Local and Global Deference

J. DMITRI GALLOW *

Abstract: A norm of local expert deference says that your credence in an arbitrary proposition A, given that the expert's probability for A is n, should be n. A norm of global expert deference says that your credence in A, given that the expert's entire probability function is E, should be $E(A)$. Gaifman (1988) taught us that these two norms are not equivalent. Stalnaker (2019) conjectures that Gaifman's example is "a loophole". Here, I substantiate Stalnaker's suspicions by providing characterisation theorems which tell us precisely when the two norms give different advice. They tell us that, in a good sense, Gaifman's example is the *only* case where the two norms differ. I suggest that the lesson of the theorems is that Bayesian epistemologists need not concern themselves with the differences between these two kinds of norms. While they are not strictly speaking equivalent, they are equivalent for all philosophical purposes.

1 | INTRODUCTION

Principles of expert deference play a prominent role in Bayesian epistemology.[1] For an example of a principle of expert deference: Lewis (1980)'s *Principal Principle* tells you that, in the absence of an extraordinary form of evidence, you should defer to the future objective chances. That is, given that the objective chance of an arbitrary proposition, A, is n, your own subjective probability, or credence, in A should be n, too. That is, if C is your credence function, and $\langle \mathcal{C}h_t(A) = n \rangle$ is the proposition that the time t chance of A is n, then you should satisfy the equality:[2]

$$C(A \mid \langle \mathcal{C}h_t(A) = n \rangle) = n$$

That's one principle of expert deference. For another: *Rational Reflection* says that you should defer to the *ideally rational* credences for someone with your evidence to

Final Draft. Forthcoming at *Philosophical Studies*.

* Thanks to Kevin Dorst, Ben Levinstein, and Bernhard Salow for feedback on earlier versions of this manuscript.

1. Principles of expert deference are discussed by Gaifman (1988), Talbott (1991), Lewis (1994), Hall (1994, 2004), Thau (1994), Hall & Arntzenius (2003), Arntzenius (2003), Ismael (2008, 2015), Briggs (2009), Pettigrew (2012), Christensen (2010), Lasonen-Aarnio (2015), Dorst (2020), Dorst *et al.* (2021), and Levinstein (forthcoming), among several others.

2. Or, maybe more carefully, we should say: $C(A \mid \langle \mathcal{C}h_t(A) = n \rangle)$ should be n *whenever it is defined*. To avoid interrupting the exposition with constant reminders about this proviso, I'll adopt the non-standard convention of treating an equality as trivially true whenever one side is undefined. Thus, even when $C(A \mid E)$ is undefined, I will treat the equality $C(A \mid E) = n$ as true for all n.

have. (For discussion, see Christensen, 2010, Elga, 2013, and Lasonen-Aarnio, 2015.) That is, given that the rational credence function for you to have is R, your credence in any proposition A should be $R(A)$.

$$C(A \mid \langle \mathcal{R} = R \rangle) = R(A)$$

(Here, '$\langle \mathcal{R} = R \rangle$' says that the rational credence function for someone with your evidence is R.)

These principles both tell you to defer to some expert probability function, but they take different forms. The first tells you to defer to the expert conditional on their views about *any* proposition; whereas the second tells you to defer to the expert conditional on their views about *every* proposition (that is: conditional on their entire probability function). We can call the first a norm of *local* expert deference, and the second a norm of *global* expert deference.

Local Deference You *locally* defer to an expert, $\mathcal{E}$, iff, for any proposition, A, and any number n, your credence in A, given that $\mathcal{E}$'s probability for A is n, is n.

$$C(A \mid \langle \mathcal{E}(A) = n \rangle) = n$$

Global Deference You *globally* defer to an expert, $\mathcal{E}$, iff, for any proposition A, and any probability function E, your credence in A, given that $\mathcal{E}$'s entire probability function is E, is whatever probability E gives to A.

$$C(A \mid \langle \mathcal{E} = E \rangle) = E(A)$$

It's not obvious what the relationship is between these two different ways of showing deference to an expert. It's natural to think that they're equivalent, in the sense that you will globally defer to an expert function $\mathcal{E}$ if and only if you locally defer to $\mathcal{E}$. However, as we'll see in §2 below, this isn't quite right. While globally deferring to $\mathcal{E}$ entails locally deferring to $\mathcal{E}$, an example from Gaifman (1988) teaches us that the converse is not true. In some cases, you can locally defer to $\mathcal{E}$ without globally deferring to $\mathcal{E}$.

