orderings that are extensions of those produced by CC. Informally, the argument shows that while it may be permissible to update in a way that deviates slightly from CC, any 'large violations' of CC lead one away from the path of instrumental rationality.

Now, one could well search for alternative/additional betting principles in order to reinforce the argument and achieve a full justification of CC.[32] But the argument presented above suffices for current purposes. Although it doesn't fully justify the normative status of CC, it strongly constrains the range

[31] I stress that my focus here is on those aspects of the argument that are peculiar to the current comparative formulation. For present purposes, I am happy to ignore those aspects of the argument which are shared with standard probabilistic formulations (i.e. that the bookie knows the agent's updating strategy, that the strategy is deterministic etc). In that sense, I am interested only in evaluating whether the diachronic Dutch book for CC is *as good as or better than* the corresponding arguments for Bayesian conditionalisation, rather than on whether any of those arguments work at all.

[32] In ongoing work, I conduct a more thorough survey of the betting principles that can be used to ground Dutch book arguments for CC and related rules.

of CC violations that are permissible from the perspective of instrumental rationality, and does so by assuming a very weak decision theory for comparative confidence judgements.

Secondly, it should be stressed that much of the heavy lifting in the argument is done by the comparative betting principles CBP2 and CBP3. Once these principles are accepted, the argument follows quite straightforwardly from the same assumptions that are at play in the standard diachronic Dutch book argument for Bayesian conditionalisation. Thus, the argument is only as compelling CBP2 and CBP3. And at first blush, these are both plausible principles. Both cohere naturally with the idea that agents with probabilisitic credences should accept all and only bets with positive expected utility. But of course, this observation alone is not a dialectically acceptable justification, since we are concerned with investigating the prospects of normatively justifying CC in a purely comparative setting, where it is not assumed that agents come equipped with probabilistic credences. At this stage, it is worth emphasising that CBP2 and CBP3 are purely qualitative principles, which do not assume that agents have numerical credences of any kind, and are therefore applicable in a far broader setting than the injunction to buy all and only bets with positive expected utility. For now, I will not devote further efforts to justifying CBP2/CBP3 on a priori grounds, and am content to observe that any decision theory for comparative confidence that entails CBP2 and CBP3 will automatically yield a diachronic Dutch book argument for (a slight weakening of) CC. Since these principles are intuitively compelling and are entailed not only by the theory of probabilistic expected utility, but also by numerous theories of rational decision making for non probabilistic credence (see e.g. Denoeux (2019)), I take this observation to constitute significant evidence in favour of CC.

Finally, I should reiterate that the argument presented in this section is applicable only to the special case in which Opinionation is assumed. Whether the argument generalises to the non opinionated setting will depend crucially on how one takes failures of Opinionation to interact with the norms of rational decision making. This topic is unfortunately beyond the scope of the present enquiry (but see Liu, forthcoming).

6 Comparative Expected Inaccuracy

In the previous sections, I presented one pragmatic argument and one epistemic/evidentialist argument for the normative status of CC. I turn now to exploring the possibility of constructing another purely epistemic justification of CC. Again, the strategy will be to identify an extant justification for Bayesian conditionalisation, and investigate whether it can be transferred to the comparative setting to ground a normative justification for CC. Perhaps the single most influential epistemic justification for Bayesian conditionalisation in the current literature is the argument from expected inaccuracy, first articulated by Graves and Wallace (2006). The basic idea behind the argument is the following.

Let P be the probability distribution that encodes the prior credences of a Bayesian agent A. Upon learning the evidential proposition e, A needs to adopt a new probability function P^* such that $P^*(e) = 1$. But how should they choose amongst the infinitely many functions which satisfy this condition? Greaves and Wallace propose that A should choose the function P^* that satisfies $P^*(e) = 1$ and minimises the quantity

