Philosophers' Imprint

volume 16, no. 11

june, 2016

UPDATING ON THE CREDENCES OF OTHERS: DISAGREEMENT, AGREEMENT, AND SYNERGY

Kenny Easwaran, Luke Fenton-Glynn†, Christopher Hitchcock‡, Joel D. Velasco§[1]*

**Texas A&M University,*
† University College London,
‡ California Institute of Technology,
§Texas Tech University

1. We recommend the acronym EaGlHiVe, pronounced "eagle hive".

Abstract

We introduce a family of rules for adjusting one's credences in response to learning the credences of others. These rules have a number of desirable features. 1. They yield the posterior credences that would result from updating by standard Bayesian conditionalization on one's peers' reported credences if one's likelihood function takes a particular simple form. 2. In the simplest form, they are symmetric among the agents in the group. 3. They map neatly onto the familiar Condorcet voting results. 4. They preserve shared agreement about independence in a wide range of cases. 5. They commute with conditionalization and with multiple peer updates. Importantly, these rules have a surprising property that we call *synergy* — peer testimony of credences can provide mutually supporting evidence raising an individual's credence higher than any peer's initial prior report. At first, this may seem to be a strike against them. We argue, however, that synergy is actually a desirable feature and the failure of other updating rules to yield synergy is a strike against *them*.

1. Introduction

How should you rationally respond when you discover that others hold beliefs that are either similar to or different from your own? The answer depends upon, among other things, your judgment about the reliability of those others when it comes to topics like the one at hand. Let us suppose that you consider these others to be *epistemically responsible*; you believe that they have formed their opinions in a reasonable way on the basis of evidence and reasoning that probably differs from your own. We will call such people your *epistemic peers*.[2]

To make our question more precise, suppose that P has a *credence* function P, representing P's probabilistic degrees of belief. (As a gen-

2. We take this term from Kelly (2005), who attributes it to Gutting (1982). However, our notion of an epistemic peer is somewhat different from, and weaker than, that commonly employed in the literature on the epistemology of disagreement. See section 2.3 below for more discussion.

eral convention, we will use uppercase Roman letters to represent epistemic peers, and upper case Italic letters to represent their corresponding credence functions.) Let A be some proposition in which P has an initial credence $P(A) = p$. P then meets Q, R, S,..., whom she regards as epistemic peers with respect to A,[3] and learns of their credences $Q(A) = q, R(A) = r, \ldots$. How should P revise her credence in A to arrive at a new credence $P^+(A) = p^+$? More generally, P might begin with a credence distributed over the members of a partition $\{A_1, \ldots, A_k\}$ of exclusive and exhaustive possibilities, and then learn how Q, R, S,... distribute their credences over these possibilities. Again, how should P revise her credences?

We will call this the problem of *updating on the credences of others*, or just *updating* for short. This problem is closely related to, but also importantly different from, several other problems that have received considerable attention: the problem of responding to *testimony*, the problem of judgment *aggregation* or opinion pooling, and the problem of *disagreement* and higher-order evidence.

One answer to the problem of updating is the Bayesian one: the credences of others are evidence, and when one gets evidence, one ought to update one's own credences by conditionalizing on that evidence. Conditionalizing takes into account one's views on the reliability of others because this information is encoded in one's prior *likelihoods*. Consider the simplest case with one binary (yes-no) proposition A, and two epistemic peers. If $P(A)$ is P's prior probability for A, then upon learning that $Q(A) = q$, P should update to

$$P^+(A) = P(A|Q(A) = q).$$

$P^+(A)$ is P's *posterior probability* for A.

We take Bayesian conditionalization to be normatively correct for epistemic agents that already have the requisite conditional degrees of belief. However, conditionalization imposes unrealistic cognitive demands on agents. An agent must have opinions not only about possible states of the world, but also about the possible opinions of any peer she might meet.

The complexity of conditionalization has led a number of authors to seek simple heuristics that agents can use to produce reasonable, if not perfect, answers to the problem of updating. Perhaps the most straightforward — and widely discussed — shortcut rule is linear averaging.[4] In the simplest case, where there are two peers, P and Q, who have prior credences $P(A) = p$ and $Q(A) = q$, this rule states that, upon learning each other's credence in A, the two agents should come to have a credence in A given by

$$P^+(A) = Q^+(A) = \frac{p+q}{2}.$$

This rule has the virtue of being computationally undemanding. Also, it is readily generalizable to updates on the credences of multiple peers and to allow for the agents' initial credences to be assigned unequal weights in the determination of a posterior credence.

Nonetheless, linear averaging has a number of well-known drawbacks. It almost never preserves judgments of probabilistic independence. It does not commute with conditionalization. Nor is it commutative or associative with respect to updates on multiple peers: one's final credence will depend upon the order in which one updates on one's peers' opinions. We will discuss these problems in more detail in Section 4.

We will argue that linear averaging has another important shortcoming. The agents' post-update credences can never be more extreme than the most extreme of the agents' initial credences: in the two-agent case described above, the agents' post-update credences must

3. We allow that one might be a peer with respect to some propositions but not others.

4. For recent discussions, see e.g., Jehle and Fitelson 2009; Kelly 2010; Cohen 2013; Elga 2007; Christensen 2009; Christensen 2011; Steele 2012; de Ridder 2014; Bradley 2015; Romeijn 2015; and Staffel 2015.

lie within the interval $[p, q]$. This rules out the possibility of what we shall call "synergy": that is, of the agents' credences providing mutually supporting evidence that raises their posterior credences higher than either of their initial credences. As we will show in Section 6, there are cases in which synergistic responses appear to be rational.

For these reasons, we take it to be worthwhile to seek an alternative "shortcut" rule that, while still being less computationally demanding than conditionalization, shares a greater range of the desirable features that conditionalization has than does linear averaging. For one thing, such a rule would seem likely to help us better predict and explain the behavior of otherwise rational agents operating under the sorts of constraints that prevent conditionalization.

With a view to this, we introduce and defend a multiplicative rule that we call *Upco*, short for 'updating on the credences of others'. In the simplest case, described above, where there are two peers, P and Q, who have credences for the propositions A and $\neg A$, *Upco* implies that both should update to

$$P^+(A) = Q^+(A) = \frac{pq}{pq + (1-p)(1-q)}. \qquad (Upco^\dagger)$$

(The above formula is labelled $Upco^\dagger$ to indicate that it is a special case of our rule *Upco*.) In the fully general case, where peers P, Q, R,... have credences over the partition $\{A_1, \ldots, A_k\}$, with $P(A_i) = p_i, Q(A_i) = q_i$, etc., our rule becomes:

$$P^+(A_i) = \frac{p_i q_i r_i \ldots}{\sum_{j=1}^{k} p_j q_j r_j \ldots} \qquad (Upco)$$

Upco, like linear averaging, is computationally undemanding. Nonetheless, it has several key advantages over linear averaging.

Firstly, although the agents' likelihoods are not invoked in this formula, it is equivalent to conditionalization when the likelihood ratios of the agents take a particular simple form (Section 7 and Section 8). We will thus say (Section 7.1) that our rule *mimics* these likelihoods (or, more precisely, that it mimics the result of conditionalization for a Bayesian agent who has these likelihoods), even though it does not require the agent to have these numbers available to them to plug into the formula. In fact, we'll see (Section 7 and Section 8) that *Upco* is compatible with many different likelihoods and thus might be expected to yield reasonable results even for agents whose likelihood function is not fully specifiable.

A second advantage of *Upco* is that, when agents update in accordance with this rule, the agents' post-update credences do not lie in the $[p, q]$ interval where it is not reasonable for them to do so (see Section 6). Thirdly, as we shall show in Section 9, updates in accordance with this rule commute with conditionalization (Section 9.3) and are commutative and associative with updates upon the credences of multiple peers. Finally, updating in accordance with *Upco* preserves shared agreement about independence in a wide range of cases (Section 9.2).

Upco is intended as a heuristic; we do not claim that it yields reasonable results in all cases: only conditionalization does *that*. Nor do we claim that *Upco* is normative for agents that do not have the wherewithal to use conditionalization. As we shall see in Section 10, there are cases in which the simple likelihood function that we will use to motivate *Upco* is implausible. However, in that section, we show that there are some generalizations of *Upco* that accommodate a range of such cases. Updating according to these generalized rules places more demands upon an agent, but the rules are still not as demanding as conditionalization. In Section 10 we will argue that another key advantage of *Upco* over (weighted) linear averaging is its heuristic value: it is suggestive of a greater range of generalizations, which allows one to handle a greater range of cases.

2. Related Problems

The problem of *updating on the credences of others*, which is our focus, is closely related to three other problems, which we will call the problems of *testimony*, *aggregation*, and *disagreement*. In this section, we briefly survey these problems, and highlight their differences from the problem of updating.

2.1 *Testimony*

The problem of testimony concerns how one should respond to the reports of partially reliable witnesses. For instance, suppose that m witnesses testify that proposition A is true, while n witnesses testify that A is false. If we have some kind of probabilistic measure of how reliable each witness is, what should be our credence for A? A closely related problem concerns jury deliberations. Suppose that there are $m + n$ jurors, with $m > n$, and each juror has a certain probability of reaching a correct verdict. If at least m jurors must agree in order to reach a verdict, what is the probability that the jury as a whole will reach the correct verdict? The jury problem is usually formulated as a problem about the *objective probability* that a jury will reach the correct verdict, rather than how an agent should revise her beliefs in light of the jury's vote.

There is a very long history of work on probabilistic accounts of testimony. Questions about how a jury should use probability theory to reach a verdict go back at least to Leibniz, and the late 18th and early 19th centuries in particular are filled with numerous Bayesian analyses of testimony — much of it driven by Hume's widely read *Of Miracles* (originally Book X of *An Enquiry Concerning Human Understanding* — Hume [1748] — and later published separately). Condorcet produced his famous *jury theorem* in 1785. For a good historical survey, see Zabell (1988).

The problem of testimony assumes that witnesses report full beliefs rather than credences. That is, each witness reports his belief that A is true or A is false. In many cases, we assume that the witness has some given probabilistic reliability p. We could then try to link the testimony problem to the updating problem by mapping the witnesses' report that A is true onto a credence of p for A, and $1 - p$ for $\neg A$. (A separate issue is what we mean by saying that a witness has a reliability of p. For example, does it mean that $P(A|\text{witness reports } A) = p$, or $P(\text{witness reports } A|A) = p$, or something else? And what if the probabilities are different for A and $\neg A$?) This will generate reasonable results in many cases, but in general, a person's degree of belief in A is not the same as their reliability if they testify that A.

In addition, the more general case of the problem of updating, where one's peers report their credences over a partition $\{A_1, \ldots, A_k\}$, does not have any obvious analog in the problem of testimony. There has been some work on generalizing the problem of testimony to include cases where witnesses have different reliabilities and where there are numerous possible answers to a question, rather than just testimony that A is true or A is false (see Van Cleve [2011] for an overview of some of these attempts). But all such attempts still rely on a number of simplifying assumptions and do not get at the heart of the subjective nature of the problem of updating. The logic of the testimony of full beliefs is the wrong place to start if one hopes to understand the general case of updating.

2.2 *Aggregation*

The problem of *aggregation* (sometimes called the problem of *opinion aggregation* or *opinion pooling*) involves a group of agents who have credences in some proposition A, or over a range of propositions. The goal is to find an appropriate way to represent the opinion of the group as a whole. Perhaps the group has to decide whether to take a certain investment opportunity (as in Russell et al. 2015).

There are two important differences between the aggregation problem and the updating problem. The first is that the aggregation problem concerns how best to represent the collective opinion of the group. It does not presuppose that any one member is rationally required to revise her credences in light of the credences of the others in the group.

Second, a common stipulation in the aggregation problem is that the group be in a state of "dialectical equilibrium" (Lehrer and Wagner 1981, 19; Genest and Zidek 1986, 125) or "reflective equilibrium" (Steele 2012, 984). Practicing statisticians might understand this idea in behavioral terms. The agents meet, explain the reasons behind their credences, and revise their own credences. Equilibrium is reached when no agent wishes to revise her credences further (regardless of whether she *ought to* further revise her credences).

The updating problem arises earlier in the process. One encounters others whom one regards as epistemic peers, and whose credences may differ from one's own. One then might revise one's own credences. This revision constitutes the dynamics by which equilibrium is reached. The updating problem comprises the question of how these dynamics are rationally constrained.[5]

2.3 *Disagreement*

The problem of epistemic disagreement has received a great deal of attention recently. This problem concerns how one should respond to disagreement with an epistemic peer over some proposition A. Crucially, however, the notion of *peerhood* that is employed in this literature is much stronger than ours. Typically, it assumes that your peer has precisely the same evidence regarding A that you do, and that you regard your peer as equally competent with respect to the subject matter of A.

Some authors address the problem of disagreement in full beliefs: I believe that A, while my peer believes that $\neg A$. See for example Kelly (2005), Feldman (2006), Lackey (2010), and van Wietmarschen (2013). Others, such as Jehle and Fitelson (2009), Wagner (2011), and Steele (2012), frame the notion of disagreement in terms of credences as, sometimes, do Elga (2007), Kelly (2010), de Ridder (2014), and Christensen (2011, 2010, 2009, 2007).

Many of the responses to the problem of disagreement have fallen into one of two camps. Defenders of *steadfast* responses, e.g., Kelly (2005, 2010), claim that it is reasonable to retain your original belief or credence in the face of disagreement. Others, such as Elga (2010) and Christensen (2007, 2011), advocate *conciliatory* responses to peer disagreement, in which you revise your belief or credence to accommodate your peer's opinion. One form of the conciliatory approach is the *Equal Weight View*, which advocates giving your peer's opinion the same weight as your own. We will discuss the Equal Weight View in more detail in Section 5 below. We will add, at least as logical possibilities, *deferential* views, in which you defer to the opinion of your peer (see, e.g., Joyce 2007).

The problem of disagreement is closely connected to the problem of how to respond to *higher-order evidence* (Christensen 2010). Higher-order evidence for a proposition A concerns your evaluation of the evidence for or against A. For instance, if someone shows you that you have systematically underestimated the probability of propositions like A in the past, this might cause you to adjust your credence in A upward. If your peer has the same evidence for A that you do, but nonetheless assigns a different credence to A, should you take this as evidence that you have misevaluated the evidence for A? [6]

We are skeptical about the prospects of clearly unpacking the stipulation that epistemic peers share the same evidence. For example, suppose that P and Q are chicken-sexers. According to philosophical lore, chicken-sexers learn to distinguish male from female chicks fairly reliably by being shown a large number of each. However, the chicken-sexers cannot explain why one chick looks male or female, they do not consciously apply any criteria in reaching their judgments, and they cannot communicate to others what to look for.[7] P and Q look at a chick, and P's credence that it is male is .7, while Q's credence is .9. Do they share the same evidence? They both have access to the same

5. This distinction is essentially the same as Bradley's distinction between "deliberation" and "aggregation" (Bradley 2007).

6. Here we have benefited from helpful discussion with Paulina Sliwa and Leon Leontyev.

7. Philosophical lore is not entirely accurate, as there are several methods for distinguishing the sexes of chicks, although ambiguous cases can still occur.

sensory information. But, for all they know, they are each responding to different sensory cues. In this case, does it even make sense to decompose P's credence into evidence and response to evidence (and likewise for Q)?

In any event, in our treatment of the updating problem, we do not assume that the epistemic peers share the same evidence.[8] In Section 10.2 we show how our update rule may be modified to reflect shared background knowledge among the peers.

Another respect in which the updating problem differs from the problem of disagreement is that we do not assume that the epistemic peers actually disagree, in the sense of assigning different credences to some proposition(s). As we shall see, the case of peer agreement actually turns out to be more interesting and complicated than might be expected.

2.4 *Updating*

The problems of updating, testimony, aggregation, and disagreement are different enough that there is no guarantee that a good solution to one will be a good solution to the others. At the same time, the problems are similar enough to borrow concepts and terminology from each other. For example, it is clear enough what a *steadfast* or *deferential* solution to the problem of updating would be, even though these terms are imported from the literature on disagreement.

It remains possible that one might solve one of these problems by means of a solution to another. For example, there are at least two attempts to solve the problem of aggregation by means of a solution to the problem of updating. Lehrer and Wagner (1981) propose that the peers first update on each other's credences, and then iterate the updating process until a consensus is reached. The consensus then represents the aggregated opinion of the group.[9] Another approach to the aggregation problem, dubbed *supra-Bayesianism* by Keeney and Raiffa (1976), has its origins in the work of Winkler (1968) and Morris (1974, 1977). The proposal is to represent the aggregate opinion of a group by supposing that the members inform a hypothetical decision-maker (the *supra-Bayesian*) of their opinions.

While we claim only to offer a useful heuristic for addressing the problem of updating, we encourage those working on these other problems to make use of our ideas.

With these important clarifications in mind, we are now ready to describe some existing answers to the updating problem, before proceeding to propose our own, novel answer.

3. Bayesian Conditionalization

A familiar rule for revising one's beliefs in the light of new evidence is *Bayesian conditionalization*. If $P(A)$ is P's initial credence in A, and she then learns that proposition E is true (and nothing else), her new credence would become $P^+(A) = P(A|E)$. $P(A)$ is P's *prior* probability for A, and $P^+(A) = P(A|E)$ is her *posterior* probability.