Stalnaker (2019, pp. 111–12) speculates that Gaifman's example is "a loophole—a contrived case where [a principle of local deference] is satisfied without its usual motivation". Here, I will substantiate Stalnaker's suspicions. I'll argue that the differences between local and global deference are so incredibly slight as to be philosophicaly negligible—there is no good reason to accept the weaker local deference norm without accepting the stronger global deference norm. To that end, I will precisely characterise the situations in which global and local deference principles come apart. This characterisation will show us that Gaifman's original example of an expert who may be deferred to locally but not globally is—in a good sense—the *only* expert like this. So the kinds of situations in which it is possible to defer locally without deferring globally are *incredibly* singular and fragile. And there is no reason to think that these kinds of cases are epistemologically singular. The upshot is that Bayesians should have no qualms about moving freely back and forth between global and local formulations of principles of expert deference. While they are not strictly speaking equivalent, they are equivalent for all philosophical purposes.

2 | HOW LOCAL AND GLOBAL DEFERENCE NORMS DIFFER

I'm going to take for granted here that your credence function, C, is a countably additive probability function, defined over subsets of a space of possible worlds, $\mathcal{W}$. For the sake of simplicity, I'm going to assume that $\mathcal{W}$ is at most countably infinite. I'll call any $A \subseteq \mathcal{W}$ a 'proposition', and since $\mathcal{W}$ is at most countably infinite, we can suppose that C gives a probability to every proposition.

I'll suppose that you are certain that the expert's probability function is defined over exactly the same algebra of propositions as your own, namely the powerset of $\mathcal{W}$, $\mathbb{P}(\mathcal{W})$. And I'll suppose that we have a function from worlds in $\mathcal{W}$ to probability distributions over $\mathbb{P}(\mathcal{W})$, which I'll write '$\mathcal{E}$'. The value of this function, given the argument w—which I'll write '$\mathcal{E}_w$'—will be interpreted as the probability function the expert has at the world w. With this function, we can form the proposition that the expert's probability function is E (for some probability distribution E), by gathering together all the worlds $w \in \mathcal{W}$ such that $\mathcal{E}_w = E$.

$$\langle \mathcal{E} = E \rangle \stackrel{\text{def}}{=} \{ w \in \mathcal{W} \mid \mathcal{E}_w = E \}$$

We may likewise form the proposition that $\mathcal{E}$'s probability for A is n by gathering together all of the worlds $w \in \mathcal{W}$ such that $\mathcal{E}_w(A) = n$.

$$\langle \mathcal{E}(A) = n \rangle \stackrel{\text{def}}{=} \{ w \in \mathcal{W} \mid \mathcal{E}_w(A) = n \}$$

Given this setup, if you defer to $\mathcal{E}$ globally, then you will defer to $\mathcal{E}$ locally as well. To appreciate this, just notice that $\langle \mathcal{E}(A) = n \rangle$ is partitioned by the set of all propositions of the form $\langle \mathcal{E} = E \rangle$, for some E that gives a probability of n to A. It then follows from conglomerability that, if $C(A \mid \langle \mathcal{E} = E \rangle) = n$ for each E such that $E(A) = n$, then $C(A \mid \langle \mathcal{E}(A) = n \rangle)$ must also be n.[3]

So global deference implies local deference. But the converse is false.