$$Exp(P^*|P) = \sum_{w \in W} P(w)\mathfrak{I}(P^*, w)$$

where W is the set of possible worlds and $\mathfrak{I}$ is the strictly proper scoring that is assumed to encode A's conception of the inaccuracy of a credal state. Intuitively, this quantity is supposed to encode the *expected inaccuracy* of the credal state represented by the probability function P^* according to A's prior credal state P. By minimising this quantity, A will find the function P^* which, by the lights of A's own initial credences, is expected to be the most accurate function that satisfies the new evidential constraint. Greaves and Wallace show that the expected inaccuracy of P^* is uniquely minimised (as a function of P^*) when $P^* = P(-|e)$, i.e. updating by conditionalisation will always produce the posterior probability function with the lowest possible expected inaccuracy. Assuming that rational agents should always attempt to minimise the inaccuracy of their posterior credences, we then reach the conclusion that conditionalisation is the unique rational updating rule for agents with probabilistic credences. To date, this is the most influential epistemic justification of Bayesian conditionalisation as a principle of ideal rationality.

The idea now is to study whether a similar argument can be made for CC, i.e. to check whether CC always leads to those comparative confidence judgements that an agent initially expected to be the most accurate in the comparative context. In order to do this, we need access to a suitable scoring rule for quantifying the inaccuracy of an agent's comparative confidence judgements. Happily, the problem of identifying such a scoring rule has already been addressed by Fitelson and McCarthy (unpublished), whose framework I will now briefly introduce.

6.1 Accuracy and Confidence Orderings

Just as there is an intuitive sense in which an agent's numerical credences can be more or less accurate, there is likewise an intuitive sense in which an agent's comparative confidence judgements can be more less accurate. One might hope that, by making this intuitive notion formally precise, it will be possible to obtain epistemic justifications for norms governing the comparative confidence judgements of rational agents. This is precisely the project taken up by Fitelson and McCarthy (unpublished), who base their formalisation of accuracy for confidence orderings on the following premise,

> [A] confidence ordering is (qualitatively) inaccurate (at w) if and only if it fails to rank all the truths strictly above all the falsehoods (at w). (Fitelson and McCarthy (unpublished): 6)[33]

The idea is that an agent A makes an epistemic mistake (at a world w) when they fail to be strictly more confident in a proposition that is true at w than they are in a proposition that is false at w. There are two ways in which this can happen:[34]

(Case 1) A makes the judgement $[q \succ p]$, but $w \models p \wedge \neg q$.

(Case 2) A makes the judgement $[q \sim p]$, but $w \models \neg(p \equiv q)$.

Intuitively, A's epistemic mistake in Case 1 is worse than their epistemic mistake in Case 2. For, in Case 2, A merely fails to be more confident in a truth than they are in a falsehood, while in Case 1, A is actually more confident in the falsehood than they are in the truth. Fitelson and McCarthy take these observations to suggest that the overall inaccuracy of A's comparative confidence judgements at a world w should be a weighted sum

$$\mathfrak{I}_{\geq}(\succsim, w) = \alpha_1 \cdot M_1(\succsim, w) + \alpha_2 \cdot M_2(\succsim, w)$$

where $\succsim$ is A's confidence ordering, $M_1(\succsim, w)$ and $M_2(\succsim, w)$ are the numbers of Case 1 mistakes and Case 2 mistakes that A makes at w, respectively, and $\alpha_1 > \alpha_2$.[35] At this point, we need to specify the exact strength of this inequality, i.e. we need to specify how much worse a Case 1 mistake is than a case 2 mistake. Fitelson and McCarthy show that there is actually a good argument for requiring that $\alpha_1 = 2\alpha_2$, i.e. that Case 1 mistakes are exactly twice as bad as Case 2 mistakes. To do this, they begin by assuming the Regularity condition introduced in Section 2. They then go on to show that, assuming Regularity, setting $\alpha_1 = 2\alpha_2$ is the only way to ensure that the scoring rule $\mathfrak{I}_{\geq}$ is *evidentially proper*, where an evidentially proper scoring rule is one that guarantees that whenever $\succsim$ is fully representable by a probability function P, $\succsim$ is the unique confidence ordering that minimises expected inaccuracy

[33] A brief clarificatory note is in order here. At the time of the writing of this article, Fitelson and McCarthy's manuscript on scoring rules for comparative confidence judgements remains unpublished. However, all of the material referenced herein is publicly available in draft form at the url https://davidmccarthy.org/wp-content/uploads/2018/01/fitelson-mccarthy2015.pdf.