A natural suggestion, then, is that when P learns that Q's credence in A is $Q(A) = q$, she treats the proposition $Q(A) = q$ as evidence and updates by conditionalization. Thus her new credence will be $P^+(A) = P(A|Q(A) = q)$. By Bayes' Theorem:

$$P^+(A) = \frac{P(Q(A) = q|A) \cdot P(A)}{P(Q(A) = q|A) \cdot P(A) + P(Q(A) = q|\neg A) \cdot P(\neg A)}.$$

$P(Q(A) = q|A)$ and $P(Q(A) = q|\neg A)$ are P's *likelihoods* for $Q(A) = q$. It is often helpful to express Bayes' theorem in the form of an *odds ratio*:

8. Dietrich (2010) discusses the importance of distinguishing between the cases when the peers have the same information and when they have different information in the context of the aggregation problem. He says that in the case where they have the same information, they have *symmetric* information, and when they have different information, they have *asymmetric* information. We think that this terminology is misleading, however. For example, one might use the same terminology to distinguish cases where each peer has information the other lacks (*symmetric*) from cases where one peer has strictly more information than the other (*asymmetric*).

9. For criticisms of this proposal, see Loewer and Laddaga (1985, 86), Martini et al. (2013), and Elga (2010).

$$\frac{P^+(A)}{P^+(\neg A)} = \frac{P(A)}{P(\neg A)} \cdot \frac{P(Q(A) = q|A)}{P(Q(A) = q|\neg A)}.$$

The odds-ratio form of Bayes' Theorem shows how your posteriors depend on your priors multiplied by the *likelihood ratio*.

As others have noted,[10] however, while conditioning in this way may be the normative standard, it would still be of great benefit to look for further, simpler formulas. In the first place, conditionalizing is computationally complicated, and this difficulty explodes when we have multiple propositions and/or the credences of multiple agents are involved. For example, suppose that P has credences over the algebra generated by the propositions $A, \neg A, B, \neg B$. In this case, there are four basic possibilities. To fully specify P's credence function, we need to specify her credence in three of these possibilities. Now suppose that P has one peer, Q, who has credences for the same propositions. For simplicity, suppose that Q's credences must be multiples of .25. There are 35 possible credence functions for Q. For P to have credences for A and B, and also over Q's various possible credences, P needs to have credences over an algebra with 140 elementary possibilities. And if P has two peers whose credences come in multiples of .1, then we have over 300,000 elementary possibilities. We quickly face a combinatorial explosion.

Worse, in real cases with real agents, it is doubtful we even *have* the requisite likelihoods.[11] For instance, suppose that you have just been introduced to someone who is a professor at the University of Helsinki or the University of Sydney (whichever is farther from you). In light of this person's position, you decide to treat her as an epistemic peer with respect to some proposition under discussion. Since you literally just met this person, there is no way that you could have prior degrees of belief about this person's credences. So a shortcut rule that is computationally tractable, easy to understand, and easy to generalize and modify is desirable.

10. See, e.g., Genest and Schervish (1985); Dawid et al. (1995, 310–311); Steele (2012, 986–987); Bradley (2015); and Romeijn (2015).
11. See, e.g. Genest and Schervish (1985, 1198–1200, 1205); Loewer and Laddaga (1985, 87); Dawid et al. (1995, 311); Bradley (2015); Romeijn (2015).

4. Linear Averaging

One very simple heuristic that has received considerable attention is *linear averaging*. Suppose that $P_1, P_2, \ldots, P_n$ are peers with respect to proposition A. Then, when they learn each other's credences, they could each update to the average of the group:

$$P_1^+(A) = P_2^+(A) = \ldots = P_n^+(A) = \frac{P_1(A) + P_2(A) + \ldots P_n(A)}{n}.$$

Jehle and Fitelson (2009, 284) call this rule *Straight Averaging*.

A common generalization of the Straight Averaging rule is one that allows weighted averaging with non-equal weights. A weighted linear averaging rule states that

$$P_1^+(A) = \sum_{i=1}^{n} w_i P_i(A),$$

where $w_1, \ldots$, and w_n respectively represent the weights of respect P_1 assigns to *herself* and to each of her $n-1$ peers, and where $\sum_{i=1}^{n} w_i = 1$. If each w_i is $1/n$, this yields the Straight Averaging rule.

The weighted linear average was proposed as a solution to the problem of aggregation by Stone (1961). Others, such as DeGroot (1974), Lehrer and Wagner (1981), and Genest and Schervish (1985), adapted the linear averaging rule to the problem of updating.[12]

12. For recent discussions, see e.g., Jehle and Fitelson 2009; Kelly 2010; Cohen 2013; Elga 2007; Christensen 2009; Christensen 2011; Steele 2012; de Ridder 2014; Bradley 2015; and Romeijn 2015.

One advantage of the linear averaging strategy, besides its simplicity, is that it provides a natural framework for articulating the different positions that have been defended in response to the problem of disagreement. An epistemic agent P is being steadfast if she assigns herself a weight close to one; she is being conciliatory if she assigns her peers substantial positive weights; and she is being deferential if she assigns herself a weight close to zero. The natural way to cash out the Equal Weight View, in this framework, would be in terms of the Straight Averaging rule, where each peer receives equal weight. As we will argue in the next section, however, it would be a mistake to conclude that this is the only way to cash out the Equal Weight View.

Unfortunately, linear averaging suffers from a number of well-known drawbacks. The first is that it is neither commutative nor associative, as we illustrate with the following example:

Example 4.1. *Suppose that peers P, Q, and R have credences in A of* .3, .5, *and* .8 *respectively, and that P updates on the opinion of her peers by the Straight Averaging rule. If P meets Q first, she averages* .3 *and* .5 *to get* .4*. Then, when P meets R, she averages* .4 *with* .8 *to get a final value of* .6*. However, if she first meets R, she averages* .3 *and* .8 *to get* .55*. Then, when she meets Q, she averages* .55 *with* .5 *to get final value of* .525*. Finally, if she meets both peers at once, and averages the three numbers together, she gets a new credence* .533*.*

This result is undesirable; thus, the defender of linear averaging has to say something more about the case of multiple peers. Cohen (2013), for example, proposes that

> ...as I encounter new peers, I should revise by continually averaging over all of them. So regardless of the order in which I encounter them, I will end up with the same credence. (Cohen 2013, 116)

It is far from clear exactly what this proposal amounts to. At the very least, this would seem to require that for every proposition in which I have a credence, I keep track of how many peers have had input into my credence for that proposition. Gardiner (2014, 92–93) raises worries about the possibility of adequately addressing this issue, as well as offering compelling arguments against a number of alternative responses that the linear averager might give.

A second problem with linear averaging is that it is not always compatible with conditionalization. That is, if P updates on the opinions of her peers by taking a weighted average of the group's credences, it is not always possible to find likelihoods for P such that conditionalization using those likelihoods yields the same result as averaging. Dawid et al. (1995) and Bradley (2015) discuss the problems that arise in the specific case where a peer is perfectly deferential. But the incompatibility of linear averaging with conditionalization also follows from the failure of linear averaging to be commutative and associative. Bayesian conditionalization *is* commutative and associative. If we view conditionalization as the normative standard, then compatibility with conditionalization seems like a straightforward desideratum of any heuristic. Lasonen-Aarnio (2013) argues that compatibility is impossible, unless one makes restrictive and implausible assumptions about the agent's prior credences.

Third, linear averaging does not preserve judgments of independence. That is, if P and Q both judge propositions A and B to be independent, and P^+ is a linear average of P and Q, then A and B will not be independent in P^+ except in rare cases (e.g. when P and Q are identical, or where the averaging rule used assigns one of the peers zero weight).[13] This is closely related to Simpson's Paradox. We illustrate this property with the following example:

Example 4.2. *Two coins are about to be flipped, so that the algebra of propositions under consideration is* HH, HT, TH, TT*, and their disjunctions. Let P have credences indicating a credence* $1/3$ *of each coin coming up heads, and independence of the flips. Thus,* $P(HH) = 1/9$*,* $P(HT) = P(TH) = 2/9$*, and* $P(TT) = 4/9$*. Similarly, let Q have credences indicating a credence*

13. See, e.g. Laddaga (1977) and Loewer and Laddaga (1985, 89–90).

$2/3$ *of each coin coming up heads, and independence of the flips. Thus,* $Q(HH) = 4/9$, $Q(HT) = Q(TH) = 2/9$, *and* $Q(TT) = 1/9$. *P updates using Straight Averaging, so* $P^+(HH) = P^+(TT) = 5/18$ *and* $P^+(HT) = P^+(TH) = 2/9 = 4/18$. *This distribution doesn't treat the two flips as independent, but instead treats them as slightly correlated — for independent flips, the probability of HT would have to be in between the probabilities of HH and TT.*

Wagner (1985) responds to a similar coin flipping case from Loewer and Laddaga (1985) by saying that "obviously" the correct thing to do is to first average the biases of the individual coin tosses, and then calculate the probabilities of various sequences of tosses on the assumption that the tosses are independent. We agree that there are some cases where this is the obvious thing to do. But how the formal apparatus of linear averaging can achieve this is entirely mysterious. Other than saying "When you get the wrong answer, don't do it", we don't think that Wagner can say much else here. Wagner (2011) shows how, if we wanted to, we could preserve the independence of any fixed finite family of countable partitions. This move seems rather ad hoc. Of course one can preserve specific independences if one specifically sets out to do so. But what we want is a general rule that we can follow where it simply falls out of the rule that certain kinds of agreed-upon independences are preserved.

A fourth problem with linear averaging is that it does not commute with conditionalization.[14] If P meets Q, updates her credence in A on the basis of Q's credence, and then learns E and conditionalizes, she will arrive at a different credence for A than if she performed the operations in the reverse order. We can illustrate this property using the previous example of the coin. If P learns that the outcome of the first coin toss was heads and conditionalizes, her credence for the second coin coming up heads will remain $1/3$. She then meets Q, whose credence for heads on the second toss is $2/3$ (regardless of whether or not he also learned the outcome of the first toss). Using Straight Averaging, P's new credence for heads on the second toss will be $1/2$. On the other hand, suppose she first meets Q and updates on the basis of Q's credence. Then (as shown above) her new credences will be $P^+(HH) = P^+(TT) = 5/18$ and $P^+(HT) = P^+(TH) = 2/9 = 4/18$. Now, if she learns that the first coin landed heads and she conditionalizes, her new credence for heads on the second toss will be $5/9$.

In Section 6, we will discover yet another drawback of linear averaging. A number of these properties are subjects of various impossibility theorems; we will return to these issues in Section 9 below. In the meantime, however, we have provided ample motivation to search for an alternative shortcut rule for updating.

5. The Equal Weight View

One mainstream view in the disagreement literature is known as the *Equal Weight View* or EWV (Elga 2007). The informal idea behind the EWV is that, in cases of disagreement with an epistemic peer, one should treat one's own judgment as exactly on a par with that of one's peer. This view is motivated by the very strong sense of peerhood that is invoked in the disagreement literature.

As we noted in Section 4, linear averaging provides a natural way to formulate the EWV. The EWV corresponds to Straight Averaging, i.e. linear averaging with equal weights. Indeed, the EWV has often been taken to be synonymous with Straight Averaging. For example, Kelly (2010) and Cohen (2013) simply define the EWV in this way, while Elga (2007, 484, 486n) assumes that the EWV implies that, in the case of a disagreement with a peer, one ought to move (halfway) "in the direction" of one's peer's view.[15] Likewise, in their paper "What is the 'Equal Weight View'?", Jehle and Fitelson (2009) answer the titular question by examining (only) the linear averaging rule with equal weights and some minor variants thereof.

14. See, e.g. Loewer and Laddaga (1985, 88); Jehle and Fitelson (2009, 285–287); Wilson (2010, 323ff); and Russell et al. (2015).

15. Wilson (2010) adduces further evidence that Elga takes EWV to entail (linear) "averaging" (with equal weights).

By contrast, we take the essence of the EWV to be the condition Russell et al. (2015) call *anonymity*.[16] This condition states that the rule for updating one's credences be a function only of the credences themselves, without regard to which credence belongs to which peer. For example, if $P(A) = p$ and $Q(A) = q$, then a rule satisfying anonymity would return the same value for $P^+(A)$ as it would if $P(A) = q$ and $Q(A) = p$. Straight Averaging exhibits this symmetry, but it is certainly not the only rule to do so. Our rule, which is described in Section 7 below, satisfies anonymity and thus respects the intuition that we take to underlie defenses of the EWV.

6. Synergy

We will say that P_1 responds *synergistically* to learning the initial credences of her epistemic peers P_2, ..., and P_n if her new credence lies outside of the interval $[min\{P_i(A)\}, max\{P_i(A)\}]$. Linear averaging will never produce synergy so long as all of the weights are between zero and one.[17] While synergy might seem, at first, like an undesirable feature of an updating rule, we will argue in this section that it is in fact a desirable feature at least in some cases.

Christensen (2009) describes a case where synergy seems warranted:

> Suppose, for example, that I am a doctor determining what dosage of a drug to give my patient. I'm initially inclined to be very confident in my conclusion, but knowing my own fallibility in calculation, I pull back a bit, say, to 0.97. I also decide to ask my equally qualified colleague for an independent opinion. I do so in the conciliatory spirit of using her reasoning as a check on my own. Now suppose I find out that she has arrived — presumably in a way that also takes into account her fallibility — at 0.96 credence in the same dosage. Here, it seems that the rational thing to do is for me to increase my confidence that this is the correct dosage, not decrease it as difference-splitting would require. But this is not inconsistent with giving equal weight to my colleague's opinion . . . . (p. 759; a similar example is given in Christensen 2011, 3n)

We agree with Christensen that in this case one ought to increase one's confidence in one's conclusion. This example is just one of a very large number in which a synergistic response is rational. Such a response will be rational in any case in which, despite the fact that one's credence in A differs from those of one's peers, the different credences nevertheless represent a kind of agreement about whether the evidence favors A or $\neg A$. In some cases of disagreement, the credences of one's peers are evidence against one's view; here, they are evidence for it. Loewer and Laddaga (1985, 86) note the rational possibility of synergistic responses to learning the credences of one's peers. Dietrich (2010) argues that if the members of a group do not all share the same information, then aggregating the group's opinion should sometimes produce synergy.

Here is a second example of synergy. Even critics of the EWV agree that perceptual judgments can provide powerful motivations for EWV (compare Kelly 2010, 150–152). A standard example is that in which you and your peer view the same horse race; you believe that Horse A is the winner, but then discover that your peer believes Horse B is the winner. It seems that this discovery should make you less confident that Horse A won — perhaps even abandoning your belief that it did. However, suppose that after viewing the race but before conferring,

16. See also Moss (2011) and Lasonen-Aarnio (2013) for different ways of cashing out EWV. Moss proposes that the problem of disagreement could be handled by maximizing the average *epistemic utility* of the agents. Then, EWV would amount to the proposal that each agent's epistemic utility receive equal weight. Lasonen-Aarnio proposes a framework in which an agent's own credences are not transparent to herself. Then, she learns both her own credences and those of a peer. In this framework, Lasonen-Aarnio understands the EWV to entail that the agent is as certain that her peer's credence is "correct" as that her own credence is "correct". She leaves it open-ended what it would be for a credence to be "correct" in the relevant sense.

17. Genest and Schervish (1985), e.g., explore linear averaging rules where the weights are not so constrained.

you each have a credence of .8 that Horse A was the winner. In such a case, it seems clear that each of you should become even more confident that Horse A was indeed the winner. Linear averaging cannot deliver this verdict.

We may support these intuitions with some elementary considerations from confirmation theory. Suppose that you think your peer is somewhat reliable in judging the outcomes of horse races. Specifically, suppose that you think your peer is more likely to report a credence of .8 for A's victory if A in fact won than if A lost. That is, suppose that your credence P in your peer's credence Q satisfies the inequality $P(Q(A) = .8|A) > P(Q(A) = .8|\neg A)$. Then, the proposition $Q(A) = .8$ *confirms* A for you, and if you conditionalize on this proposition, your credence in A will go up. This is true regardless of your prior credence in A. In particular, even if you have credence $P(A) = .8$, your credence in A will go up when you learn that $Q(A) = .8$.

We also note that synergy is a familiar feature of solutions to the problem of testimony. Even if each witness has a fixed probability $p < 1$ of being correct, if enough independent witnesses testify to the truth of A, the probability that A is true can be made arbitrarliy close to one.

Having defended the general principle that a rational updating rule should allow for the possibility of synergy, we will now introduce our update rule, which prescribes synergy in the above cases.

7. The *Upco* Rule

In this section, we introduce our proposed update rule, *Upco*, in the special case where an agent updates on a peer's credence on an algebra generated by the propositions A, $\neg A$, first for the case of two peers, and then for multiple peers. In Section 8, we generalize to updates on credences over more complex algebras. The versions of the rule described initially will be appropriate to apply in the case of agents with completely independent evidence, who are each equally reliable in their response to this evidence. In Section 10, we will suggest ways to apply the rule (with some modifications) to other cases as well, including many of the traditional cases considered for peer disagreement.

7.1 *Two Peers*

To introduce our update rule, let us consider in more detail how an ideal Bayesian agent would respond to learning the initial credences of her epistemic peers. Specifically, suppose that this agent P has initial credence $P(A) = p$ in proposition A. P has a peer, Q, whose opinion she accords some weight. This peer has credence $Q(A) = q$. How should P revise her credence in light of learning her peer's opinion about A?