Example 1 (Gaifman, 1988)**.** *There are three worlds in $\mathcal{W}$, which we will call '1', '2', and '3'. At world 1, the expert gives 50% probability to 1 and 50% probability to 2. At world 2, the expert gives 50% probability to 2 and 50% probability to 3. At world 3, the expert gives 50% probability to 3 and 50% probability to 1.*

We can represent the expert from example 1 with a square matrix, where the entry

3. Conglomerability tells us that, if Q is partitioned by $\{P_1, P_2, \ldots\}$, and $C(A \mid P_i) = n$, for each P_i, then $C(A \mid Q) = n$ as well. So long as we assume that $\mathcal{W}$ is at most countably infinite, conglomerability follows from our assumption that C is a countably additive probability.

in the rth row and the cth column gives us the probability which the expert gives to world c at the world r, $\mathcal{E}_r(c)$. (Throughout, I'm going to adopt the convention of using expressions like '$\mathcal{E}_1(3)$' and '$\mathcal{E}_2(1 \vee 3)$' for $\mathcal{E}_1(\{3\})$ and $\mathcal{E}_2(\{1, 3\})$, respectively.)

$$
\begin{array}{c} \\ \mathcal{E}_1 \\ \mathcal{E}_2 \\ \mathcal{E}_3 \end{array}
\begin{array}{c}
\begin{array}{ccc} 1 & 2 & 3 \end{array} \\
\begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix}
\end{array}
$$

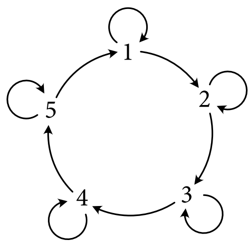
Gaifman's example is interesting because, if you spread your credences uniformly—$C(1) = C(2) = C(3) = 1/3$, then you will defer to $\mathcal{E}$ *locally*, but not *globally*. For instance, your credence in $1 \vee 2$, given $\langle \mathcal{E}(1 \vee 2) = 1/2 \rangle$, is just $C(1 \vee 2 \mid 2 \vee 3)$ (since $\mathcal{E}$'s credence in $1 \vee 2$ is $1/2$ at worlds 2 and 3), and if your credences are uniform, then $C(1 \vee 2 \mid 2 \vee 3)$ is $1/2$. Moreover, as you can check for yourself, this works for every $A \subseteq \{1, 2, 3\}$ and every n. $C(A \mid \langle \mathcal{E}(A) = n \rangle) = n$ whenever $\langle \mathcal{E}(A) = n \rangle$ is given a credence greater than 0. So, with the uniform credence distribution, you defer to $\mathcal{E}$ locally. But you do not defer globally, since $C(2 \mid \langle \mathcal{E} = \mathcal{E}_2 \rangle) = C(2 \mid 2) = 1$, even though $\mathcal{E}_2$'s credence in 2 is only $1/2$.

We can pull the same trick with more worlds. For instance, if $\mathcal{W} = \{1, 2, 3, 4, 5\}$, and the expert function is given by this matrix,

$$
\begin{array}{c} \\ \mathcal{E}_1 \\ \mathcal{E}_2 \\ \mathcal{E}_3 \\ \mathcal{E}_4 \\ \mathcal{E}_5 \end{array}
\begin{array}{c}
\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{bmatrix}
\end{array}
$$

Then the uniform credence distribution (the one which gives credence $1/5$ to every world) will defer locally, but not globally, to this expert.

Another helpful way of looking at an expert function, $\mathcal{E}$, is with a Kripke frame $(\mathcal{W}, R)$, where we stipulate that world w 'sees' a world, x, wRx, iff the expert at w gives positive probability to x, $\mathcal{E}_w(x) > 0$. For illustration, the expert from the 5 world model above gives rise to the following frame.



Call any collection of worlds like this—a collection $\mathscr{C}$, containing at least 3 worlds, such that each world in $\mathscr{C}$ bears R to itself and exactly one other world, and every $w \in \mathscr{C}$ bears R^+ (the transitive closure of R) to every other world in $\mathscr{C}$—a *cycle*. If

$\mathscr{C}$ is a cycle and, moreover, for every $w \in \mathscr{C}$, $\mathcal{E}_w$ gives exactly half of its probability to w, then I'll say that $\mathscr{C}$ is a 'half-cycle'. Finally, if the frame $\mathcal{E}$ gives rise to contains *some* half-cycle, then I'll say that $\mathcal{E}$ is a *half-cyclic* expert.