[34] Note that Fitelson and McCarthy assume Opinionation.

[35] i.e. $M1 = |\{(p, q) \in \mathfrak{B} \times \mathfrak{B} | (q \succ p) \wedge (w \models p \wedge \neg q)\}|$, $M2 = |\{(p, q) \in \mathfrak{B} \times \mathfrak{B} | (q \sim p) \wedge (w \models p \wedge \neg q)\}|$.

according to P. If our scoring rule were not evidentially proper, then there would be orderings that were fully represented by probability functions, but which didn't minimise expected inaccuracy according to those functions.[36] Such orderings would appear to be 'self-undermining' in the sense that their probabilistic representations would expect other orderings to be more accurate.[37] For example, we can imagine a case in which an agent is informed of the relevant chance distribution determining the objective probabilities of the propositions under consideration. If they apply the Principal Principle, they will then adopt credences that match this distribution. These credences will define a corresponding confidence ordering, and it would be peculiar if the agent expected (relevant to their actual credences) another ordering to be more accurate than their own. It should be noted that this justification does not assume that agents are always equipped with probabilistic credences, but rather than *whenever an agent has such credences*, they should expect their own confidence ordering to be maximally accurate.

So the requirement that one's scoring rule be evidentially proper is, at first blush, a natural and persuasive one, and it motivates the following explicit definition,

$$\mathfrak{I}_{\geq}(\succsim, w) = 2 \cdot M_1(\succsim, w) + M_2(\succsim, w)$$

This is the final form of the scoring rule proposed by Fitelson and Mccarthy. It is important to stress that this scoring rule follows directly from two compelling axioms: evidential propriety and the requirement that inaccuracy should be a linear sum of the inaccuracies of the individual judgements encoded by an agent's ordering. As I've already stressed, violations of evidential propriety would seem to commit agents to a problematic form of epistemic modesty, on which they expect their own confidence orderings to be less accurate than some other possible orderings. The linearity assumption is also familiar from standard epistemic utility theory for probabilistic credences, where it is standardly assumed that global inaccuracy should be a weighted sum of the 'local inaccuracies' generated by individual credal judgements (see e.g. Pettgirew (2016), pp39–40 for a discussion and motivation of this assumption).

With their scoring rule in hand, Fitelson and McCarthy employ the following accuracy dominance avoidance norm (familiar from the epistemic utility theoretic justification of probabilism) to provide epistemic justifications for some of the synchronic rationality requirements described in section 2.

[36]Where the expected inaccuracy of an ordering $\succsim$ according to a probability function P is defined in the obvious way, i.e. $Exp_{\geq}(\succsim|P) = \sum_{w \in W} P(w) \cdot \mathfrak{I}_{\geq}(\succsim, w)$.

[37]This is of course similar to the sense in which immodest credences are said to be 'self-undermining' (see e.g. Joyce (1998)).

Weak Accuracy Dominance Avoidance (WADA): $\succsim$ should not be weakly dominated in accuracy, i.e. there should not exist any $\succsim^*$ such that

$$\text{(i) } (\forall w)[\mathfrak{I}_{\geq}(\succsim^*, w) \leq \mathfrak{I}_{\geq}(\succsim, w)]$$

$$\text{(ii) } (\exists w)[\mathfrak{I}_{\geq}(\succsim^*, w) < \mathfrak{I}_{\geq}(\succsim, w)]$$

Intuitively, an ordering $\succsim$ that is weakly dominated by another ordering $\succsim^*$ is guaranteed to be alethically sub-optimal in the sense that it can never be more accurate than $\succsim^*$, but it can be less accurate. Assuming that rational agents aim to have accurate epistemic states, it is clear that they should strive to ensure that their comparative confidence judgements can not be weakly accuracy dominated. Fitelson and McCarthy prove the following results.