A correct Bayesian analysis would suggest that P conditionalize on $Q(A) = q$. Assuming that P has likelihoods of the form $P(Q(A) = q|A)$ and $P(Q(A) = q|\neg A)$, she can use them to update by a simple application of Bayes' theorem. As mentioned above, P's posterior probability will be:

$$P^+(A) = \frac{P(Q(A) = q|A) \cdot p}{P(Q(A) = q|A) \cdot p + P(Q(A) = q|\neg A) \cdot (1 - p)}. \qquad \text{(Post)}$$

An equivalent formulation in terms of the odds ratio is:

$$\frac{P^+(A)}{P^+(\neg A)} = \frac{p}{1-p} \cdot \frac{P(Q(A) = q|A)}{P(Q(A) = q|\neg A)}. \qquad \text{(Odds)}$$

Our proposed rule, *Upco*, is designed to give the same result as Bayesian conditionalization when P's *likelihoods* — her beliefs about what credences Q will have, conditional on A and $\neg A$ — take a particular, simple form.

Suppose that P believes that Q is *reliable*. This means that P thinks that Q is more likely to report higher credences in A when A is true, and more likely to report lower credences in A when A is false. One very simple way to implement this idea would be for P's likelihoods for Q's credences to be *linear* in Q's credences. Specifically:

$$P(Q(A) = q|A) \propto q;$$
$$P(Q(A) = q|\neg A) \propto 1 - q.$$

There are several ways in which P's likelihoods could have this form. We will first consider the discrete case.

Example 7.1. *Suppose that P knows that Q will report his credence rounded off to the nearest tenth. Her likelihoods are as follows:*

$$P(Q(A) = 0|A) = 0;$$
$$P(Q(A) = .1|A) = 1/55;$$
$$P(Q(A) = .2|A) = 2/55;$$
$$\vdots$$
$$P(Q(A) = .9|A) = 9/55;$$
$$P(Q(A) = 1|A) = 10/55;$$

(The denominator 55 is chosen because $0 + 1 + 2 + \cdots + 9 + 10 = 55$*.) This likelihood has the form of a step function, and is depicted graphically in Figure 1. Analogously, suppose that P's likelihoods, conditional on* $\neg A$*, take the form of a decreasing step function:*

$$P(Q(A) = 0|\neg A) = 10/55;$$
$$P(Q(A) = .1|\neg A) = 9/55;$$
$$P(Q(A) = .2|\neg A) = 8/55;$$
$$\vdots$$
$$P(Q(A) = .9|\neg A) = 1/55;$$
$$P(Q(A) = 1|\neg A) = 0;$$

Updating on the Credences of Others

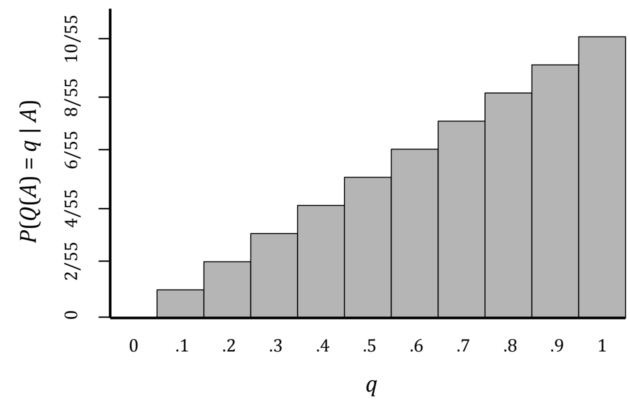

Figure 1: P's credences about Q's credences in A take the form of an increasing step function.

Now, when P learns that Q's credence in A is q, we can plug these likelihoods into Odds *to get:*

$$\frac{P^+(A)}{P^+(\neg A)} = \frac{p}{1-p} \cdot \frac{q}{1-q}. \qquad (Upco^{\ddagger})$$

Letting $O_R(X) =_{def} \frac{R(X)}{R(\neg X)}$*, this formula simplifies to:* $O_{P+}(A) = O_P(A) \cdot O_Q(A)$*. That is, P's posterior odds ratio is just the product of P's prior odds ratio with Q's prior odds ratio.*

We can also plug these likelihoods into Post *to get the corresponding posterior probability:*

$$P^+(A) = \frac{pq}{pq + (1-p)(1-q)}. \qquad (Upco^{\dagger})$$

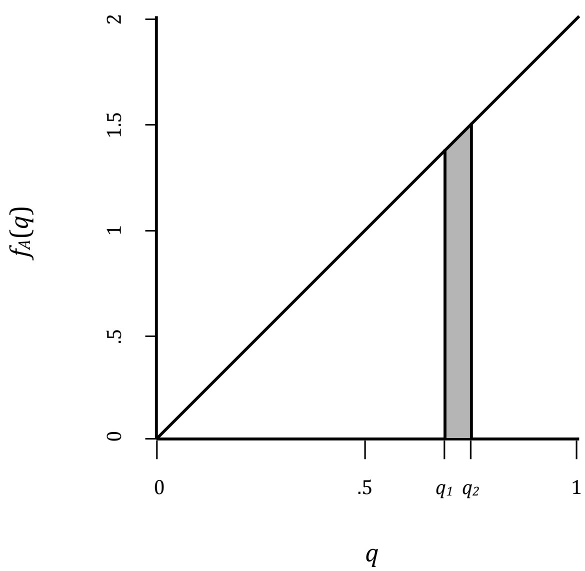

Figure 2: P's credences about Q's credences in A are governed by a linear density function.

This, at last, is our proposed *Upco* rule. More precisely, it is a special case of our *Upco* rule, as indicated by the dagger in the label $Upco^{\dagger}$. $Upco^{\dagger}$ is the formula for the case in which there are two peers, P and Q, who have credences on the simple algebra generated by A, $\neg A$. While conditionalization is computationally demanding, $Upco^{\dagger}$ is extremely simple and can be easily computed. $Upco^{\ddagger}$ is our updating rule expressed in terms of odds ratios. The two formulations are equivalent.

In Example 7.1, we assumed that P's likelihood was discrete. That is, P assigned positive probability only to finitely many values of Q's credences. Although not essential for motivating *Upco*, let us now consider the case where P's likelihoods are continuous.

Example 7.2. *Suppose that P's likelihoods for Q's credences conditional on A are governed by the* density *function f_A, where $f_A(q) = 2q$. This means that:*

$$P(q_1 \leqslant Q(A) \leqslant q_2|A) = \int_{q_1}^{q_2} f_A(q)\, dq$$

This is illustrated in Figure 2. The density function f_A is represented by the diagonal line. The probability $P(q_1 \leqslant Q(A) \leqslant q_2|A)$ is represented by the shaded region — i.e. by the region under the diagonal line representing the function $f_A(q)$, and between the values q_1 and q_2 of q. In the particular case where $f_A(q) = 2q$, integration will yield a value of $q_2^2 - q_1^2$.

Assume, analogously, that P's credences conditional on $\neg A$ are governed by the density $f_{\neg A} = 2(1-q)$.

Now suppose P learns that Q's credence in A is q. Technically, P can't conditionalize on the proposition $Q(A) = q$, since this proposition has probability 0. (Since Q could have any real-valued credence between 0 and 1, the probability of his having any specific real-valued credence q is zero.) There are a couple of possible fixes available. One would be to suppose that P conditions on Q's credence being in some small interval $(q-\epsilon, q+\epsilon)$. A second would be to expand the usual rule of updating by conditionalization by allowing P to use the ratio of the density functions $\frac{f_A(q)}{f_{\neg A}(q)}$ in place of the ratio of likelihoods $\frac{P(Q(A)=q|A)}{P(Q(A)=q|\neg A)}$ when calculating her new odds ratio. In fact, if we calculate the posterior using the first method, and then take the limit as ϵ goes to 0, we will get the same result as the second method. The result is that P will multiply her odds ratio by $\frac{q}{1-q}$, and her new credence will be that implied by $Upco^{\dagger}$, namely $P^+(A) = \frac{pq}{pq+(1-p)(1-q)}$.

In fact, many different likelihoods will yield the $Upco^{\dagger}$ rule. As before, suppose that P's credences about Q's credences in A, conditional on A and $\neg A$, are governed by the density functions f_A and $f_{\neg A}$, respectively. Let them have the following forms:

$$f_A(q) = cqh(q),$$
$$f_{\neg A}(q) = c(1-q)h(q),$$

where h is any strictly positive function on the interval $[0,1]$, and c is a normalization constant ensuring that the total probability adds up to 1. Then conditionalizing on $Q(A) = q$ will yield the same result as $Upco^{\dagger}$. (When calculating the odds ratio, the function $h(q)$ will cancel, as will the constant c.) For example, instead of thinking that Q is likely to report higher credences in A when A is true, P might believe that Q is very likely to be circumspect, and report a credence close to .5. This would be reflected by choosing a function $h(q)$ that is sharply peaked at $q = .5$. As long as P still believes that departures from .5 in the positive direction are more probable than similar departures from .5 in the negative direction when A is true, and vice versa when $\neg A$ is true, and her credences for these departures are in the appropriate ratio, then she will replicate the results of $Upco^{\dagger}$.[18]

We will introduce a quasi-technical term and say that $Upco^{\dagger}$ *mimics* likelihood functions of this form. (Equivalently, $Upco^{\dagger}$ mimics a rational agent who has likelihoods of this form, and it mimics conditionalization using these likelihoods.)

$Upco^{\dagger}$ tells P how to update her credences in A and $\neg A$ when she learns of Q's credence in A (and by inference, in $\neg A$ as well). But P may well have credences in further propositions, besides A and $\neg A$. How should P incorporate her new credences $P^{+}(A)$ and $P^{+}(\neg A)$ into her overall credence function? The most natural suggestion is that P should revise the rest of her credences by *Jeffrey conditionalizing* on her new credences for A and $\neg A$.[19] Thus, if B is an arbitrary proposition to which P assigns a credence, $P^{+}(B) = P(B|A)P^{+}(A) + P(B|\neg A)P^{+}(\neg A)$.

18. Thanks to David James Barnett for extremely helpful discussion about the issues discussed in this paragraph.
19. This proposal is made, e.g., by Wagner (2011).

While this particular proposal is not central to the update rule, it will be relevant to discussions in sections 9.3 and 10.9.

To our knowledge, the formula $Upco^{\dagger}$ first appeared in an article by George Boole in 1857 on the problem of testimony (Boole 1952). It is also a special case of different *multiplicative rules* that have been defended by Morris (1974, 1977), Bordley (1982), Dietrich (2010), and Russell et al. (2015) in the context of the aggregation problem.[20]

With our rule in hand, we can now note some simple properties that it possesses. These can be exhibited (and justified) with reference to a few clear, simple cases.

Example 7.3. *Christensen's doctor case.* $p = .97$ *and* $q = .96$. *Thus, according to* $Upco^{\dagger}$ *the doctors should update to* $p^{+} = \frac{.97 \cdot .96}{.97 \cdot .96 + .03 \cdot .04} \approx .998$.

This shows that $Upco^{\dagger}$ exhibits synergy. Our posterior credence is extremely high in this case because it is extremely unlikely that both doctors have misjudged the situation so badly. That our rule allows for synergy is especially clear when we formulate things in terms of odds ratio $Upco^{\ddagger}$. The odds ratio corresponding to a probability of .97 is approximately 32, and the odds ratio corresponding to a probability of .96 is exactly 24, so the rule tells us that the posterior odds ratio should be about $32 \times 24 = 768$. And in fact, as long as both odds ratios are above 1 (meaning that the probabilities are above .5), it is clear that the posterior odds ratio (and thus the posterior probability) will be higher than either prior.[21]

Example 7.4. *My friend and I are never sure whether Montevideo is the capital of Uruguay or Paraguay. I have credence of .6 that it is Uruguay. My friend also has confidence .6 that it is Uruguay.*

20. See also (Genest and Schervish 1985, 1210) and (Dawid et al. 1995, 281).
21. For results analogous to synergy when multiplicative rules are applied to the aggregation problem, see Morris (1974, 1239–1240), Morris (1977, 692), and Bordley (1982, 1142–1145).

$Upco^\dagger$ says that we should update and move to a credence of ≈ .69. In odds ratio form, our prior odds ratios are both 3/2, so the posterior odds ratio is 9/4, implying a probability of 9/13.

When the two initial credences are the same, as in the last example, any form of averaging tells us to stay put as if the other person's report is no evidence at all. In discussing necessary conditions on good aggregation rules, Lehrer and Wagner (1983) say "A restriction which comes immediately to mind is that F [the amalgamation function] ought to respect a consensus on the probabilities assigned to any atomic proposition." By this they mean that the credence pair (x, x) ought to result in agents remaining at credence x.[22] We do not think that this is a good rule for updating, since it means that you can never treat your peer's agreement as a form of evidence. Our view allows one to treat the other agent's credence as additional evidence, even if it happens to be numerically the same as one's own initial credence. Our rule does imply three special cases of preservation of the first agent's credences at $(0, y)$ (for $y < 1$), $(x, .5)$, and $(1, y)$ (for $y > 0$).[23] But cases of the form (x, x), where $x \neq 0, .5$, or 1, exhibit synergy. Of course, according to $Upco^\dagger$, the peers will have identical post-update credences, but the posterior credences that they share will be different from either of their priors. We think it is clear in these simple cases that our rule intuitively does better than linear averaging.

Example 7.5. *P's credence in A is $P(A) = p$, and Q's is $Q(A) = q = 1 - p$. $Upco^\dagger$ yields $P^+(A) = \frac{p(1-p)}{p(1-p)+(1-p)p} = .5$. Due to the symmetry of the peers' opinions toward A and $\neg A$, they end up with credences of .5 in each proposition.*

22. Here and throughout the paper "credence pair (x, y)" means that one agent has a credence of x while a second agent has a credence of y in the same proposition.

23. Compare Morris (1974, 1239–1240) and Bordley (1982, 1142–1145) for the analogous result in multiplicative judgment aggregation. Since our rule exhibits a built-in symmetry between agents, the *second* agent's credence will likewise be preserved at $(x, 0)$ (for $x < 1$), $(.5, y)$, and $(x, 1)$ (for $x > 0$).

In this special case, our rule happens to exactly agree with the rule of linear averaging. In fact, any rule that is symmetric among the peers P and Q (satisfying what Russell et al. (2015) call *anonymity*), and in the propositions A and $\neg A$ (satisfying what Russell et al. [2015] call *neutrality*), will have this result. Since these symmetric cases make up a large proportion of the examples considered by other philosophers, we think that they have misled people about the general case and, in particular, into thinking that the EWV should be formalized by linear averaging (with equal weights). Discussions of binary belief cases of disagreement between peers tend to focus on cases where your peer is "equally confident" in A as you are in $\neg A$. (For examples, see Elga [2007], Christensen [2007], and Kelly [2010, 122].)

Example 7.6. *P has a credence in A of $P(A) = p$, and Q has credence $Q(A) = .5$. Then $Upco^\dagger$ yields $P^+(A) = \frac{p(.5)}{p(.5)+(1-p)(.5)} = p$.*

If one of the agents has an odds ratio of 1 (and thus a credence of .5), then the posterior is identical to the prior of the other agent. $Upco^\dagger$ mimics a likelihood in which a peer's having a credence of .5 is equally likely whether the proposition is true or false. Thus, a credence of .5 in A provides no evidence for or against A. For this reason, we will say that .5 is a *break-even* point. (In Section 10, we give a generalization of $Upco$ that allows for other break-even points.)[24]

Example 7.7. *P has a credence in A of .4, and Q has a credence of .9. Then $Upco^\dagger$ yields $P^+(A) = \frac{(.4)(.9)}{(.4)(.9)+(.6)(.1)} \approx .86$.*

In this example, P moves almost all the way to Q's point of view, while Q changes very little.

While $Upco^\dagger$ adheres to the EWV in the sense of satisfying *anonymity*, it produces very different results depending upon the credences of the peers P and Q. In some cases, it will lead to P moving halfway toward Q's credence; in others it will lead to P's deferring to

24. See also Morris (1974, 1239–1240), Morris (1977, 692), and Bordley (1982, 1142–1144).

Q's credence; in yet others, it will lead to P steadfastly holding on to her original credence; and there will also be cases that produce synergy, where P moves *past* Q's original credence.

7.2 *Multiple Peers*

The rule $Upco^\dagger$ generalizes quite straightforwardly to interactions with more than one peer. Suppose that there are n peers, $P_1, P_2, \ldots, P_n$, with $P_i(A) = p_i$ for all i. Then the natural generalization of $Upco^\dagger$ is:

$$P^+(A) = \frac{p_1 p_2 \ldots p_n}{p_1 p_2 \ldots p_n + (1-p_1)(1-p_2)\ldots(1-p_n)}. \qquad (Upco^\star)$$

Or equivalently:

$$O_{P+}(A) = O_{P_1}(A) \cdot O_{P_2}(A) \cdot \ldots \cdot O_{P_n}(A). \qquad (Upco^*)$$

The star indicates that this is still a special case of our most general rule, since we are still assuming that the credences of the peers are on a simple algebra generated by the propositions A, $\neg A$. $Upco^\star$ gives the credence that would result from P_1 conditionalizing on the credences of $P_2, \ldots, P_n$, if her likelihood function has two features. First, for each i, her credence that P_i reports a credence of p_i, conditional on A being true, is proportional to p_i. That is, P_1's likelihood function for each of her peers has the same linear form discussed above.[25] Second, her credences for different peers are independent, conditional on A and $\neg A$. Thus, conditional on A being true, P_1's credence that P_i and P_j will report credences in A of p_i and p_j is proportional to $p_i p_j$. Note that, on this assumption, P_1's credences for P_i and P_j will not be *unconditionally* independent. If P_i reports a high credence for A, that provides evidence that A is true, which in turn makes it more likely that P_j will report a high credence for A.