Half-Cyclicity An expert $\mathcal{E}$ is *half-cyclic* if and only if the frame it generates contains a cycle $\mathscr{C}$ such that, for every $w \in \mathscr{C}$, $\mathcal{E}_w(w) = 1/2$.

Whenever an expert is half-cyclic, it will be possible to defer to them locally but not globally. In the appendix, I prove the following theorems:

Theorem 1. *If $\mathcal{E}$ is half-cyclic, then C will defer to $\mathcal{E}$ locally but not globally if it spreads its credence uniformly over each half-cycle and gives a probability of* 0 *to any world not in a half-cycle.*

Theorem 2. *If $\mathcal{E}$ is half-cyclic, then C defers to $\mathcal{E}$ locally only if C is uniform over every half-cycle.*

When else is it possible to defer locally but not globally? *Never*. The half-cyclic experts are *the only ones* to whom you can defer locally without deferring to them globally. In the appendix, I offer a proof of the following theorem:

Theorem 3. *If $\mathcal{E}$ is not half-cyclic, then C defers to $\mathcal{E}$ locally iff C defers to $\mathcal{E}$ globally.*

This tells us that Gaifman's example is *incredibly* singular. We can vary the size of the half-cycles, but that's it. In no other kind of case do the local and global norms pull apart.

3 | WHY THE DIFFERENCE IS PHILOSOPHICALLY NEGLIGIBLE

In my view, this theorem teaches us something helpful. It teaches us that we don't have to concern ourselves with the differences between local and global norms of deference. For it teaches us that there is no philosophically plausible reason anyone could have to endorse a norm of local deference while denying the corresponding norm of global deference. I'll give two independent reasons to think that such a position is implausible in §3.1 and §3.2 below.

3.1 | Drawing New Distinctions

Suppose that we begin with the model from example 1, and we simply introduce a new distinction. Perhaps, for each world w, we introduce two new worlds, w_H and w_T, where w_H is the possibility previously represented by w, plus the additional information that a flipped coin landed heads, and w_T is the possibility previous represented by w, plus the additional information that the coin landed tails. And suppose that each possible expert gives a probability of 1/2 to the coin landing heads and a probability of 1/2 to the coin landing tails, and takes the outcome of the coin flip to be independent of whether $1, 2$, or 3. Then, including this additional distinction gives us the following expert:

$$
\begin{array}{c} \\ \mathcal{E}_{1_H} \\ \mathcal{E}_{1_T} \\ \mathcal{E}_{2_H} \\ \mathcal{E}_{2_T} \\ \mathcal{E}_{3_H} \\ \mathcal{E}_{3_T} \end{array}
\begin{array}{c} \begin{array}{cccccc} 1_H & 1_T & 2_H & 2_T & 3_H & 3_T \end{array} \\
\begin{bmatrix}
1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\
1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\
0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\
0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\
1/4 & 1/4 & 0 & 0 & 1/4 & 1/4 \\
1/4 & 1/4 & 0 & 0 & 1/4 & 1/4
\end{bmatrix} \end{array}
$$

And theorem 3 assures us that, while the half-cyclic expert from example 1 could be deferred to locally, this non-half-cyclic expert cannot.[4] Attending to an additional distinction like whether a coin landed heads or tails should only make a difference to whether the expert $\mathcal{E}$ is deserving of epistemic deference if there is something irrational about the probabilities $\mathcal{E}$ assigns to the coin landing heads or tails. But in this case, there is nothing irrational about $\mathcal{E}$'s probabilities. The coin is fair and independent of whether $1, 2,$ or 3. So conditional on 1, conditional on 2, and conditional on 3, the expert *should* divide their probability evenly between heads and tails. So, if a half-cyclic expert is deserving of epistemic deference, then, after we introduce a new, independent distinction—dividing each former possibility into an equally likely 'heads' and 'tails' possibility—the new expert should also be deserving of epistemic deference.