Theorem 5 *(1) If $\succsim$ violates $\mathfrak{C}1$, then $\succsim$ is weakly accuracy dominable. (2) If $\succsim$ violates $\mathfrak{C}2$, then $\succsim$ is weakly accuracy dominable. (3) $\succsim$ can violate $\mathfrak{C}3$ without being weakly accuracy dominable.*

Thus, the norm **WADA** alone is sufficient to ground a purely epistemic justification for the requirement that $\succsim$ be fully representable by a Dempster-Schafer belief function. However, it is not sufficient to justify the stronger requirement that $\succsim$ be fully representable by a probability function.

Now, recall that in the presence of the Regularity assumption, Fitelson and McCarthy's scoring rule is entailed by two compelling axioms (linearity and evidential propriety). Our current aim is to study whether or not this scoring rule can be used to articulate an expected inaccuracy argument for CC. But in order to do so, we of course need to drop the Regularity assumption, since CC always produces orderings which violate that assumption. And, as it turns out, dropping the Regularity assumption not only undermines the justification for Fitelson and McCarhty's scoring rule. It actually renders it implausible, in light of the following result.

Proposition 6 *In the absence of the Regularity assumption, $\mathfrak{I}_{\geq}$ is not evidentially proper.*[38]

Far from being entailed by evidential propriety, $\mathfrak{I}_{\geq}$ is actually inconsistent with evidential propriety when Regularity is not assumed. Together with Fitelson and McCarthy's observation that $\mathfrak{I}_{\geq}$ is the only evidentially proper linear scoring rule when Regularity *is* assumed, this entails that there is no generally evidentially proper linear scoring rule for comparative confidence judgements.[39] And this

[38]In fact, the proof of Proposition 6 shows that in the absence of Regularity, it is possible for a probabilistically representable confidence ordering to be weakly dominated by another ordering that is not probabilistically representable, and hence that Theorem 5 relies crucially on the Regularity assumption.

[39]It is also worth drawing attention to another salient problem for Fitelson and McCarthy's scoring rule that arises even when Regularity is assumed, namely that it ignores a third kind of epistemic mistake that should also be taken into account

makes it clear that one can never hope to use a linear scoring rule to ground a plausible expected inaccuracy argument for CC. For, if one uses a linear scoring rule, then one is obliged to countenance violations of evidential propriety, which means that one is happy to adopt confidence orderings that are guaranteed to have sub optimal expected inaccuracy, even in synchronic contexts. But if you don't care whether your confidence ordering has high expected inaccuracy *now*, why should you care about whether the confidence orderings you will have in the future also have high expected inaccuracy (relative to your current judgements)? It would be bizarre for one to happily adopt orderings that one knows to have high expected inaccuracy now, but then to feel compelled to change one's judgements in a way that minimises expected inaccuracy diachronically. One remaining possibility is to try to use non linear scoring rules to ground an expected inaccuracy argument for CC. I will make no general remarks about the prospects for this project, but it is worth noting that the one extant example of a non linear scoring rule for confidence orderings, introduced in Spohn and Raidl (2020), will be of no use to the CC advocate, since it also violates evidential propriety (as Spohn and Raidl acknowledge) and therefore cannot be a suitable formalisation of epistemic utility for those that care about synchronic or diachronic expected inaccuracy.

Overall then, it is clear that existing formalisations of accuracy for comparative confidence orderings will not be able to provide an expected inaccuracy argument for CC. Any such argument will necessarily rely on a radically new approach to quantifying the inaccuracy of comparative confidence judgements.