25. As noted in the previous section, $Upco^\dagger$ will also be consistent with conditionalization using a range of different likelihoods as well. That holds true of $Upco^\star$ too.

We will illustrate some of the properties of $Upco^\star$ by means of example. The first example involves peers with the same credences as in Example 4.1 above:

Example 7.8. *Peers P, Q, and R have credences in A of .3, .5, and .8, respectively. P meets Q, updates on his credence, then meets R, and updates on hers. After meeting Q, P will have credence* $P^+(A) = \frac{(.3)(.5)}{(.3)(.5)+(.7)(.5)} = .3$. *This is an example of the phenomenon we saw in example 7.6. Then, upon meeting R, P will have credence* $P^{++}(A) = \frac{(.3)(.8)}{(.3)(.8)+(.7)(.2)} \approx .63$. *But suppose that P had met R first. Then she would first have* $P^+(A) = \frac{(.3)(.8)}{(.3)(.8)+(.7)(.2)} \approx .63$. *And upon meeting Q, her credence will remain unchanged (as we saw in 7.6). Thus P's final credence does not depend upon the order in which she meets her peers. Readers can also verify that P will arrive at the same credence if she updates by* $Upco^\star$ *on both peers at once.*

This property holds in general. Unlike linear averaging, $Upco^\star$ doesn't depend on the order in which one responds to the credence of peers. On the odds ratio form of our rule, $Upco^*$, one updates upon meeting each peer by multiplying one's current odds ratio by the odds ratio of that peer. The fact that multiplication is commutative means that it doesn't matter in what order one updates in light of one's peers, and the fact that multiplication is associative means that it doesn't matter if one updates in light of two peers at once or in sequence.

$Upco^\star$ also permits another kind of sequential updating.

Example 7.9. *Peers P, Q, and R have credences in A of* .8, .6, *and* .6, *respectively. Q meets R and updates his credence using* $Upco^\star$, *so that* $Q^+(A) \approx .69$. *Then P meets Q, and updates using* $Upco^\star$. *P's new credence* $P^+(A)$ *will be .9. This is the same credence that would have resulted had P updated on the original credences of Q and R.*

In this example, P still gets the benefit of learning from R's opinion, since R's opinion is reflected in Q^+. However, if P meets Q and R *after* they have updated on each other's opinion, and uses $Upco^\star$, she will have a new credence of $P^+(A) \approx .95$. This is the same credence that would result from meeting *four* peers with credences of .6 for A.

Since R's original opinion is now reflected in Q's new opinion, and vice versa, the original credences of Q and R are being double-counted. $Upco^\star$ is designed to mimic a likelihood in which the credences of Q and R are *independent* conditional on A and $\neg A$. That is not reasonable in this case, where Q and R have synchronized their credences. We will return to this problem in section 10.6 below.

Example 7.10. *There are $m + n$ peers. m have a credence in A of p, while the other n have a credence of $1 - p$. Then, when the peers update on each other's credences, they will arrive at a new credence of* $P^+(A) = \frac{p^m(1-p)^n}{p^m(1-p)^n + (1-p)^m p^n} = \frac{p^{m-n}}{p^{m-n} + (1-p)^{m-n}}$. *This depends only on the difference $m - n$. Moreover, the result will often be synergistic. For example, if $p > .5$, and $m - n > 1$, the final credence will be higher than p.*

This result is similar to the famous "Condorcet Jury Theorem". If there is a jury of individuals that have imperfect but positive reliability on some proposition, then the result of a majority vote is more reliable than any of the individual jurors. In our framework, if each agent has a reliability of $p > .5$, then we can represent them all as having credence either p or $(1 - p)$ in the proposition. That is, we can represent a juror who votes for A as an epistemic peer with a credence in A equal to p; and we can represent a juror who votes against A as a peer with credence $1 - p$. If m of them claim that the proposition is true, and n of them claim that it is false, then the posterior odds ratio will be O^{m-n} where $O = p/(1 - p)$. The final degree of confidence we should have depends on the *difference* between the number of jurors that vote for or against the proposition, and will generally be substantially *higher* than the confidence of any individual juror. People often find this result counterintuitive. For example, suppose that each individual juror is 80% reliable. Then, if ten jurors unanimously judge that A is true, the probability of A is the same as if 1,010 jurors judged that A is true, and 1,000 judged that it is false. This result is normatively correct (on the assumption that the jurors are independent), although many people would (without calculating) judge A to be more probable in the first case. Linear averaging cannot capture the results of the Condorcet theorem. On linear averaging, the posterior will depend on the *ratio* between the number that vote for and against, and will generally be substantially *lower* than the maximum confidence to be found among the individual jurors.

Note that the $Upco^\star$ rule will break down if the p_i's include both 1's and 0's (we will get a posterior credence of 0/0).[26] However, the likelihood that is mimicked by our rule says this happens with probability zero.[27] Suppose, however, that some number of peers have credences of ϵ, and others have credences of $1 - \epsilon$, where ϵ is small. Taking limits as ϵ goes to 0, we see that an equal number of 0's and 1's will cancel out, but if one is in excess of the other, then it will dominate. In this regard, our formula works like Condorcet's: six 1's and five 0's will give the same result as a single 1.

8. Updating on Credences over Partitions

The rule as we have proposed it so far is useful when the question under discussion involves only a single proposition and its negation. But this is importantly limited. Imagine that you have no idea when Bob's birthday is. Your own credence is (roughly) 1/365 for each possible day. Alice then tells you that it is either April 8th or April 9th — she has credence .5 in each. If you just focus on whether or not it is April 8th and apply $Upco^\dagger$, you will get $P^+(April8) = \frac{(1/365)(.5)}{(1/365)(.5)+(364/365)(.5)} = 1/365$, which is clearly the wrong answer. You have treated Alice's report as if it contained no useful information. Even worse, it is not clear this procedure is even coherent. Focusing on a different proposition — such as whether Bob's birthday is December 25th — leads to an incompatible result. $Upco^\dagger$ and $Upco^\star$ are partition-dependent rules.

26. For the analogous result when a multiplicative rule is applied to the aggregation problem, see Bordley (1982, 1145).

27. Since peer P_1 has a likelihood *density* for the credences of her peers, she assigns probability zero to to her peers reporting any specific credence. But the case described in the text has probability zero in a stronger sense: the density itself will be zero.

In the present section, we show how to properly extend these rules to credences over a partition $\{A_1, A_2, \ldots, A_k\}$ instead of a binary choice $\{A, \neg A\}$.

Suppose that we have a partition of n propositions $\{A_1, A_2, \ldots, A_k\}$. P has probabilities $p_1, p_2, \ldots, p_k$ while her peers have $q_1, q_2, \ldots, q_k; r_1, r_2, ..., r_k$; etc. The natural generalization of our rule is:

$$P^+(A_i) = \frac{p_i q_i r_i \ldots}{\sum_{j=1}^{k} p_j q_j r_j \ldots} \qquad (Upco)$$

Since this (at last!) is the fully general version of our rule, the name *Upco* appears unadorned.

When there are many propositions involved, there is an odds ratio for each given pair. The rule says that the odds ratio between any two elements of the partition after the update is the product of the odds ratios each peer had between those two elements of the partition before the update. That is:

$$\frac{P^+(A_i)}{P^+(A_j)} = \frac{p_i}{p_j} \cdot \frac{q_i}{q_j} \cdot \frac{r_i}{r_j} \ldots$$

The general version of *Upco* mimics the result of conditionalization with likelihoods similar to those we have encountered before, although a bit more complicated. Suppose that P's likelihoods for each peer's credence, conditional on A_i, are governed by a density f_i having the following form:

$$f_i(q_1, \ldots, q_k) = \begin{cases} k! q_i, & \text{if } q_1 + \ldots + q_k = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Suppose also that P takes the credences of her peers to be independent, conditional on each A_i. Then, the result of updating by conditionalization on the credences of her peers will be the same as the result of computing using *Upco*. Note that f_i yields a joint likelihood for a complete credence distribution for a given peer, rather than a separate likelihood for the peer's credence in each proposition.

Upco will be consistent with conditionalization on other likelihoods as well. It suffices if P's likelihoods are governed by densities of the form:

$$f_i(q_1, \ldots, q_k) = \begin{cases} c_Q h_Q(q_1, \ldots, q_k) q_i, & \text{if } q_1 + \ldots + q_k = 1 \\ 0, & \text{otherwise;} \end{cases}$$

where h_Q is a strictly positive function that may vary for different peers, but not for different members of the partition, and c_Q is a normalization constant. [28]

We will now illustrate some of the properties of *Upco*.

Example 8.1. *We have a partition with three propositions* $\{A_1, A_2, A_3\}$. *Peer P has probabilities* $(.5, .3, .2)$ *while Q has probabilities* $(.4, .3, .3)$. *Then the result of updating will be approximately* $(.57, .26, .17)$.

28. This form of likelihood is applicable in cases where the likelihood of the other person's credences has a particular sort of symmetry. As a toy example where this might fail, consider a case where two peers are both calculating the amount they owe on splitting the bill at a restaurant. Let A_n be the hypothesis that they owe \$$n$. This form of likelihood means that the first agent thinks that a credence function for Q with $q_{43} = .01, q_{44} = .05, q_{385} = .9, q_{386} = .05$ is equally likely whether A_{44} or A_{386} is true. An important part of the argument of Vavova (2014) involves noticing that we should find this set of credences incredibly implausible if A_{44} is true, but not so implausible if A_{386} is true. Thus, it is likely that our proposal makes more sense in cases where the partition doesn't exhibit this sort of asymmetry — an agent is no more likely to concentrate her credence on one false hypothesis than another, perhaps because each hypothesis differs from others in similar ways. This may be a useful approximation in cases where we can assume that the range of hypotheses is just between A_{40} and A_{45}, but is probably not a good approximation in cases like the one described.

This example illustrates that synergy still arises. In fact, it can often be even more pronounced than in the case of binary propositions. If the peers involved agree about which element of the partition has the highest probability, then this element of the partition will have an even higher probability after the update than before. Even if all agents have credence less than .5 in a proposition, it can still go up in credence for each agent.

Example 8.2. *P and Q have credences over the partition $\{A_1, A_2, \ldots, A_k\}$. Q assigns each member of the partition a credence of $1/k$. Then when P updates on Q's credences, her own credences will remain unchanged.*

One should be careful to note that, while credence $1/k$ for each element of a partition of k elements is the break-even *distribution*, it does not follow that $1/k$ is a break-even credence for one element of the partition taken individually. In the case of a partition with > 2 elements, there is no such thing as a break-even credence *for one element*. Rather break-even credences are really distributions over the whole partition. It is possible for every peer to have probability $1/k$ in one element of the partition and the updated group probability to be either higher or lower than this value, as the next two examples illustrate.

Example 8.3. *P and Q have credences over the three-element partition $\{A_1, A_2, A_3\}$. P has credences $(1/2, 1/3, 1/6)$, and Q has credences $(1/6, 1/3, 1/2)$. After the update, the credences are proportional to $(1/12, 1/9, 1/12)$, which makes them $(.3, .4, .3)$ — the second proposition had credence $1/3$ for both agents, but afterwards it has gone up.*

Example 8.4. *P and Q both have credences $(1/2, 1/3, 1/6)$ over the partition $\{A_1, A_2, A_3\}$. Then when they update, the posterior is proportional to $(1/4, 1/9, 1/36)$, which gives approximately $(.64, .29, .07)$. In this case, the shared value of $1/3$ has gone down to $.29$.*

These illustrations also show that *Upco* is *partition-dependent*. In both 8.3 and 8.4, the peers had exactly the same credences over the partition $\{A_2, \neg A_2\}$, namely $(1/3, 2/3)$. Thus, if P were to update on Q's credences on the partition $\{A_2, \neg A_2\}$, she would get the same result in both 8.3 and 8.4. However, as we have seen, when P updates on Q's credences over the finer partition $\{A_1, A_2, A_3\}$, she gets a different result in the two examples. We will discuss the phenomenon of partition-dependence in greater detail in Sections 9.1 and 10.9.

Example 8.5. *P and Q have credences over the partition $\{A_1, \ldots, A_k\}$. However, Q is befuddled, and his credences $q_1, \ldots, q_k$ do not sum to one. (Technically, then, they are not credences, but merely "weights" or "degrees of belief". We will ignore this terminological nicety here.) Nonetheless, we can apply Upco directly to P's and Q's credences, so that $P^+(A_i) = \frac{p_i q_i}{\sum_{j=1}^{k} p_j q_j}$. P^+ will be a probability function. Moreover, P^+ will be the same probability function we would have ended up with had we first normalized Q's credences and then updated.*

The fact that we can combine unnormalized weights in this way will prove useful in many of the modifications we will consider in Section 10.

We will discuss some more features and modifications of *Upco* in the remaining sections of the paper.

9. Other Properties of Updating Rules

In this section, we aim to compare *Upco* to linear averaging by examining more general properties that update rules might have.

9.1 The Context-Free Assumption

As before, let $P(A)$ and $Q(A)$ be the credences that P and Q initially assign to A, and let $P^+(A)$ and $Q^+(A)$ be the credences that P and Q assign to A after updating on each other's credences (and nothing more). One principle that has been proposed for rules for combining the credences of different agents is that there should be some function f such that for all P, Q, and A, $P^+(A) = f(P(A), Q(A))$. That is, the resulting credence in a given proposition should depend only on the credences that those two agents have in that proposition, regardless of which proposition, A, is involved, or what credences either of the two agents has in various other propositions. Versions of this assumption have been called "irrelevance of alternatives" (Jehle and Fitelson 2009), "the strong setwise function property" (McConway 1981), "strong label neutrality" (Wagner 1982), and the "context-free assumption" (Bordley and Wolff 1981). We'll abbreviate the latter phrase, calling this principle CFA.

Upco satisfies CFA in the special case of a two-element partition into A and $\neg A$: the same formula is applied to the credences in A and the credences in $\neg A$. But CFA is relatively trivial in this case. In the case of a partition with more than two members $\{A_1, \ldots, A_k\}$, *Upco* satisfies a weaker condition that Russell et al. (2015) call *neutrality*. This condition says that permuting the order of the A_is doesn't change the way in which the rule is applied. Specifically, it says that if we apply the same permutation to the credences of each peer and then use *Upco*, we get the same result as if we first use *Upco* and then permute the resulting credences.

As we saw in Section 8, however, our rule doesn't obey CFA in general. In fact, it violates CFA in two different ways. Suppose P and Q have credences over the partition $\{A_1, \ldots, A_k\}$. Then the updated credence in A_1 will not just depend upon $P(A_1)$ and $Q(A_1)$, but will depend upon the distribution of the credences over the other members of the partition. Second, the formula in *Upco* applies only to the members of the partition, and not to other propositions such as $A_1 \vee A_2$. In fact, McConway (1981) and Wagner (1982) independently proved that weighted linear averages are the only possible rules for combining two probability functions into a single probability function that satisfy CFA.

However, we think that this is not a problem for our rule. CFA is far too strong. For a clear counterexample to it, imagine a pair of examples where you and your peer witness a horse race with three horses, A, B, and C. In Case 1, both you and your peer have $P(A) = P(B) = P(C) = 1/3$. After learning each other's credences, neither of you should make any changes, and so you both remain at (1/3, 1/3, 1/3). In Case 2, you and your peer have independent, strong evidence in favor of horse B winning at the expense of horse C. You both assign $P(A) = 1/3, P(B) = 2/3 - \epsilon, P(C) = \epsilon$, for some small ϵ. Here, intuitively, we should see some synergistic effects, and both you and your peer should increase your confidence that horse B was the winner. If ϵ is small enough, then the only way to do that is at the expense of horse A. Our rule happens to recommend that you each end up at credences proportional to $(1/9, [2/3 - \epsilon]^2, \epsilon^2)$ or approximately $(.2, .8, 0)$, but any plausible rule will have to decrease the posterior credence in A. In both cases we had each agent's prior probability in A equal to 1/3. But, in the first case, their credences in A stay the same; in the second case, they both went down. This was entirely appropriate. Thus CFA is a bad principle.

9.2 *Preservation of Independence*

A second principle that has been considered for updating rules is the preservation of independence PI. The basic version, debated (for instance) in Laddaga (1977) and Lehrer and Wagner (1983), says: For any A, B, if, for all i, $P_i(A \wedge B) = P_i(A) \cdot P_i(B)$, then, for all i, $P_i^+(A \wedge B) = P_i^+(A) \cdot P_i^+(B)$. (Note that PI is relevant only when updating beliefs over partitions of more than two elements, since in an algebra generated by $A, \neg A$, no propositions other than the tautology or the contradiction can be independent of anything.) A stronger requirement, proposed by Jehle and Fitelson (2009), says: For any A, B, C, if, for all i, $P_i(A \wedge B|C) = P_i(A|C) \cdot P_i(B|C)$, then, for all i, $P_i^+(A \wedge B|C) = P_i^+(A|C) \cdot P_i^+(B|C)$. This is the preservation of *conditional* independence (PCI) (from which (PI) is derived by letting C be the tautology).