However, if you endorsed a local norm of deference while rejecting the corresponding global norm of deference, you would be forced to disagree. For then, you would think that introducing this new distinction *does* make a difference to whether $\mathcal{E}$ is deserving of epistemic deference. I take that to be rather implausible; so I take it to be rather implausible that a local norm of deference holds without the corresponding global norm holding.

3.2 | *Learning the Expert's Evidence*

In the introduction, I said that Lewis's *Principal Principle* tells you to locally defer to the future objective chances. That's true, but it's slightly misleading, because it *also* tells you to *globally* defer to the future objective chances. Lewis's principle has the form of what we can call a *conditional* local deference principle. It says that your initial or ur-prior credence function, C_0, should locally defer to the future objective

4. To see this without slogging through the proof of theorem 3, note that your credence in $\{1_H, 1_T, 2_H\}$, conditional on the expert's credence in $\{1_H, 1_T, 2_H\}$ being $1/4$, is your credence in $\{1_H, 1_T, 2_H\}$, conditional on $\{2_H, 2_T\}$ (since these are the possibilities in which $\mathcal{E}(\{1_H, 1_T, 2_H\}) = 1/4$). So, if you are going to defer to $\mathcal{E}$ locally, then your credence in 2_H must be 1/4 your credence in 2_T. But then, take the proposition $\{1_H, 1_T, 2_T\}$. Since $\langle \mathcal{E}(\{1_H, 1_T, 2_T\}) = 1/4 \rangle$ is $\{2_H, 2_T\}$, you can defer to $\mathcal{E}$ locally only if your credence in 2_T is 1/4 your credence in 2_H. But this can only happen if your credence in both 2_H and 2_T is zero. But there was nothing special about 2. Run the same argument swapping '3' for '2', '2' for '1', and '1' for '3', and you get that your credence in both 3_H and 3_T must be zero. Swap the labels again, and the same argument gets that your credence in both 1_H and 1_T must be zero. But then your credences aren't probabilistic.

chances *conditional on any admissible evidence*. That is: for any proposition A, any future time t, any number n, and any admissible evidence proposition F, you should satisfy the equality

$$C_0(A \mid \langle \mathcal{C}h_t(A) = n \rangle \cap F) = n$$

If your total evidence is admissible, then conditionalisation says that $C(-)$ should be $C_0(- \mid F)$, so this norm implies the one from the introduction.

Lewis thought (back in 1980, at least) that propositions about the time t chances were themselves admissible. And he thought that admissibility was closed under conjunction. So we can take any probability function ch such that $ch(A) = n$, and any admissible evidence F, and the *Principal Principle* will require that

$$C_0(A \mid \langle \mathcal{C}h_t(A) = n \rangle \cap \langle \mathcal{C}h_t = ch \rangle \cap F) = n$$

Now, notice that $\langle \mathcal{C}h_t(A) = n \rangle \cap \langle \mathcal{C}h_t = ch \rangle$ is just $\langle \mathcal{C}h_t = ch \rangle$, and n is just $ch(A)$, so this is equivalent to a conditional global norm which requires that

$$C_0(A \mid \langle \mathcal{C}h_t = ch \rangle \cap F) = ch(A)$$

which is why, in his original 1980 article, Lewis was able to freely move back and forth between a local and a global version of the *Principal Principle*.

There's a general lesson here. For we often don't just want to suggest that you should defer to an expert *now*, given the evidence you *currently* have. We generally want to say that you should *continue* to defer to them, even after you've received certain kinds of evidence. Taking Lewis's lead, call this kind of evidence 'admissible'. Then, consider the following two ways of showing deference:

Conditional Local Deference You conditionally locally defer to an expert, $\mathcal{E}$, iff, for any proposition A, any number n, and any admissible evidence F, your credence in A, given that $\mathcal{E}$'s probability for A is n, and given F, is n.

$$C(A \mid \langle \mathcal{E}(A) = n \rangle \cap F) = n$$

Conditional Global Deference You conditionally globally defer to an expert, $\mathcal{E}$, iff, for any proposition A, any probability function E, and any admissible evidence F, your credence in A, given that $\mathcal{E}$'s entire probability function is E, and given F, is whatever probability E gives to A.

$$C(A \mid \langle \mathcal{E} = E \rangle \cap F) = E(A)$$

Intuitively, evidence F is *admissible* iff you should continue deferring to $\mathcal{E}$ even after you have F as your total evidence. Here's a general principle about admissible evidence that we should want to accept in a wide variety of cases: if F might be the *expert's* total evidence, then F is admissible.