7 Conclusion and Future Work

Let's recap. In Section 3, I introduced and characterised CC as a rule for updating one's comparative confidence judgements on the basis of novel evidence. In Section 4, I showed that CC follows directly from a fundamental norm regarding the relation between an agent's judgements and their evidence, and demonstrated that this evidentialist argument for CC is more general and in some ways stronger than analogous arguments for Bayesian conditionalisation. In Section 5, I showed that two intuitively compelling and generally applicable decision making norms for agents that make comparative confidence judgements are sufficient to ground a diachronic Dutch book argument for (a slight weakening of) CC. Again, it is important to note that these decision making norms are far weaker than the decision making

when assessing the inaccuracy of confidence orderings. Specifically, one might think that it is alethically non optimal for an agent to make the judgement $p \succ q$ when $p \equiv q$ is actually the case. Certainly, an omniscient being would never make such a judgement in this case. And the current scoring rule doesn't take these mistakes into account. Again though, it turns out that evidential propriety is the reason for ignoring this third class of epistemic mistake. If we add a third term to the scoring rule corresponding to the number $M3$ of these mistakes, and weight that number by some factor α_3, then the resulting rule will be evidentially proper if and only if $\alpha_3 = 0$.

norms that are typically used in pragmatic justifications of Bayesian conditionalisation, and that they are also entailed by alternative theories of rational decision making for agents with non probabilistic credences. Finally, in Section 6, I showed that existing approaches to formalising the inaccuracy of comparative confidence judgements are not suitable for the articulation of an expected inaccuracy argument for CC. Any attempt to form such an argument will require a formalisation of epistemic utility that is radically different to either the linear rules considered by Fitelson and McCarthy or the non linear rule proposed by Spohn and Raidl.

One major implication of all this is that the normative force of CC follows from relatively weak assumptions about the synchronic coherence norms for comparative confidence judgements. Firstly, the evidentialist argument for CC presented in Section 4 relies only on the synchronic norm that a rational agent's comparative confidence judgements should always be representable by a set of Dempster Schafer belief functions.[40] This requirement is far weaker than the requirements that an agent's comprative confidence judgements be representable by either a single probability function (as assumed by standard Bayesianism) or even a set of probability functions (as assumed by imprecise Bayesianism). Secondly, the Diachronic Dutch book argument presented in Section 5 requires only that agents satisfy the betting principles CBP and CBP3, both of which are logical consequences of the far more stringent requirement that rational agents should always make decisions via expected utility calculations based on probabilistic credences. So in comparison to Bayesian conditionalisation, whose justifications typically presuppose that rational agents always have numerical credences that conform to the strict mandates of probability theory, CC is a significantly more realistic diachronic norm for agents with limited cognitive capacities (like me). Its normative force follows from pre-theoretically compelling qualitative reasoning and decision making norms that are implied not only by Bayesianism, but also by alternative non probabilistic approaches to uncertainty.

In closing, I draw the reader's attention to a major problem that remains to be solved: namely to generalise the diachronic norm CC to deal with a broader range of possible evidence. In its current form, CC applies only to agents who learn the truth of a proposition $e \in \mathfrak{B}$ *with certainty.* It says nothing about how agents should revise their confidence orderings upon acquiring more equivocal evidence. For example, an agent might learn only that p is more likely to be true than q is, or that e is actually evidentially independent of p. In the Bayesian setting, subtle evidential constraints like these can be integrated by means of Jeffrey conditionalisation and distance minimisation methods, both of which reduce to Bayesian conditionalisation in the special case where a proposition is learned

[40] In fact, it follows from even weaker assumptions than this, but I phrased the result in terms of Dempster Schafer functions for ease of exposition.

with certainty. Of course, no analogous techniques exist for comparative confidence judgements, and the task of generalising CC to obtain methods like these is a pressing one that I will return to in a sequel to this paper.

Appendix

Proofs

Proof of Proposition 1: $\top \succ_e \bot \Leftrightarrow (e \wedge \top) \succ (e \wedge \bot) \Leftrightarrow e \succ \bot$. So $\succsim_e$ satisfies A1 if and only if $e \succ \bot$. To see that $\succsim_e$ satisfies A2 as long as $\succsim$ does, let $p \vdash q$. Then $(e \wedge p) \vdash (e \wedge q)$. So $p \succsim_e q \Leftrightarrow (e \wedge p) \succsim (e \wedge q)$, which is guaranteed by $\succsim$ satisfying A2. $\blacksquare$