We noted in Section 4 above that linear averaging does not satisfy PI, except in rare cases. Lehrer and Wagner (1983) respond by claiming that PI is too strong a principle to accept. In particular, they show that it, together with zero preservation (the principle that if both agents have credence 0 then the updated credence will also be 0) and CFA, entail that an update rule must be dictatorial. That is, the final credences are always equal to the credences of one of the initial agents. Indeed, it is impossible to satisfy PCI together with CFA. For a survey of many such results, see Genest and Zidek (1986). Of course, since CFA is too strong, these arguments against PI/PCI may not be good ones.

As the next example illustrates, *Upco* preserves independence in many cases where linear averaging doesn't.

Example 9.1. *Let us return to the coin-flipping case of Example 4.2. Two coins will be tossed. P has a credence* $1/3$ *of each coin coming up heads, and takes the flips to be independent. Thus,* $P(HH) = 1/9$, $P(HT) = P(TH) = 2/9$, *and* $P(TT) = 4/9$. *Q has a credence* $2/3$ *for heads on each flip, and also takes them to be independent. Thus,* $Q(HH) = 4/9$, $Q(HT) = Q(TH) = 2/9$, *and* $Q(TT) = 1/9$. *If P updates using* $Upco$, *her new credence for each of the four propositions —* HH, HT, TH, *and* TT *— will be proportional to the products of the initial probabilities the individual agents have for them. In this case, these products are all equal to* $4/81$, *so* $P^+(HH) = P^+(HT) = P^+(TH) = P^+(TT) = 1/4$. *P's new credences continue to exhibit independence for the two flips.*

This is not just a coincidence due to the symmetrical numbers, as the next theorem shows.

Theorem 9.2. *Consider two partitions* $\{A_1, \ldots, A_k\}$ *and* $\{B_1, \ldots, B_r\}$, *and the joint partition generated by conjunctions* $A_i \wedge B_j$. *Let P and Q be two agents who each think the two partitions are independent, so that* $P(A_i \wedge B_j) = P(A_i) \cdot P(B_j)$ *and* $Q(A_i \wedge B_j) = Q(A_i) \cdot Q(B_j)$ *for all* i, j. *Then combining the probability functions over the joint partition using* $Upco$ *gives a probability function such that the two partitions are still independent. That is,* $P^+(A_i \wedge B_j) = P^+(A_i) \cdot P^+(B_j)$.

Proof. Let $P(A_i \wedge B_j) = p_{ij}$ and $Q(A_i \wedge B_j) = q_{ij}$, and let $S = \sum_{i=1}^m \sum_{j=1}^n p_{ij}q_{ij}$. Then our rule says $P^+(A_i \wedge B_j) = \frac{p_{ij}q_{ij}}{S}$. Let $P(A_i) = p_{i*}$ and $P(B_j) = p_{*j}$, and similarly let $Q(A_i) = q_{i*}$ and $Q(B_j) = q_{*j}$. Then the assumption of independence of the two partitions for P and for Q means that $p_{ij} = p_{i*}p_{*j}$ and $q_{ij} = q_{i*}q_{*j}$.

Thus, $P^+(A_i \wedge B_j) = \frac{p_{i*}q_{i*} \cdot p_{*j}q_{*j}}{S}$. Summing the conjuncts, we see that $P^+(A_i) = \frac{p_{i*}q_{i*}}{S} \sum_{j=1}^n p_{*j}q_{*j}$ and $P^+(B_j) = \frac{p_{*j}q_{*j}}{S} \sum_{i=1}^m p_{i*}q_{i*}$. Multiplying, we see that $P^+(A_i) \cdot P^+(B_j) = \frac{p_{i*}q_{i*}}{S} \frac{p_{*j}q_{*j}}{S} \sum_{i=1}^m p_{i*}q_{i*} \sum_{j=1}^n p_{*j}q_{*j}$. But the last two sums can be multiplied out to give $\sum_{i=1}^m \sum_{j=1}^n p_{i*}p_{*j}q_{i*}q_{*j} = \sum_{i=1}^m \sum_{j=1}^n p_{ij}q_{ij} = S$. Thus, we get $P^+(A_i) \cdot P^+(B_j) = \frac{p_{i*}q_{i*}}{S} \frac{p_{*j}q_{*j}}{S} S = \frac{p_{i*}q_{i*} \cdot p_{*j}q_{*j}}{S} = P^+(A_i \wedge B_j)$. □

Thus, when there are two partitions that are agreed upon as being independent by two agents, this independence will be preserved when the agents update over the joint partition generated by these two partitions. This is important for many cases like the coin flip example. For instance, this is often the situation when we are considering the rolls of two dice, or when we are considering two questions, such as 'What is the capital of Uruguay?' and 'Which country will win the FIFA World Cup in 2018?' For each agent individually, the answer to one question is unrelated to the answer to the other question; this should hold for the updated credence as well.

However, our rule doesn't preserve independence of individual pairs of propositions in general. For instance, we can have two partitions, $A, \neg A$ and B_1, B_2, B_3. If two agents both think that A is independent of B_1, but one thinks that A is positively correlated with B_2 while the other thinks that A is positively correlated with B_3, then the result of combining the probability functions will be a distribution on which A is correlated (perhaps positively or negatively) with B_1. Lehrer and Wagner (1983), Genest and Wagner (1987), and Wagner (1985, 2011) argue that these cases are ones in which it is not advantageous to preserve independence, since the independence here is, in a sense, accidental. However, while Wagner uses these cases as motivation to give up any sort of general preservation of independence, we think that there is good reason to preserve independence at least in cases involving full partitions. Full preservation of independence would be too much, but linear averaging preserves too little. And importantly, our rule preserves independence in cases where it ought to do so without the sort of ad hoc maneuvers that Wagner appeals to.

We should also note that the preservation of independence is partition-relative. For example, suppose that we have three partitions $\{A_1, \ldots, A_k\}$, $\{B_1, \ldots, B_r\}$, and $\{C_1, \ldots, C_s\}$. Peers P and Q both think that the A's are independent of the B's. Then, as we have seen, if they update on each other's credences over the partition generated by the A's and the B's, the A's and B's will remain independent in the revised probability distribution. However, if they update on each other's credences over the partition generated by *all three* original partitions, A's, B's, and C's, the revised probability may not preserve the independence of the A's and B's. This can happen if the C's are correlated with *both* the A's and the B's in each peer's credences. (This is consistent with the A's and B's being independent. It can happen, for example, if the A's and B's are independent causes of the C's.) Synergistic effects may increase the probabilities of some of the C's, and this will create a correlation between the A's and the B's.

9.3 *Relationships with Conditionalization*

Some other proposed principles for combining probability functions involve the relation between a rule proposed for doing so and the standard Bayesian rule of update by conditionalization. Since any rule for updating on learning another agent's credences is an update rule, it is natural to ask under what circumstances the proposed rule agrees or disagrees with conditionalization. It is also natural to ask how the rule interacts with conditionalization on other pieces of evidence.

As we have seen, *Upco* is specifically designed to yield the same result as conditionalization on the credences of one's peers, if one's likelihoods have a particular form. By contrast, we saw in Section 4 that linear averaging is not compatible with conditionalization in general.

We also would like to see how an update rule interacts with conditionalization. A natural thought is that updating on the credences of one's peers should commute with conditionalization. That is, it should not matter whether one first updates on the credences of one's peers and then conditionalizes on some proposition, or whether one first conditionalizes and then updates on the credences of the peers. As we saw in Section 4, linear averaging does not commute with conditionalization. But does *Upco* commute with conditionalization?

The answer, essentially, is 'yes'. But the details take some care to state. Suppose that we have peers P, Q, R, ... The peers have credences over the partition $\{A_1, \ldots, A_k\}$, and also for the proposition E. Consider the following three procedures:

1. P updates on her peers' credences over the partition $\{A_1, \ldots, A_k\}$. Then she learns E and conditionalizes on it.
2. P learns E and conditionalizes, but her peers do not. P then learns her peers' credences over $\{A_1, \ldots, A_k\}$ and updates according *Upco*.
3. P learns E and conditionalizes, and so do all of her peers. P then learns her peers' posterior credences over $\{A_1, \ldots, A_k\}$ and updates according to *Upco*.

When we ask whether *Upco* commutes with conditionalization, we might be asking whether procedures 1 and 2 yield the same outcome, or we might be asking whether procedures 1 and 3 yield the same outcome. The answers need not be the same. Most authors have understood commutativity to mean that procedures 1 and 3 (or their analogs for other updating rules) are the same. The requirement that an updating rule commute with conditionalization in this sense has been called *external Bayesianism* (see, e.g., McConway 1981, Dietrich 2010, and Russell et al. 2015.) But the equivalence of 1 and 2 seems like a perfectly sensible thing to mean by the commutativity of conditionalization and *Upco*.[29]

Now we need to consider two separate cases:

- The proposition E is equivalent to some disjunction of the A_i's;
- E is not equivalent to such a disjunction.

29. Dietrich 2010 calls the equivalence of procedures 2 and 3 *internal Bayesianism*.

In the first case, the proposition E can be expressed in the "language" of the partition $\{A_1, \ldots, A_k\}$. One particularly simple case of this sort occurs when the partition is of the form $\{AE, A\neg E, \neg AE, \neg A\neg E\}$. When E is equivalent to a disjunction of A_i's, $Upco$ commutes with conditionalization in *both* senses described above. That is, procedures 1, 2, and 3 all yield the same result. Intuitively, this is because once anyone learns E, they completely eliminate the A_i that are incompatible with it, and don't change anything else about their distribution. Then, when the products are taken in $Upco$, the elements of the partition that are incompatible with E will receive a new weight of zero. Thus, it doesn't matter whether one person or many learn E. On the other hand, if E is not equivalent to some disjunction of the A_i's, then $Upco$ will commute with conditionalization in the sense that procedures 1 and 2 yield the same result, but not in the sense that 1 and 3 yield the same result. Intuitively, this is because once Q, R, . . . conditionalize on E, their new credences over the partition $\{A_1, \ldots, A_k\}$ will reflect their prior evidence as well as their beliefs about how E is correlated with each member of the partition, so procedure 3 will count each of these correlations as additional evidence. However, in procedures 1 and 2, this information does not get incorporated into P's credence over $\{A_1, \ldots, A_k\}$.

We now provide proofs of these results.

Theorem 9.3. *Suppose that peers P, Q, R, . . . have credences over the partition $\{A_1, \ldots, A_k\}$. Suppose also that the proposition E is a disjunction of A_i's. Then procedures 1, 2, and 3 above all yield the same outcome.*

Proof. Let A_i be any element in the partition that is being updated on, let P's initial credence in A_i be $P(A_i) = p_i$, Q's initial credence in A_i be $Q(A_i) = q_i$, and so on for the other peers. Define p_j, q_j, and so on analogously. Since E is a disjunction of elements in the partition, either A_i entails E or it is incompatible with E. If it is incompatible with E, then on all three procedures, it ends up with probability 0. Whenever P conditionalizes on E, her credence in A_i will go to zero. Now suppose A_i and A_j are both compatible with E. Then, A_i and A_j must both entail E. In this case, conditionalizing on E will not affect the ratio of the probabilities of A_i and A_j. For example, if P conditionalizes on E, we will have $\frac{P(A_i)}{P(A_j)} = \frac{P(A_i|E)}{P(A_j|E)}$. This is true for P's credences regardless of whether she has updated on the credences of her peers, and it is true for her peers' credences as well.

Now consider how the ratio of peers' credences in A_i and A_j will evolve in each of the procedures described above. On procedure 1, P updates on the credences of her peers first, so the ratio becomes $\frac{P^+(A_i)}{P^+(A_j)} = \frac{p_i q_i r_i \cdots}{p_j q_j r_j \cdots}$. Then, when P conditionalizes on E, this ratio will remain unchanged.

On procedure 2, P first conditionalizes on E, which leaves the ratio unchanged. Then, she will update on the credences of her peers, yielding a ratio of $\frac{p_i q_i r_i \cdots}{p_j q_j r_j \cdots}$.

On procedure 3, all of the peers update on E, which leaves the ratio of their credences unchanged. Then, when P updates on the posterior credence of her peers, the ratio will again be $\frac{p_i q_i r_i \cdots}{p_j q_j r_j \cdots}$.

But P's final credence over the partition $\{A_1, \ldots, A_k\}$ is entirely determined by the ratios of her credences in the members of the partition that get non-zero probability. So P's credence over the partition will be the same in all three procedures. Specifically, her final credence for an arbitrary member of the partition A_i is $P^+(A_i) = \frac{\delta_i p_i q_i r_i \cdots}{\sum_{j=1}^k \delta_j p_j q_j r_j \cdots}$, where δ_j is a function that takes the value 1 if A_j is compatible with E, or 0 otherwise. □

This proof generalizes straightforwardly to the following corollary:

Corollary 9.4. *Suppose that peers P, Q, R, . . . have credences over the partition $\{A_1, \ldots, A_k\}$. Suppose also that the proposition E is a disjunction of A_i's. Suppose that some of P's peers learn E and conditionalize, while others do not. P then updates on the credences of some of her peers. Then P learns E and conditionalizes. Finally, P updates on the credences of her remaining peers. P's final credences over the partition $\{A_1, \ldots, A_k\}$ will be the same as if she had followed any of procedures 1 through 3, namely $P^+(A_i) = \frac{\delta_i p_i q_i r_i \cdots}{\sum_{j=1}^k \delta_j p_j q_j r_j \cdots}$.*

Theorem 9.5. *Suppose that peers P, Q, R, ... have credences over the partition $\{A_1, \ldots, A_k\}$, and also for the proposition E, which is not a disjunction of A_i's. Then procedures 1 and 2 above yield the same outcome, but procedure 3 need not.*

Proof. As with the previous proof, it suffices to consider the ratio of P's credences in arbitrary elements of the partition A_i, and A_j. Let P, Q, p_i, p_j, q_i, q_j,... be defined as above. Let p'_i be the likelihood $P(E|A_i)$, and likewise for $p'_j, q'_i, q'_j, \ldots$. In the case where E is a disjunction of partition elements, these likelihoods are all 0 or 1, but in the present case, they need not be.

On procedure 1, P learns her peers' initial credences and updates, giving her credence P^+. Thus $\frac{P^+(A_i)}{P^+(A_j)} = \frac{p_i q_i r_i \cdots}{p_j q_j r_j \cdots}$. Since E is not in the algebra defined by the partition, P now needs to update her credence in E by Jeffrey conditionalizing with her new credences over the partition, $\{P^+(A_1), \ldots, P^+(A_k)\}$. Call her new credence P^{++}. Note that Jeffrey conditionalization will leave the conditional probability of E on each member of the partition unchanged, hence $P^{++}(E|A_i) = p'_i$, and likewise for other members of the partition. Now P learns E, and conditionalizes, giving her a final credence of P^{final}. By the odds formulation of Bayes' theorem $\frac{P^{final}(A_i)}{P^{final}(A_j)} = \frac{P^{++}(A_i)p'_i}{P^{++}(A_j)p'_j} = \frac{p'_i p_i q_i r_i \cdots}{p'_j p_j q_j r_j \cdots}$. Note that P's likelihoods p'_i and p'_j appear in this formula, but the likelihoods of the other peers do not.

On procedure 2, P learns E and conditionalizes, giving her a new credence P^+. By the odds formulaiton of Bayes' theorem $\frac{P^+(A_i)}{P^+(A_j)} = \frac{P(A_i|E)}{P(A_j|E)} = \frac{p_i p'_i}{p_j p'_j}$. Then P learns the credences of her peers over the partition, and updates using *Upco*. This gives her a final credence of P^{final}. (Since we are not interested in her credences in any further propositions, we do not need to Jeffrey conditionalize.) Thus we will have $\frac{P^{final}(A_i)}{P^{final}(A_j)} = \frac{P^+(A_i)q_i r_i \cdots}{P^+(A_j)q_j r_j \cdots} = \frac{p'_i p_i q_i r_i \cdots}{p'_j p_j q_j r_j \cdots}$. This is the same result as for procedure 1.

On procedure 3, P learns E and conditionalizes, giving her a new credence P^+, with $\frac{P^+(A_i)}{P^+(A_j)} = \frac{p_i p'_i}{p_j p'_j}$ as before. Moreover, Q, R,... also learn E and conditionalize. Q will have a new credence Q^+, with $\frac{Q^+(A_i)}{Q^+(A_j)} = \frac{q_i q'_i}{q_j q'_j}$, and analogously for the other peers. P now updates on the new credences of her peers, leading to her final credence P^{final}. *Upco* tells us that $\frac{P^{final}(A_i)}{P^{final}(A_j)} = \frac{p_i p'_i q_i q'_i r_i r'_i \cdots}{p_j p'_j q_j q'_j r_j r'_j \cdots}$. This differs from the results of applying the other two procedures, since it contains the product of the ratios $q'_i/q'_j, r'_i/r'_j, \ldots$. Since these ratios need not be one or zero, P's final credences will typically be different in this case. □

One special case is where all of the peers have the same likelihoods p'_i. Then, the result of following procedure 3 will be to magnify the ratio $\frac{P^{final}(A_i)}{P^{final}(A_j)}$ by a factor of $(p'_i/p'_j)^{n-1}$ (where n is the number of peers). This suggests that the evidence E is being doubly (or rather n-tuply) counted. This shows the way that our rule implicitly assumes that the evidence of the different agents is independent. In Section 10 we will suggest a modification of *Upco* to handle this kind of case of shared evidence, either by focusing just on the *additional* evidence that each peer provides, or by focusing on a level at which the peers provide independent *interpretations* of the evidence.