Admissibility of Expert Evidence For any possible world w such that $C(w) > 0$,

$\mathcal{E}$'s total evidence at w is admissible.

In other words, if you should show epistemic deference to $\mathcal{E}$, then for any possible world w with positive credence, after learning $\mathcal{E}$'s total evidence at w, you should continue to show epistemic deference to $\mathcal{E}$.

If we accept the admissibility of expert evidence, then there will be no difference between a norm of conditional local deference and a norm of conditional global deference. To appreciate this, notice that a norm of conditional local deference says not only that $C(-)$ should locally defer to $\mathcal{E}$, but also that, for any admissible F, $C(- \mid F)$ should locally defer to $\mathcal{E}$, too. But theorem 3 teaches us that the only way it could be possible for $C(- \mid F)$ to defer to $\mathcal{E}$ locally but not globally is if $\mathcal{E}$ is a half-cyclic expert. But then, $\mathcal{E}$'s evidence at every world w in a half-cycle is $w \vee wR$, where 'wR' is w's successor in the cycle. If expert evidence is admissible, then for any world w with positive credence, $w \vee wR$ is admissible, and a norm of conditional local deference will require that $C(- \mid w \vee wR)$ locally defer to $\mathcal{E}$. Since $C(- \mid w \vee wR)$ only gives positive probability to two worlds within a cycle, it does not spread its probability uniformly over every cycle. So theorem 2 assures us that it does not locally defer to $\mathcal{E}$. So, if expert evidence is admissible, then it is impossible to conditionally locally defer to a half-cycle expert. So, if expert evidence is admissible, then it is possible to conditionally *locally* defer to all and only the experts it is possible to conditionally *globally* defer to. And whenever you conditionally *locally* defer, you will also conditionally *globally* defer.

That is: if expert evidence is admissible, then there is no difference between a norm of conditional local deference and the corresponding norm of conditional global deference. It is not plausible to think that some expert $\mathcal{E}$ is deserving of epistemic deference, but that $\mathcal{E}$ might not be deserving of deference, were you to learn what $\mathcal{E}$'s evidence is. So it is not plausible to endorse a norm of local deference without endorsing the corresponding global norm.

A | PROOFS

Recall, an expert function $\mathcal{E}$ generates a Kripke frame $(\mathcal{W}, R)$ where, for any two $w, x \in \mathcal{W}$ wRx iff $\mathcal{E}_w(x) > 0$. If wRx, I'll say colloquially that w sees x.

Lemma 1. *If C defers to $\mathcal{E}$ locally, then, for any world $w \in \mathcal{W}$, if $C(w) > 0$, then wRw.*

Proof. Suppose otherwise. Then, $w \in \langle \mathcal{E}(w) = 0 \rangle$, so $C(w \mid \langle \mathcal{E}_w(w) = 0 \rangle)$, is defined and not equal to 0. So you don't defer to $\mathcal{E}$ locally. Contradiction. □

Theorem 1. *If $\mathcal{E}$ is half-cyclic, then C will defer to $\mathcal{E}$ locally but not globally if it spreads its credence uniformly over each half-cycle and gives a probability of 0 to any world not in a half-cycle.*

Proof. Assume that $\mathcal{E}$ is half-cyclic and that C gives only positive credence to the worlds in some half-cycle. Suppose further than, for any two worlds in the same half-cycle, w and x, $C(w) = C(x)$. We will now show that C defers to $\mathcal{E}$ locally.