Proof of Proposition 2: Let $\succsim$ satisfy the condition. Proposition 2 guarantees that $\succsim_e$ satisfies A1 and A2. So we just need to show A3. Let $p \vdash q$, $q \succ_e p$ and let $q \vdash \neg r$. This implies that $(e \wedge p) \vdash (e \wedge q)$, $(e \wedge q) \succ (e \wedge p)$ and $(e \wedge r)$ is incompatible with $(e \wedge q)$. So, since $\succsim$ satisfies A3, we get that $(e \wedge q) \vee (e \wedge r) \succ (e \wedge p) \vee (e \wedge r)$, i.e. $e \wedge (q \vee r) > e \wedge (p \vee r)$, i.e. $(q \vee r) >_e (p \vee r)$, which proves that $\succsim_e$ satisfies A3, as desired. $\blacksquare$

Proof of Proposition 3: By definition, if $\succsim$ is fully representable by a set S of probability functions, then $\succsim_e$ is fully representable by the set $S_e = \{P(-|e)|P \in S\}$, which proves the proposition. $\blacksquare$

Proof of Proposition 4: Fix $\mathfrak{B}$ and let U be an updating rule for $\mathfrak{B}$, i.e. a function that takes a partial preorder $\succsim$ over $\mathfrak{B}$ and a proposition $e \in \mathfrak{B}$ and returns a partial preorder $\succsim_{U(\succsim,e)}$ on $\mathfrak{B}$ such that $e \sim_{U(\succsim,e)} \top$ (where it is assumed that $\succsim$ and $\succsim_{U(\succsim,e)}$ satisfy $\mathfrak{C}2$). We show that U satisfies EN if and only if $\succsim_{U(\succsim,e)} = \succsim_e$ for all $e \in \mathfrak{B}$ (where $\succsim_e$ denotes the posterior ordering produced by CC).

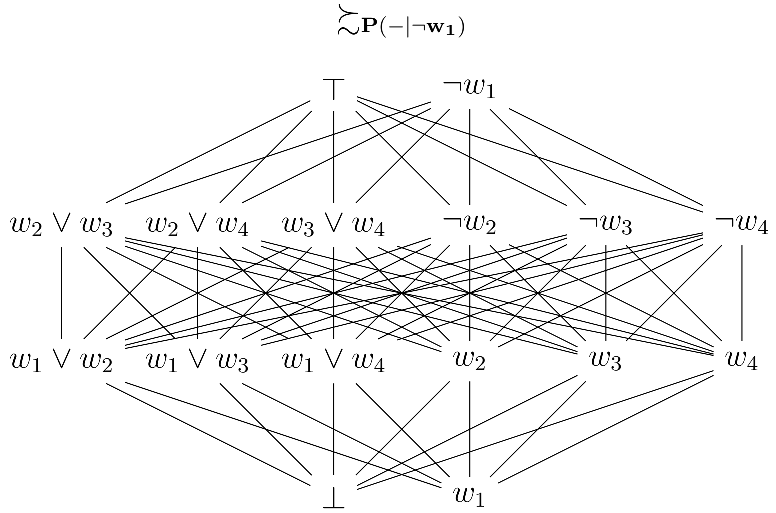
First of all, suppose that U does coincide with CC. Then for any $p, q, e \in \mathfrak{B}$ with $p \vdash e$, $q \vdash e$,