There has been some tendency in the literature to formulate conditions such as independence-preservation and commutation with conditionalization, and treat them in an all-or-nothing fashion. In the case of independence-preservation, impossibility theorems have led researchers to give up on satisfying the condition altogether. What the foregoing discussion shows is that it may still be possible to formulate restricted versions of these conditions and find simple, plausible rules that satisfy them.

10. Modifications to *Upco*

Upco mimics Bayesian conditionalization when the likelihoods take very specific forms. In the case of an algebra generated by the propositions A, $\neg A$, *Upco* mimics likelihoods in which P's credence that Q will report a credence in A of q, conditional on A, is linearly proportional to q. That is, $P(Q(A) = q|A)$ is governed by the density function $f(q) = 2q$. In the more general case where we have k exclusive and exhaustive propositions $A_1, \ldots, A_k$, $Upco$ mimics likelihoods conditional on A_j governed by the density $f(q_1, \ldots, q_k) = k!q_j$, for $q_1 + \ldots + q_k = 1$ (and zero otherwise). (As we noted in Section 7.1 and Section 8, $Upco$ is compatible with many other likelihoods as well.) This means that $Upco$ can be expected to yield reasonable results in cases where it would be reasonable to have likelihoods approximately like these (even if P does not in fact have precise likelihoods). In many cases, however, there are reasons to think that it would be unreasonable to have these likelihoods. In this section, we consider a variety of such cases, and suggest some modifications to $Upco$.

The introduction of such modifications will inevitably involve trade-offs. The motivation for introducing heuristic rules like $Upco$ (and rival rules like linear averaging) is that such rules are much simpler than conditionalizing. By introducing modifications to $Upco$, we allow it to give reasonable answers in a wider range of cases. But at the same time, we sacrifice some of the simplicity that made $Upco$ attractive to begin with. The modifications will be more complex in at least two respects. First, the formulas themselves will be more complicated. More importantly, however, these modifications do not allow one to mechanically plug in the credences of one's peers and crank out a number; they require judgment about whether to employ a modified version of $Upco$, and if so, which one. They also require the agent to supply further parameters. Nonetheless, these modified versions of $Upco$ are still a great deal simpler than full-blown conditionalization. In particular, they don't require the agent to have complete likelihoods. Rather, they require the agent to have parameters that can be thought of as partial specifications of a likelihood function. For example, many of the modifications merely require the agent to provide a "break-even" point.

We do not assume that there is a uniquely correct trade-off between simplicity and flexibility. We offer the original $Upco$ rule, and the modifications described in the present section, as distinct options that make different trade-offs.

The general motivation for most of these alternatives is the same. One way to think of the original rule is that there is a "default" uniform distribution over the elements of the partition, and that each agent has independent, reliable evidence that they then want to combine. The strength of each agent's evidence is measured by the likelihood ratio, and the strength of the combined evidence is just the product of these ratios. For the first two modifications that we will consider, the change is that the "default" distribution is taken to be something other than the uniform distribution, either because one expects that the other agent started from a different default, or because the other agent shares some of one's own initial evidence. For the third modification, the change is that the other agent is taken to have evidence that is a superset of one's own. We then consider updating when the peers are experts (or anti-experts), and finally, we examine cases of estimating chances and cases of higher-order evidence where certain synergistic effects of $Upco$ would seem inappropriate.

10.1 Optimistic Peers

Suppose that P and Q are fans of the Los Angeles Clippers basketball team, and they are wondering whether they will win tonight's game. P wants to adjust her credence in this outcome, A, by considering Q's credence. However, P thinks that Q is over-optimistic and likely to have an unsuitably high credence.

According to $Upco^{\dagger}$ and $Upco^{\star}$, a credence in A of .5 is the "break-even" point. That is, if an epistemic peer reports a credence greater than .5 for A, that is taken to be evidence in favor of A; if she reports a credence less than .5, that is taken to be evidence against A; and a credence of .5 provides no evidence one way or the other. (In the case of a partition with k members, a uniform distribution that assigns a credence of $1/k$ to each possibility serves as the "break-even" point.)

A natural way to accommodate an optimistic peer would be to shift the break-even point upwards. For example, P might shift the break-even point for Q to .6. This means that if Q reports a credence of .6, P would take this to provide no evidence for or against A.

More generally, she might take the break-even point to be some value q'. A simple way to do this is to weight each term in the likelihood ratio by the inverse of the respective break-even point.[30] That is:

$$\frac{P(Q(A) = q|A)}{P(Q(A) = q|\neg A)} = \frac{(q/q')}{(1-q)/(1-q')} = \frac{q(1-q')}{(1-q)q'}$$

This ratio takes the value 1 when $q = q'$ — i.e. when Q's credence is at the break-even point q'. Suppose Q reports a credence in A of q. Then conditionalization will yield a revised credence for P of:

$$P^+(A) = \frac{pq(1-q')}{pq(1-q') + (1-p)(1-q)q'}$$

This is just the result of accommodating a third peer who reports credence $Q'(A) = 1 - q'$. That is, P can counteract the effects of an optimistic peer by adding a fictitious peer who is suitably pessimistic. Equivalently, it is the result of accommodating a third peer with non-normalized credences $Q(A) = 1/q$ and $Q(\neg A) = 1/(1-q)$. This latter formula will be useful when we generalize to the case of a partition. More generally, if P's break-even points for peers $Q, R, \ldots,$ are $q', r', \ldots,$ respectively, $Upco^\star$ will be modified by adding fictitious peers with credences $(1-q'), (1-r'), \ldots$.

In the case of a partition $\{A_1, \ldots, A_k\}$, if P wants to make the break-even distribution for Q to be $\{q'_1, \ldots, q'_k\}$, and analogously for the other peers R, S,..., $Upco$ can be modified to:

$$P^+(A_i) = \frac{(p_i q_i r_i \ldots / q'_i r'_i \ldots)}{\sum_j (p_j q_j r_j \ldots / q'_j r'_j \ldots)}.$$

This is equivalent to adding peers with (non-normalized) credences $\{1/q'_1, \ldots, 1/q'_k\}$, $\{1/r'_1, \ldots, 1/r'_k\}$,....

Note that this revised formula violates the neutrality condition of Russell et al. (2015). $P^+(A)$ will not just be a function of $P(A), Q(A), R(A), \ldots$, but it will also depend on A. For example, in the case of a binary proposition A, and two epistemic peers P and Q, if P takes the break-even point for Q's credence to be different from .5, the formula for $P^+(A)$ will be different from the formula for $P^+(\neg A)$. If the break-even point is .6, a report of $Q(A) = .6$ will leave P's credence unchanged, while a report of $Q(\neg A) = .6$ will cause P to revise her credence in $\neg A$ upward.

But because it behaves as though we are adding peers with credences related to the break-even points, things like preservation of independence will still hold in cases where all peers *and the break-even points* have the feature (e.g. initial agreement on the independence of two propositions).

10.2 Common Background Knowledge

Suppose that epistemic peers P, Q, and R live in Southern California, and they are planning a day at the beach next Saturday. They are interested in whether it will be sunny that day. They share the background knowledge that it is almost always sunny in Southern California. Let's say that, in the absence of any specific knowledge about Saturday's

30. As usual, there are many different likelihood densities that yield this likelihood ratio. One pair of densities that works assigns $P(Q(A) = q|A)$ the density $f(q) = (n+1)q^n$; while the likelihood $P(Q(A) = q|\neg A)$ has the density $g(q) = n(n+1)[q^{(n-1)} - q^n]$; where $n = q'/(1-q')$. This can be found by imposing the constraints that (i) $f(q)/g(q)$ has the appropriate ratio, (ii) that $f(q)$ has the form $f(q) = mq^n$, and (iii) $\int_0^1 f(q) = \int_0^1 g(q) = 1$, and then solving for n. A second pair results from applying the same pair of formulas, reversing A and $\neg A$, and reversing q' and $(1-q')$. As is often the case, the likelihood ratios are much simpler than the likelihoods themselves. Thanks to Hendrik Rommeswinkel for helpful discussion on this topic.

weather (such as from a weather report), they would each have a credence of .8 in the proposition that it will be sunny (since it is sunny 80% of the time). Thus, if Q and R reported credences of .8, P would take this to provide no new evidence about the prospects for sunshine, and her credence would remain unchanged. That is, P takes the break-even point to be .8. In this example, P shifts the break-even point upward, not because she thinks Q and R are over-optimistic, but because the nature of the proposition together with shared background knowledge make .8 the default probability. Now, if Q and R report credences of q and r, P will revise her credence to:

$$P^+(A) = \frac{(pqr/.8^2)}{(pqr/.8^2) + (1-p)(1-q)(1-r)/(.2^2)}$$

$$= \frac{pqr}{pqr + 16(1-p)(1-q)(1-r)}.$$

This formula will be symmetric for P, Q, and R, so long as they share the same background knowledge, and hence all take the break-even point to be the same.

Generalizing to the case where n peers share background knowledge about a partition $\{A_1, \ldots, A_k\}$, reflected in a default probability distribution of $\{d_1, \ldots, d_k\}$, the formula will be:

$$P^+(A_i) = \frac{(p_i, q_i, r_i, \ldots,)/d_i^{(n-1)}}{\sum_j p_j, q_j, r_j, \ldots /d_j^{(n-1)}}.$$

This formula will be appropriate in any situation of the following sort: The peers come together, share information, and reach a consensus that the appropriate credence distribution is $Prob(A_i) = d_j$. Then they each go off and independently collect evidence. They return, and exchange information about their new credences. The new evidence collected will be reflected in differences between their new credences and the earlier consensus credence. Our treatment of this kind of case echoes the treatment of Dietrich (2010).

Recall that at the end of Section 9 we discussed the following kind of case: Peers P, Q, and R have credences over a partition $\{A_1, \ldots, A_k\}$, as well as for proposition E, which is not equivalent to a disjunction of the A_i's. The peers all have the same conditional credences for E given each member of the partition. That is, $P(E|A_i) = Q(E|A_i) = R(E|A_i) \cdots = p'_i$. The peers all conditionalize on E, and then update upon learning each other's new credences over the partition $\{A_1, \ldots, A_k\}$. It was suggested there that the result seemed to overweight the evidence E, counting it n times instead of once (where n is the number of peers). Specifically, it seemed that after updating on each other's credences, the ratio $\frac{p^{final}(A_i)}{p^{final}(A_j)}$ was magnified by a factor of $(p'_i/p'_j)^{n-1}$. The present proposal can correct for this magnification by letting the (normalized) likelihoods $\{p'_1, \ldots, p'_k\}$ play the role of the default distribution $\{d_1, \ldots, d_k\}$. This is just the distribution that would result from updating the uniform distribution $\{1/k, \ldots, 1/k\}$ by conditionalization on E, with the likelihoods given.

10.3 Structured Partitions

Suppose that $\{A, B\}$ is a partition, so that $B \equiv_{def} \neg A$. Then, as we have seen, $Upco$ takes the break-even distribution to be the one that assigns each possibility a probability of .5. Now suppose that we subdivide A into A_1 and A_2, yielding a new partition $\{A_1, A_2, B\}$. Then, $Upco$ will take the break-even distribution over this new partition to be the one that assigns each of these possibilities a probability of 1/3. When A is subdivided, the default probability for A changes from 1/2 to 2/3, and the default probability for B changes from 1/2 to 1/3. This is another way in which our $Upco$ rule is partition-dependent. Given the way in which the partition $\{A_1, A_2, B\}$ was constructed, by subdividing one element of a binary partition, it might be more natural to modify *Upco* by using a break-even distribution that assigns the cells probabilities of $1/4, 1/4$, and $1/2$, respectively. The strategy for modifying the break-even distribution described in the previous sections can be used to do this.

The issues here are closely connected to familiar worries about the *Principle of Indifference*, which is at the heart of the classical interpretation of probability.[31] We do not offer here any solution to these long-standing problems. However, *Upco* is easily modified to accommodate the preferred solution in any specific case.

10.4 Expert Peers

We have been exploring cases where P takes the break-even point to take different values. One special case is where P takes her own credence to be the break-even point. That is, if peer Q reports a credence for A that is higher than P's, P treats this as evidence for A; if Q reports a credence for A that is lower than P's, P treats this as evidence against A; and if Q's credence is the same as P's, that is evidence neither for A nor against A.

Now, when P learns that $Q(A) = q$, and applies $Upco^{\dagger}$ with p as the break-even point, she gets:

$$P^{+}(A) = \frac{p(q/p)}{p(q/p) + (1-p)(1-q)/(1-p)} = q.$$

In this case, P regards Q as an *expert*. That is, the modified version of $Upco^{\dagger}$ mimics a likelihood for which conditionalization yields $P(A|Q(A) = q) = q$. This is the standard definition of what it is for P to regard Q as an expert (with respect to A.) The philosophical literature on rational credences contains two well-known examples of expert principles. David Lewis's *Principal Principle* (Lewis 1980) says that (in the absence of inadmissible information) one should treat objective chance as an expert. Conditional on the chance of A being q, one should set one's credence in A to q. Bas van Fraassen's *Reflection Principle* (van Fraassen 1984) says that we should treat our future credences as experts. Conditional upon my coming to have credence q in A, I should set my current credence in A to q.

This generalizes to the case of a partition $\{A_1, \ldots, A_k\}$ where P takes her own credence distribution as the break-even distribution:

$$P^{+}(A_i) = \frac{p_i(q_i/p_i)}{\sum_j p_j(q_j/p_j)} = q_i.$$

As we saw with the original version of *Upco*, if P begins with a credence of .5 in a binary proposition A, or a credence of $1/k$ for each element of a partition $\{A_1, \ldots, A_k\}$, she will defer to the credence of her epistemic peer. If P has different initial credences, we can get the same deferential behavior by setting the break-even point to those credences.

This formula will be appropriate in the following kind of case: P and Q meet and share information. They arrive at a consensus about the appropriate credence for A (or for a partition). Then Q goes off and collects additional evidence, while P stays home, learning nothing new about A. Q returns and reports her new credence. Since Q has strictly more information than P (Q isn't a forgetful type, and P knows this), P adopts Q's new credence as her own.

Note that when P regards Q as an expert, it is a bit misleading to refer to Q as an epistemic peer, since P regards Q's opinion as wholly preferable to her own. Nonetheless, we will continue to use the word 'peer' in a technical sense, meaning someone whose credences P accords some weight.

Note also that, in the case where P learns the credence of one expert, Q, the updating rule does not yield synergistic effects. This is obvious: since P adopts the credence of Q, her credence will never be higher than Q's.

Suppose, however, that P has multiple epistemic peers $Q_1, \ldots$, and

31. See for example Salmon (1967, pp. 65–68) and Zabell (2016) for discussion.

Q_n, all of whom she regards as experts.[32] When she learns of their credences in a binary proposition A, she will revise her credence to:[33]

$$P^+(A) =$$

$$\frac{p(q_1/p)\ldots(q_n/p)}{p(q_1/p)\ldots(q_n/p) + (1-p)((1-q_1)/(1-p))\ldots((1-q_n)/(1-p))} =$$

$$\frac{q_1\ldots q_n/p^{(n-1)}}{q_1\ldots q_n/p^{(n-1)} + (1-q_1)\ldots(1-q_n)/(1-p)^{(n-1)}}$$

In this case, there can still be synergistic effects. Suppose, for example, that P's credence in A is .7, while both Q_1 and Q_2 report a credence of .8. Then P's new credence will be:

$$P^+(A) = \frac{.8^2/.7}{[.8^2/.7] + [.2^2/.3]} \approx .87.$$

We may think of this case in the following way: P, Q_1, and Q_2 meet and share information. They reach a consensus that the appropriate credence in A is .7. Then Q_1 and Q_2 go out and independently collect additional evidence, while P stays home. Q_1 and Q_2 return, and each reports that they have collected evidence that confirms A. After hearing from both Q_1 and Q_2, P now has more evidence in favor of A than either Q_1 or Q_2 had prior to returning.

32. This case is discussed by Dawid et al. (1995) and Bradley (2015). In this case, the linear averaging strategy is incompatible with Bayesian conditionalization if all experts are given positive weights. Bradley takes this result to be a reason to reject linear averaging.

33. This formula is very similar to the multiplicative rule offered by Bordley (1982, 1137), where P plays the role of the "decision-maker", and Q_1 through Q_n the "experts" whom P polls. Bordley's formula also includes exponential weights for each of the experts.