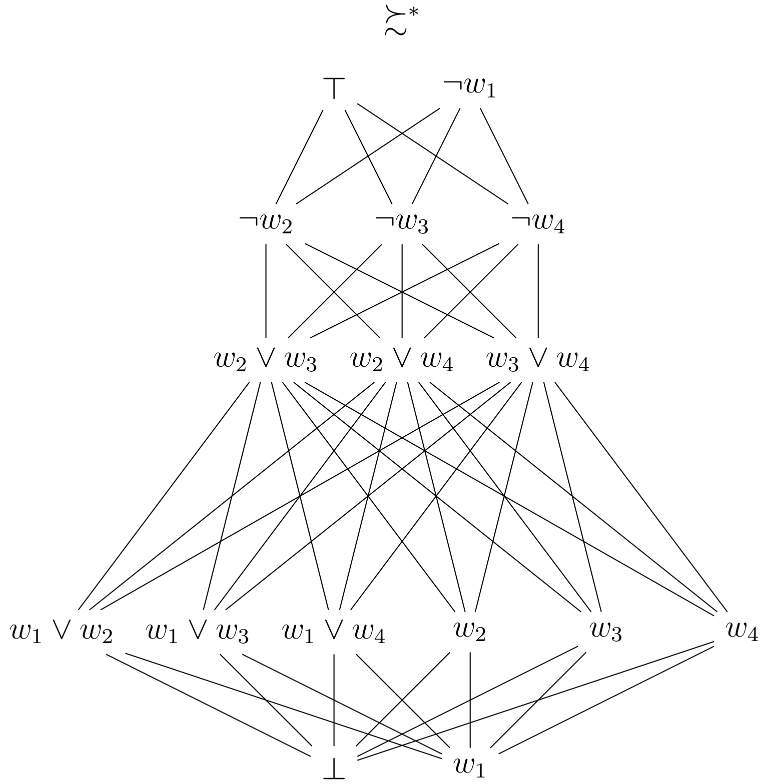
$$p \succsim_{U(\succsim,e)} q \Leftrightarrow (e \wedge p) \succsim (e \wedge q) \Leftrightarrow p \succsim q$$

Which shows that U satisfies EN. Conversely, let U satisfy EN. Then, since $\succsim_{U(\succsim,e)}$ satisfies $\mathfrak{C}2$, $\mathfrak{C}2$ ensures that $e \sim \top$ implies $e \sim (e \wedge p)$ for all $p \in \mathfrak{B}$, and $e \wedge p \vdash e$, $e \wedge q \vdash e$,

$$p \succsim_{U(\succsim,e)} q \Leftrightarrow (e \wedge p) \succsim_{U(\succsim,e)} (e \wedge q) \Leftrightarrow p \succsim q \quad \blacksquare$$

Proof of Proposition 6: Let $\mathfrak{B}$ be the algebra generated by the worlds $\{w_1, w_2, w_3, w_4\}$ and define P by $P(w_1) = 0$, $P(w_2) = P(w_3) = P(w_4) = \frac{1}{3}$. Compare the confidence ordering $\succsim_P$ with the ordering $\succsim^*$ depicted below.





It is easy to verify that on Fitelson and McCarthy's scoring rule, $\succsim^*$ weakly dominates $\succsim_P$, and hence that P expects $\succsim^*$ to be more accurate than $\succsim_P$, which shows that the rule violates evidential propriety. ∎

References

Alon, Shiri. and Lehrer, Ehud. (2014). Subjective multi-prior probability: A representation of a partial likelihood relation. *Journal of Economic Theory*, 151: 476-492

de Finetti, Bruno. (1937). La prevision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincare*, 7: 1-68.

Briggs, R.A. and Pettigrew, Richard. (2018). An Accuracy-Dominance Argument for Conditionalisation. *Nous*, https://doi.org/10.1111/nous.12258.

Denoeux, Thierry. (2019). Decision-making with belief functions: A review. *International Journal of Approximate Reasoning*, 109:87-110.

Eva, Benjamin. (2019). Principles of Indifference. *The Journal of Philosophy*, 116(7).

Fagin, Ronald. and Halpern, Jospeh. (1991). A New Approach to Updating Belief Functions. Henrion, Kanal and Lemer (eds). *Uncertainty in Artificial Intelligence 6*. Elsevier.

Fine, Terrence. (1973). *Theories of Probability.* Academic Press.

Fishburn, Peter. (1986). The axioms of subjective probability. *Statistical Science* 1(3), 335-345.

Fitelson, Branden. and McCarthy, David. (unpublished). *Toward an Epistemic Foundation for Comparative Confidence*. Manuscript.

Forrest, Peter. (1989). The Problem of Representing Incompletely Ordered Doxastic Systems. *Synthese*, 79:279-303.

Greaves, Hilary. (2006). Justifying Conditionalization: Conditionalization Maximises Expected Utility. *Mind*, 459: 607-632.

Halpern, Joseph. (2003). *Reasoning About Uncertainty.* MIT Press.

Harrison-Trainor, Matthew., Holliday, Wesley. and Icard, Thomas. (2016). A note on cancellation axioms for comparative probability. *Theory and Decision*, 80: 159-166.

Hawthorne, James. (2016). A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions. In A. Hajek and C. Hitchcock (eds.), *The Oxford Handbook of Probability and Philosophy.* Oxford University Press: 277-295.

Icard, Thomas. (2016). Pragmatic Considerations on Comparative Probability. *Philosophy of Science*, 83: 348-370.

Joyce, James. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, 65(4): 575-603.

Joyce, James. (2010). A Defence of Imprecise Credences in Inference and Decision Making. *Philosophical Perspectives* 24(1): 281-323

Keynes, John Maynard. (1921). *A Treatise on Probability.* Macmillan.

Konek, Jason. (2019). *Comparative Probabilities.* In J Weisberg and R Pettigrew (eds), *The Open Handbook of Formal Epistemology*: 267 – 348.

Koopman, B.O. (1940). The axioms and algebra of intuitive probability. *Annals of Mathematics*, 41(2): 269-292.

Kraft, Charles., Pratt, John and Seidenberg, A. (1959). Intuitive probability on finite sets. *The Annals of Mathematical Statistics*, 30(2): 408-419.

Leitgeb, Hannes. and Pettigrew, Richard. (2010). An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science* 77: 236–272.

Lewis, David. (1976). The Probability of Conditionals and Conditional probabilities. *The Philosophical Review*, 85(3): 297-315.

Lewis, David. (1980). A Subjectivist's Guide to Objective Chance. In *Studies in Inductive Logic and Probability*, ed. R. Jeffrey, 263–293. Berkeley: University of California Press.

Lewis, David (1999). *Papers in Metaphysics and Epistemology.* Cambridge: Cambridge University Press, 1999.

Norton, John. (Forthcoming). Eternal Inflation: When Probabilities Fail. *Synthese.*

Pettigrew, Richard. (2016). *Accuracy and the Laws of Credence.* Oxford University Press.

Pettigrew, Richard. (unpublished). *What is conditionalization and why should we do it?* Manuscript.

Pettigrew, Richard. (forthcoming). *Dutch Book Arguments.* Cambridge University Press

Schafer, Glenn. (1976). *A Mathematical Theory of Evidence.* Princeton: Princeton University Press.

Scott, Dana. (1964). Measurement structures and linear inequalities. *Journal of Mathematical Psychology* 1(2): 233-247

Skyrms, Brian. (1980). *Causal Necessity.* New Haven: Yale University Press.

Skyrms, Brian. (1985) Maximum Entropy Inference as a Special Case of Conditionalisation. *Synthese*, 63(1): 55–74.

Spohn, Wolfgang (1988). Ordinal Conditional Functions: A Dynamic Theory of Epistemic States. In W.L. Harper and B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics II.* Kluwer: Dordrecht, 105–134.

Spohn, Wolfgang. (2012). *The Laws of Belief.* Oxford: Oxford University Press.

Weatherson, Brian. (2007). The Bayesian and the Dogmatist. *Proceedings of the Aristotelian Society* 107: 169-185

Williams, Peter. (1980). Bayesian Conditionalisation and the Principle of Minimum Inofmration. *British Journal for the Philosophy of Science.* 31(2): 131–144.

Wong, S. M., Y. Yao, P. Bollmann, and H. Burger (1991). Axiomatization of qualitative belief structure. *IEEE Transactions on Systems, Man and Cybernetics*, 21(4): 726-734.

Liu, Yang. (forthcoming). Incomplete Preference and Indeterminate Comparative Probability. *British Journal for the Philosophy of Science.*

Zadeh L.A. (1978) Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1: 3–28.