10.5 Anti-experts[34]

Suppose that P regards Q as an *anti-expert* with regard to proposition A. That is, P regards Q's confidence in the truth of A as reason to doubt A, and vice-versa. (Most of us know of political pundits[35] and movie critics whom we regard as anti-experts in this way.) The natural way to accommodate this would be to reverse the role of q and $1-q$ in $Upco^\dagger$. Thus:

$$P^+(A) = \frac{p(1-q)}{p(1-q) + (1-p)q}.$$

In the case where P and Q have credences over the partition $\{A_1, \ldots, A_k\}$ and P regards Q as an anti-expert, the formula will be

$$P^+(A_i) = \frac{p_i/q_i}{\sum_j p_j/q_j}.$$

If we want to include different break-even points for an anti-expert, we need to invert our normal rule for incorporating break-even points. That is, if the break-even distribution for anti-expert Q on partition $\{A_1, \ldots, A_k\}$ is $\{q'_1, \ldots, q'_k\}$, the rule would be

$$P^+(A_i) = \frac{p_i q'_i/q_i}{\sum_j p_j q'_j/q_j}.$$

The ratios q_i/q'_i are inverted throughout this formula.

34. Thanks to Mark Colyvan for comments that led to this section.

35. Thanks to Rush Limbaugh for this suggestion.

10.6 Weighting

A number of authors have dealt with relative expertise by assigning different weights to different peers (see, e.g. Joyce [2007] and Bradley [2015]). Linear averaging rules are easily adapted to allow for assigning different weights to the peers (sometimes called "degrees of respect"). If the peers are $P_1, \ldots, P_n$, and they receive weights $w_1, \ldots, w_n$, with $w_1 + \cdots + w_n = 1$, we have

$$P^+(A) = \sum_j w_j P_j(A).$$

Since *Upco* is a multiplicative rule, the natural analog would be to use weights as exponents. That is, if peers $P_1, \ldots, P_n$ have weights $w_1, \ldots, w_n$, we can modify *Upco*⋆ to:

$$P_1^+(A) = \frac{p_1^{w_1} p_2^{w_2} \ldots p_n^{w_n}}{p_1^{w_1} p_2^{w_2} \ldots p_n^{w_n} + (1-p_1)^{w_1}(1-p_2)^{w_2} \ldots (1-p_n)^{w_n}}$$

This formula returns a probability regardless of the values of $w_1, \ldots, w_n$; they don't need to sum to 1, and the formula does not need to be normalized.[36]

In the case of updating on a partition, we can generalize to add weights in the obvious way:

$$P^+(A_i) = \frac{p_{1i}^{w_1} p_{2i}^{w_2} \cdots p_{ni}^{w_n}}{\sum_j p_{1j}^{w_1} p_{2j}^{w_2} \cdots p_{nj}^{w_n}}$$

36. This is clear when we note that the calculation for $P(\neg A)$ simply reverses p and $1-p$ in the formula. This leaves the denominator unchanged, but turns the numerator into the other half of the denominator. Thus the values of $P(A)$ and $P(\neg A)$ must add up to 1.

If we raise p_i to the power of 2, we are treating P_i's report that $P_i(A) = p_i$ as equivalent to the report of two independent peers with weight 1 reporting that credence. It means that our formula is mimicking conditionalization using a density function for $P_1(P_i(A) = p_i|A)$ that has the form $f(p) = 3p^2$. In general, the higher the exponent we assign to peer P_i, the more sensitive we regard P_i as being to the truth of A. If w_i is large, we take even small departures from .5 to provide strong evidence that A is true.

It is clear that an agent should assign her own exponent to be $w_1 = 1$. Otherwise, she would interpret her own report as additional evidence for A, so that when combining with herself, her new credence would be different from her original. That is obviously undesirable.

Our rule for adding weights does have one curious effect, however. If we assign a weight to a peer that is greater than 1, it means that we think that she is undervaluing her own evidence. For instance, suppose that there are two peers, P and Q. P assigns to Q a weight of 2, meaning that she thinks Q's report is worth the reports of two "normal" independent peers. Suppose that P has a credence in A of .5. Then Q reports that his credence is .8. P's new credence will be

$$P^+(A) = \frac{(.5)(.8)^2}{(.5)(.8)^2 + (.5)(.2)^2} \approx .94.$$

So while P respects Q, and even assigns his opinion more weight than she does her own, she thinks that Q's evidence for A merits a credence in A of .94, rather than the credence of .8 that Q himself assigns. We leave it as an open question whether it is possible to incorporate weights in a way that doesn't have this feature.

When updating using *Upco* the extent to which my credence will move toward that of my peer doesn't depend only on the weighting that I assign to my peer; it also depends on my own original credence. As mentioned above, if I begin with P(A) = .5, and assign my peer a weight of 1, equal to my own, then I will completely adopt the credence of any peer that I combine with. The less opinionated I am to begin with, the more easily I adopt the credences of my peers. This is so even if I assign them weights no higher than my own.

It would also seem natural to use weights when I do not regard my peers as being completely independent.[37] For instance, suppose that I think that the opinions of my five peers are partially dependent (even conditional upon the truth or falsity of the proposition in question), so that their opinions really are equivalent to those of three fully independent peers. I might reflect this by assigning each of them a weight of 3/5. The general question of how to deal with peer update when we think there are correlations between one's peers, or indeed between oneself and one's peers, cannot be dealt with here. We can note that our opinions about these correlations will be reflected in our likelihood functions, and so updating by conditionalization is still the correct normative standard. However, whether there are any simple formulas or appropriate calculational shortcuts is an open question that awaits further work.

Example 10.1. *Let us return to the case described in Example 7.9. P, Q, and R have credences in A of .8, .6, and .6, respectively. Q and R meet, and update using Upco, arriving at new credences of* $Q^+(A) = R^+(A) = .69$. *We saw that if P now updates using Upco on both* Q^+ *and* R^+, *she will have a new credence of* $P^+(A) \approx .95$. *This is the same credence that would result from updating on* four *peers with credences of .6.*

In this example, the credences of Q and R are not independent. In fact, they are perfectly coordinated. Thus, the combined opinions of Q and R should carry the weight of only one *independent peer. We can correct by assigning each of them a weight of* 1/2*:*

37. In the context of his multiplicative approach to the aggregation problem, Bordley (1982, 1142) introduces weights into his formula to reflect not just how reliable the agents' opinions are, but also how correlated their opinions are (conditional upon the truth of the proposition in question). Compare also Morris (1974, 1239–1240) and Morris (1977, 682, 687).

$$P^+(A) = \frac{P(A)[Q^+(A)]^{\frac{1}{2}}[R^+(A)]^{\frac{1}{2}}}{P(A)[Q^+(A)]^{\frac{1}{2}}[R^+(A)]^{\frac{1}{2}} + P(\neg A)[Q^+(\neg A)]^{\frac{1}{2}}[R^+(\neg A)]^{\frac{1}{2}}}$$

The result is $P^+(A) = .9$, *which is the same credence that would result from updating on two independent peers with a credence of .6.*

10.7 Chancy Events

Suppose P and Q jointly draw a coin from an urn. They both know that the urn contains some coins that are fair, and others that are biased 2-to-1 in favor of heads. They don't know the exact composition of the urn, and each one has independent evidence informing her credences for drawing a fair or biased coin. The coin they draw will be flipped at noon tomorrow. Let H be the proposition that this coin lands heads.

Now it seems reasonable that P and Q will have credences in H between 1/2 and 2/3. Suppose that they both have credence .6 in H. If we apply $Upco^\dagger$ to their credences in H, we will get:

$$P^+(H) = \frac{.6^2}{.6^2 + .4^2} \approx .69$$

But it seems unreasonable for P to update her credence in H to .69 after learning Q's credence. After all, she knows that the bias of the coin is at most 2/3.

What has gone wrong? $Upco^\dagger$ mimics conditionalization with a likelihood in which $P(Q(A) = q|A)$ is proportional to q. In this example, a linear likelihood for Q's credence in H would not be reasonable. Even if the coin does land heads when it is flipped tomorrow, P would not expect Q to report a credence for this event that is greater than 2/3. Analogously, even if the coin does land tails, P would not expect Q to report a credence of less than 1/2.

In this case, it seems much more reasonable for P to regard Q's credence in H as resulting from a weighted average of two propositions: that the coin is fair (F) and that the coin is biased 2-to-1 in favor of heads (B). That is:[38]

$$\begin{aligned} Q(H) &= Q(H|B)Q(B) + Q(H|F)Q(F) \\ &= 2/3Q(B) + 1/2Q(F) \end{aligned}$$

Since Q reports a credence $Q(H) = .6$, P can infer that Q's credence in B is .6. Now it seems much more reasonable for P to have linear likelihoods for Q's credence in B than in H. Thus we apply $Upco^{\dagger}$ to P's and Q's credences in B. Since P also has $P(B) = .6$, this yields $P^{+}(B) \approx .69$ and $P^{+}(F) \approx .31$. We may now use these to calculate a credence for H:

$$P^{+}(H) = 2/3P^{+}(B) + 1/2P^{+}(F) \approx .63.$$

There are still synergistic effects, but they are less dramatic. In particular, they do not take us above 2/3. Adding further peers with credence .6 will drive us toward the maximum of 2/3, but not past it.

In general, when P thinks that her peers' credences in proposition A result from their credences in hypotheses about the chance of A, it may be more natural to apply $Upco$ to the hypotheses about the chance of A, rather than to A itself.

10.8 Higher-Order Evidence

We think that a similar approach may be appropriate in cases involving *higher-order evidence*. Adapting an example from Christensen 2007, suppose P and Q are meteorologists with access to current weather data provided by the National Oceanic and Atmospheric Administration, the National Weather Service, etc., and have learned to apply both calculations and judgments of similarity to figure out what the data say about rain tomorrow. Because the models and data are quite complex, neither P nor Q is completely confident about how strongly the evidence supports belief in rain. But given their imperfect confidence in their evaluation of the evidence, P comes to have degree of belief .55 in rain while Q comes to have degree of belief .45 in rain.

One way to think about disagreements like this is to try to use the strategy of the previous section and to think of the case as one where the disagreement is generated by different degrees of belief about the chances of rain. But the way Christensen thinks about the case, the issue involves disagreement about how the shared evidence *bears on* the proposition that rain will occur. Nonetheless, we think that the formal treatment will be similar.

Let X be the evidence, and assume that one's degree of belief is given by a kind of estimation of the *evidential probability*. Evidential probability functions as an *expert*, in the sense described above in Section 10.4. That is, the peers' credences are guided by their beliefs about about the evidential probabilities in the same sort of way that their credences can be guided by their beliefs about chances through the *Principal Principle*. Let $E(R|X)$ be the evidential probability of rain given the shared evidence, so that degrees of belief satisfy $P(R) = \sum x \cdot P(E(R|X) = x)$.[39]

Then, one way for the peers to arrive at their credences $P(R) = .55$ and $Q(R) = .45$ is if $P(E(R|X) = .6) = .75$ and $P(E(R|X) = .4) = .25$, while $Q(E(R|X) = .6) = .25$ and $Q(E(R|X) = .4) = .75$. This would be natural if each meteorologist is equally good, and each is certain that the evidence is such as to support rain to degree .6 or .4, and is 75% confident that she would evaluate the evidence correctly in this situation.

Once the two meteorologists realize that they have come to different interpretations of the evidence, how should they react? Just as the coin flippers above think that each other's credences are sensitive to the evidence they have about the urn rather than being sensitive to the

38. Substituting 2/3 for $Q(H|B)$ and 1/2 for $Q(H|F)$ assumes that Q satisfies the *Principal Principle* and that Q has no *inadmissible* information. See Lewis (1980) for the classic statement of these ideas.

39. However, see (Lasonen-Aarnio 2013, Section 2) for criticism of this way of conceiving of evidential probability.

actual outcome of the coin, it is natural for the meteorologists to think that their judgments about the evidential bearing of X on rain are sensitive to the actual evidential bearing of X on rain, and not sensitive to rain itself. Although their evidence about rain is not independent, their higher-order evidence about the evidential probability of rain is, in a sense, independent. Thus, it is natural to apply $Upco^{\dagger}$ to the partition into $E(R|X) = .6$ and $E(R|X) = .4$, and not to any partition involving rain.

In this particular case, the end result is the same — because their credences in the two hypotheses about the evidential bearing of X on R are equal and opposite, they should update to

$$\begin{aligned} P^+(E(R|X) = .6) = P^+(E(R|X) = .4) &= Q^+(E(R|X) = .6) \\ &= Q^+(E(R|X) = .4) \\ &= .5, \end{aligned}$$

and thus both end up with degree of belief .5 in rain, as Christensen suggests.

However, because their judgments are only imperfect evidence about the bearing of the first-order evidence, related cases can end up quite differently. If P and Q both had degree of belief .55 in rain, and both were initially 75% confident that they interpreted the evidence correctly, then we would have:

$$P^+(E(R|X) = .6) = Q^+(E(R|X) = .6) = \frac{(3/4)^2}{(3/4)^2 + (1/4)^2} = .9,$$

and thus would end up with degree of belief .58 in rain. There is some synergy from their agreement in the assessment of evidence, but it does not get as high as applying $Upco^{\dagger}$ directly to their degrees of belief in rain, which would yield $\frac{.55^2}{.55^2+.45^2} \approx .599$.

If P is in fact a perfect judge of the evidence, while Q is quite imperfect, the credences of .55 and .45 in rain could come from $P(E(R|X) = .55) = 1$ while $Q(E(R|X) = .55) = 1/3$ and $Q(E(R|X) = .4) = 2/3$. In this case, Q's initial confidence in a lower evidential probability for rain is entirely swamped by P's certainty in a higher evidential probability for rain. This would give rise to a sort of steadfastness for P and complete deference for Q. (If P and Q disagree in their estimation of how reliable each other are as judges of the evidence, we could further apply some of the techniques from Section 10.1.)

In more realistic cases where agents are uncertain of the evidential probability, agents will distribute their credence over a range of possible values rather than just two, and the interplay of synergy and disagreement will be more complex. But in many of these cases, it makes sense to work with some version of $Upco$ over the partition by particular values of evidential probability, rather than the partition over first-order propositions.

However, the special case of evidence that should be completely definitive deserves special comment. Consider Christensen's example of splitting the check:

> Suppose that five of us go out to dinner. It's time to pay the check, so the question we're interested in is how much we each owe. We can all see the bill total clearly, we all agree to give a 20 percent tip, and we further agree to split the whole cost evenly. . . . I do the math in my head and become highly confident that our shares are $43 each. Meanwhile, my friend does the math in her head and becomes highly confident that our shares are $45 each. How should I react, upon learning of her belief? (Christensen 2007, 193)

In this case, it seems natural to think that the available evidence either logically guarantees that our shares are $43, or logically guarantees that our shares are $45, or logically guarantees that they are some other value. Furthermore, we are all certain that the evidence logically guarantees that the shares are a given value iff the shares actually do have that value. Thus, the partition by particular values of evidential probability will be coextensive with the partition over first-order propositions. Applying *Upco* to the evidential partition and then propagating the values to the first-order propositions will give the same results as applying *Upco* directly to the first-order partition. This is interesting, because in this case the peers share all of the relevant evidence. Nonetheless, *Upco* yields the intuitively correct result. This is because the peers still act independently, in the sense that each peer has the same probability of committing a mathematical error, whether the other peers do so or not. Thus, it is safe to treat independent evaluation of shared, completely definitive evidence the same as one would treat disagreement at the first-order level with completely independent evidence.

10.9 Partition-dependence

The examples of the previous two sections illustrate once again the partition-dependence of *Upco*. In Section 10.7, we had two partitions $\{H, T\}$ and $\{F, B\}$, and when P updates her credences on the basis of Q's credence, we get a different result depending upon whether she applies *Upco* to Q's credences over the first partition or the second. Readers may verify that applying *Upco* to Q's credence on the joint partition $\{HF, HB, TF, TB\}$ yields yet a different result. (In this case, P's new credence in heads would be approximately .71, which again seems unreasonable given the structure of the problem.) In Section 10.8, we saw that applying *Upco* to evidential probabilities can lead to yet another different result. How can we know which procedure to use in a particular case?

In general, suppose that we have a "coarse" partition $\{A_1, \ldots, A_k\}$, and that each A_i can be further partitioned into $\{A_{i1}, \ldots, A_{im_i}\}$, yielding a "fine" partition $\{A_{11}, \ldots, A_{1m_1}, \ldots, A_{k1}, \ldots, A_{km_k}\}$. P has credences over the fine partition, and learns Q's credences over the fine partition, from which she can easily infer Q's credences over the coarse partition. Now P can apply *Upco* to Q's credences over the fine partition, or she can apply it to Q's credences over the coarse partition (and update the rest of her credences by Jeffrey conditionalization, as discussed in Section 7.1). These two procedures will (typically) produce different results, so which should she do?[40]

One way to think about partition-dependence is to recall that *Upco* mimics specific likelihoods. The likelihoods that *Upco* mimics when it is applied to the fine partition are not compatible with the likelihoods that it mimics when it is applied to the coarse partition. So one way to address the question of which partition to apply *Upco* is to think about which likelihoods it would be better to mimic. We will here present one sufficient condition for preferring the coarse partition to the fine partition.

The examples of the previous two sections provide clues. Recall that there were two hypotheses about the bias of the coin: the coin was either fair (F) or biased (B). If P had explicit likelihoods for Q's credences, it would be reasonable for her likelihoods to satisfy the following conditions:

$$P(Q(X)|FH) = P(Q(X)|FT) = P(Q(X)|F);$$
$$P(Q(X)|BH) = P(Q(X)|BT) = P(Q(X)|B);$$

where X is any proposition about the bias of the coin and/or the outcome of the toss. That is, given the bias of the coin, P thinks that Q is equally likely to have any given credence, regardless of the eventual outcome of the toss. Put another way, the bias of the coin screens off Q's credences from the outcome of the toss. These credences would be reasonable because Q does not have any information about the result of the coin toss besides her information about the bias of the coin.[41]

40. We would like to thank Julia Staffel and Peter Vranas especially for pressing this worry upon us.

41. We might derive these conditions by assuming P believes Q obeys the *Principal Principle* and has no *inadmissible* information. See Lewis (1980).

In the more general case, where we have a coarse partition $\{A_1, \ldots, A_k\}$ and a fine partition $\{A_{11}, \ldots, A_{1m_1}, \ldots, A_{k1}, \ldots, A_{km_k}\}$, we can formulate an analogous condition. Let $\vec{q} = \langle q_{11}, \cdots, q_{km_k} \rangle$ abbreviate the conjunction of propositions of the form $Q(A_{ij}) = q_{ij}$. That is, $\vec{q}$ fully specifies Q's credences over the fine partition. Then suppose:

$$P(\vec{q}|A_{ij}) = P(\vec{q}|A_i) = c_i(\vec{q}) \qquad \text{(Coarse)}$$

for all $\vec{q}$ and i, j. This says that P's credence that Q will have credence $\vec{q}$ is the same, regardless of which member of the subpartition $\{A_{i1}, \cdots, A_{im_i}\}$ is true. This reflects the idea that P thinks that Q's credence is sensitive only to which member of the coarse partition $\{A_1, \ldots, A_k\}$ is true, and not to which member of the fine partition is true. Coarse is inconsistent with the likelihoods that are mimicked by applying *Upco* to the fine partition, for those require that:

$$\frac{P(\vec{q}|A_{ij})}{P(\vec{q}|A_{il})} = \frac{q_{ij}}{q_{il}},$$

which will be different from 1 whenever $q_{ij} \neq q_{il}$. By contrast, Coarse requires that $\frac{P(\vec{q}|A_{ij})}{P(\vec{q}|A_{il})}$ always be equal to one. This shows that when P's credences satisfy Coarse it would be inappropriate to update by applying *Upco* to the fine partition.

Next we show that when P's credences satisfy Coarse, and P learns Q's credences, it does not matter whether P updates her credence in each member of the fine partition by conditionalizing on Q's credences, or whether she first updates her credence in each member of the coarse partition by conditionalizing on Q's credences and then updates her credences over the fine partition by Jeffrey conditionalizating with her new credences on the coarse partition.

Theorem 10.2. *Let $\{A_1, \ldots, A_k\}$ be a partition, the "coarse partition", and let $\{A_{11}, \ldots, A_{1m_1}, \ldots, A_{k1}, \ldots, A_{km_k}\}$, the "fine partition", be a refinement of the coarse partition, so that $A_i = \bigcup_{j=1,\ldots,m_i} A_{ij}$ for all i. Let $P(A_i) = p_i$ and $P(A_{ij}) = p_{ij}$ for all i, j. Suppose that P's credences satisfy* Coarse*, and then she learns that Q's credences over the fine partition are given by $\vec{q}$. Then the following two procedures yield the same result: 1) P updates her credence in each member of the fine partition A_{ij} by conditionalizing on $\vec{q}$; 2) P updates her credence in each member of the coarse partition A_i by conditionalizing on $\vec{q}$, and then updates her credence in each member of the fine partition A_{ij} by Jeffrey conditionalizing using her new credences in the coarse partition.*

Proof. Let A_{ij} be an arbitrary member of the fine partition. Following procedure (1):

$$\begin{aligned} P^+(A_{ij}) &= P(A_{ij}|\vec{q}) \\ &= \frac{P(A_{ij}) \cdot P(\vec{q}|A_{ij})}{P(\vec{q})} \\ &= \frac{p_{ij} \cdot c_i(\vec{q})}{P(\vec{q})} \qquad \text{(Procedure 1)} \end{aligned}$$

Now, following procedure (2), we first compute P's new credences over the coarse partition:

$$\begin{aligned} P^+(A_i) &= P(A_i|\vec{q}) \\ &= \frac{P(A_i) \cdot P(\vec{q}|A_i)}{P(\vec{q})} \\ &= \frac{p_i \cdot c_i(\vec{q})}{P(\vec{q})} \end{aligned}$$

Then we compute P's new credences over the fine partition by Jeffrey conditionalizing with the new probabilities over the coarse partition:

$$
\begin{aligned}
P^+(A_{ij}) &= P(A_{ij}|A_i) \cdot P^+(A_i) \\
&= \frac{p_{ij}}{p_i} \cdot P^+(A_i) \\
&= \frac{p_{ij}}{p_i} \cdot \frac{p_i \cdot c_i(\vec{q})}{P(\vec{q})} \\
&= \frac{p_{ij} \cdot c_i(\vec{q})}{P(\vec{q})} \qquad \text{(PROCEDURE 2)}
\end{aligned}
$$

Thus the two procedures yield the same result. □

What Theorem 10.2 shows is that when P's credences satisfy COARSE, she loses no information by updating on Q's credences over the coarse partition, rather than over the fine partition. Moreover, we saw that COARSE is incompatible with the likelihoods mimicked by applying $Upco$ to the fine partition. Thus, if P's credences satisfy COARSE, it would be more appropriate for P to update by applying $Upco$ to the coarse partition, rather than the fine partition.

Of course, we would like to be able to use $Upco$ or one of its variants even when an agent does not have explicit likelihoods. Nonetheless, COARSE gives us some sense of the types of situation where it is appropriate to update on a coarse partition. If P thinks that Q's credences are sensitive only to which member of the coarse partition is true, that Q would tend to have similar credences regardless of which sub-cell of a member of the coarse partition is true, that Q has information about the fine partition only in virtue of having information about the coarse partition, etc., then P should update on Q's credences over the coarse partition.

Example 10.3. *P is hiking in the wilderness with Hilary Putnam. They see an odd-shaped tree in the distance, and begin to debate what kind of tree it is. They both read in a guide book that the trees in this region are beeches, elms, firs, and spruces. P's credences in these four possibilities are .6, .1, .1, and .2, respectively; while Putnam's credences are .3, .3, .3, and .1. However, P recalls Putnam saying that he does not know the difference between beeches and elms (Putnam 1975). Given his arboreal ignorance, P reasons that Putnam is unlikely to know the difference between firs and spruces. But surely even Putnam knows the difference between a deciduous tree and a conifer.*[42] *Hence, P decides to update using their credences over the partition* $\{deciduous(\equiv_{def}$ *beech or elm*$), coniferous(\equiv_{def}$ *fir or spruce*$)\}$. *P's credences over this partition are .7 and .3 (respectively), and Putnam's are .6 and .4. Applying Upco, she arrives at new credences of .78 and .22 over the coarse partition. Then she Jeffrey conditionalizes over this partition. Her final credences are .67 for beech, .11 for elm, .07 for fir, and .15 for spruce.* [43]

10.10 Combinations

Of course, it is possible to combine the various modifications discussed in this section. We may have a collection of peers, some of whom are anti-experts, all of whom have different weights and break-even points, reasoning about higher-order evidence about the chances of some event.

11. Conclusion

Many rules that have been proposed for updating one's credences upon learning the credences of others lack the important property of synergy. We have motivated this property both by intuitive consideration of cases and by showing that Bayesian updating supports it. Since Bayesian updating will often be computationally complex, it is desirable to have a short-cut rule. The rule that we have proposed, which exhibits synergy, is a special case of conditionalization for particular likelihoods, which suggests that our rule is at least sometimes the right response to learning the credences of others. We have gone on to consider generalizations of our rule that may help motivate it as a

42. Although a larch would probably stump him.
43. This example is similar in structure to one that Julia Staffel presented to us.

response to the updating problem even when one doesn't have these particular likelihoods.

Our rule has various advantages over the most commonly proposed alternatives, which all involve taking some sort of (weighted) average of one's own and one's peers' credences. The first advantage is precisely that it agrees with conditionalization in a clear class of cases (namely where one's likelihoods take a particular form), and that it yields synergy precisely where conditionalization does in this class of cases (and where synergy is intuitively plausible). Averaging rules cannot yield synergy. Second, our rule exhibits a plausible flexibility when it comes to conciliatoriness and steadfastness. For example, it will recommend steadfastness when one's peers' credences are simply equal to the uninformative "break-even" credences (reflecting indifference or a lack of evidence on their part), while it recommends conciliation (in the sense of adjusting one's credences in light of one's peers' credences) when one's peers have evidence that one lacks. Averaging rules are clunkier in this regard: flexible recommendations can be achieved only by re-weighting one's peers' credences on a case-by-case basis. Thirdly, our rule commutes with conditionalization upon non-peer evidence, and is also insensitive to the timing and the order in which one updates on one's peers' credences. By contrast, averaging doesn't commute with conditionalization and is sensitive to the order in which one meets one's peers. This leads averaging proposals to deliver implausible results. Fourth, our rule preserves independence in cases where it is plausible to do so, but not in cases in which it isn't. Averaging almost never preserves independence, even in cases where preservation of independence is plausible. Finally, averaging obeys the Context-Free Assumption, despite the fact that there are clear counterexamples to CFA. Our rule steers clear of these counterexamples.

Acknowledgments

We would like to thank David James Barnett, Paul Bartha, Richard Bradley, Seamus Bradley, Liam Kofi Bright, Mark Colyvan, Antony Eagle, Alan Hájek, Dennis Lehmkuhl, Leon Leontyev, Bradley Monton, Hendrik Rommeswinkel, Paulina Sliwa, Julia Staffel, Michael Titelbaum, Peter Vranas, and two anonymous referees for *Philosophers' Imprint*; as well as audience members at the 2013 Australian Association of Philosophy annual meeting in Brisbane, the University of Wisconsin-Madison, the University of North Carolina at Chapel Hill, the London School of Economics and Political Science, and the 2015 meetings of the Society for Exact Philosophy and the Formal Epistemology Workshop.

References

Boole, G. (1952). On the application of the theory of probabilities to the question of the combination of testimonies or judgments. In *Studies in Logic and Probability*, pages 308–385. Watts and Co.

Bordley, R. F. (1982). A multiplicative formula for aggregating probability assessments. *Management Science*, 28(10):1137–1148.

Bordley, R. F. and Wolff, R. W. (1981). Note — on the aggregation of individual probability estimates. *Management Science*, 27(8):959–964.

Bradley, R. (2007). Reaching a consensus. *Social Choice and Welfare*, 29(4):609–632.

Bradley, R. (2015). Learning from others: Conditioning versus averaging. Unpublished manuscript. Available at <http://personal.lse.ac.uk/bradleyr/pdf/Learning%20from%20Others%20version%202.pdf>.

Christensen, D. (2007). Epistemology of disagreement: The good news. *The Philosophical Review*, 116(2):187–217.

Christensen, D. (2009). Disagreement as evidence: The epistemology of controversy. *Philosophy Compass*, 4(5):756–767.

Christensen, D. (2010). Higher-order evidence. *Philosophy and Phenomenological Research*, 81(1):185–215.

Christensen, D. (2011). Disagreement, question-begging and epistemic self-criticism. *Philosophers' Imprint*, 11(6):1–22.

Cohen, S. (2013). A defense of the (almost) equal weight view. In Christensen, D. and Lackey, J., editors, *The Epistemology of Disagreement: New Essays*, pages 98–117. Oxford University Press.

Dawid, A. P., DeGroot, M. H., and Mortera, J. (1995). Coherent combination of experts' opinions. *Test*, 4(2):263–313.

de Ridder, J. (2014). Why only externalists can be steadfast. *Erkenntnis*, 79:185–199.

DeGroot, M. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.

Dietrich, F. (2010). Bayesian group belief. *Social Choice and Welfare*, 35:595–626.

Elga, A. (2007). Reflection and disagreement. *Noûs*, 41(3):478–502.

Elga, A. (2010). How to disagree about how to disagree. In Feldman, R. and Warfield, T. A., editors, *Disagreement*, pages 175–186. Oxford University Press.

Feldman, R. (2006). Epistemological puzzles about disagreement. In Hetherington, S., editor, *Epistemology Futures*, pages 216–236. Oxford University Press.

Gardiner, G. (2014). The commutativity of evidence: A problem for conciliatory views of peer disagreement. *Episteme*, 11(1):83–95.

Genest, C. and Schervish, M. J. (1985). Modeling expert judgments for bayesian updating. *The Annals of Statistics*, 13(3):1198–1212.

Genest, C. and Wagner, C. (1987). Further evidence against independence preservation in expert judgement synthesis. *Aequationes Mathematicae*, 32(1):74–86.

Genest, C. and Zidek, J. V. (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114–135.

Gutting, G. (1982). *Religious Belief and Religious Skepticism*. University of Notre Dame Press.

Hume, D. (1748). *An Enquiry Concerning Human Understanding*. A. Millar.

Jehle, D. and Fitelson, B. (2009). What is the "equal weight view"? *Episteme*, 6(3):280–293.

Joyce, J. M. (2007). Epistemic deference: The case of chance. In *Proceedings of the Aristotelian Society (Hardback)*, volume 107, pages 187–206. Wiley Online Library.

Keeney, R. L. and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. John Wiley.

Kelly, T. (2005). The epistemic significance of disagreement. *Oxford studies in epistemology*, 1:167–196.

Kelly, T. (2010). Peer disagreement and higher-order evidence. In Feldman, R. and Warfield, T. A., editors, *Disagreement*, pages 111–174. Oxford University Press.

Lackey, J. (2010). A justificationist view of disagreement's epistemic significance. In Haddock, A., Millar, A., and Pritchard, D., editors, *Social Epistemology*, pages 298–325. Oxford University Press.

Laddaga, R. (1977). Lehrer and the consensus proposal. *Synthese*, 36(4):473–477.

Lasonen-Aarnio, M. (2013). Disagreement and evidential attenuation. *Noûs*, 47(4):767–794.

Lehrer, K. and Wagner, C. (1981). *Rational consensus in science and society: A philosophical and mathematical study*, volume 24. D. Reidel.

Lehrer, K. and Wagner, C. (1983). Probability amalgamation and the independence issue: A reply to Laddaga. *Synthese*, 55(3):339–346.

Lewis, D. (1980). A subjectivist's guide to objective chance. In Jeffrey, R. C., editor, *Studies in Inductive Logic and Probability*, volume 2, pages 263–293. University of California Press.

Loewer, B. and Laddaga, R. (1985). Destroying the consensus. *Synthese*, 62(1):79–95.

Martini, C., Sprenger, J., and Colyvan, M. (2013). Resolving disagreement through mutual respect. *Erkenntnis*, 78(4):881–898.

McConway, K. J. (1981). Marginalization and linear opinion pools. *Journal of the American Statistical Association*, 76(374):410–414.

Morris, P. A. (1974). Decision analysis expert use. *Management Science*, 20(9):1233–1241.

Morris, P. A. (1977). Combining expert judgements: A bayesian approach. *Management Science*, 23(7):679–693.

Moss, S. (2011). Scoring rules and epistemic compromise. *Mind*, 120(480):1053–1069.

Putnam, H. (1975). The meaning of 'meaning'. In Gunderson, K., editor, *Language, Mind, and Knowledge*, volume 7 of *Minnesota Studies in the Philosophy of Science*, pages 131–193. University of Minnesota Press.

Romeijn, J.-W. (2015). Pooling, voting, and bayesian updating. Unpublished manuscript.

Russell, J. S., Hawthorne, J., and Buchak, L. (2015). Groupthink. *Philosophical Studies*, 172(5):1287–1309.

Salmon, W. C. (1967). *The Foundations of Scientific Inference*. University of Pittsburgh Press.

Staffel, J. (2015). Disagreement and epistemic utility-based compromise. *Journal of Philosophical Logic*, 44(3):273–286.

Steele, K. (2012). Testimony as evidence: More problems for linear pooling. *Journal of Philosophical Logic*, 41(6):983–999.

Stone, M. (1961). The opinion pool. *The Annals of Mathematical Statistics*, 32(4):1339–1342.

Van Cleve, J. (2011). Can coherence generate warrant ex nihilo? Probability and the logic of concurring witnesses. *Philosophy and Phenomenological Research*, 82(2):337–380.

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

van Wietmarschen, H. (2013). Peer disagreement, evidence, and well-groundedness. *The Philosophical Review*, 122(3):395–425.

Vavova, K. (2014). Confidence, evidence, and disagreement. *Erkenntnis*, 79:173–183.

Wagner, C. (1982). Allocation, Lehrer models, and the consensus of probabilities. *Theory and Decision*, 14(2):207–220.

Wagner, C. (1985). On the formal properties of weighted averaging as a method of aggregation. *Synthese*, 62(1):97–108.

Wagner, C. G. (2011). Peer disagreement and independence preservation. *Erkenntnis*, 74(2):277–288.

Wilson, A. (2010). Disagreement, equal weight and commutativity. *Philosophical Studies*, 149(3):321–326.

Winkler, R. L. (1968). The consensus of subjective probability distributions. *Management Science*, 15(2):B61–75.

Zabell, S. L. (1988). The probabilistic analysis of testimony. *Journal of Statistical Planning and Inference*, 20(3):327–354.

Zabell, S. L. (2016). Symmetry arguments in probability. In Hájek, A. and Hitchcock, C., editors, *The Oxford Handbook of Probability and Philosophy*. Oxford University Press.