For every world in a half-cycle, w, and each $A \subseteq \mathcal{W}$, there are only three possibilities for the value of $\mathcal{E}_w(A)$: 0, 1/2, and 1. For take any $A \subseteq \mathcal{W}$ and any w in a half-cycle. Use 'wR' for the unique $x \neq w$ such that wRx. Then, either (i) both w and wR are in A; (ii) exactly one of w and wR are in A; or (iii) neither w nor wR are in A. If (i), then $\mathcal{E}_w(A) = 1$. If (ii), then $\mathcal{E}_w(A) = 1/2$. And if (iii), then $\mathcal{E}_w(A) = 0$.

Now, take any $A \subseteq \mathcal{W}$, any $n \in \{0, 1/2, 1\}$, and any half-cycle $\mathscr{C}$. We will show that

(1) $$||A \cap \langle \mathcal{E}(A) = n \rangle \cap \mathscr{C}|| = n \cdot ||\langle \mathcal{E}(A) = n \rangle \cap \mathscr{C}||$$

(where '$||A||$' is the cardinality of A.)

Start with the case $n = 0$. For every $w \in \langle \mathcal{E}(A) = 0 \rangle$ such that $C(w) > 0$, $w \notin A$ by lemma 1, hence $||A \cap \langle \mathcal{E}(A) = 0 \rangle \cap \mathscr{C}|| = 0$, for every $\mathscr{C}$.

Next, consider the case $n = 1$. For every $w \in \langle \mathcal{E}(A) = 1 \rangle$ such that $C(w) > 0$, $w \in A$ by lemma 1. So, for every half-cycle $\mathscr{C}$, $||A \cap \langle \mathcal{E}(A) = 1 \rangle \cap \mathscr{C}|| = ||\langle \mathcal{E}(A) = 1 \rangle \cap \mathscr{C}||$.

Next consider the case $n = 1/2$. Every world with positive credence is in a half-cycle, so for every $w \in \langle \mathcal{E}(A) = 1/2 \rangle$ such that $C(w) > 0$, exactly one of w and wR are in A. If $w \notin A$ but $wR \in A$, then call w an *entrance world*. If $w \in A$ but $wR \notin A$, then call w an *exit world*. As we travel around each half-cycle $\mathscr{C}$, we must enter A as many times as we leave it, so for each half-cycle $\mathscr{C}$, there are as many entrance worlds in that cycle as there are exit worlds. $\langle \mathcal{E}(A) = 1/2 \rangle$ contains only entrance and exit worlds, and the exit worlds are exactly those members of $\langle \mathcal{E}(A) = 1/2 \rangle$ which are in A. So, for each half-cycle $\mathscr{C}$,

$$||A \cap \langle \mathcal{E}(A) = 1/2 \rangle \cap \mathscr{C}|| = 1/2 \cdot ||\langle \mathcal{E}(A) = 1/2 \rangle \cap \mathscr{C}||$$

With (1) established, we can show that, if C is uniform over every half-cycle, and gives only positive credence to worlds in half-cycles, then C defers locally to $\mathcal{E}$. For, in that case, there is a collection of weights $\lambda_{\mathscr{C}}$, one for each half-cycle $\mathscr{C}$, such that, for every $A \subseteq \mathcal{W}$,

$$C(A) = \sum_{\mathscr{C}} \lambda_{\mathscr{C}} \cdot ||A \cap \mathscr{C}||$$

In particular, for $n \in \{0, 1/2, 1\}$,

$$C(A \cap \langle \mathcal{E}(A) = n \rangle) = \sum_{\mathscr{C}} \lambda_{\mathscr{C}} \cdot \|A \cap \langle \mathcal{E}(A) = n \rangle \cap \mathscr{C}\|$$

and

$$C(\langle \mathcal{E}(A) = n \rangle) = \sum_{\mathscr{C}} \lambda_{\mathscr{C}} \cdot \|\langle \mathcal{E}(A) = n \rangle \cap \mathscr{C}\|$$

By (1), then,

$$C(A \cap \langle \mathcal{E}(A) = n \rangle) = n \cdot \sum_{\mathscr{C}} \lambda_{\mathscr{C}} \cdot \|\langle \mathcal{E}(A) = n \rangle \cap \mathscr{C}\|$$

So, for any $A \subseteq \mathcal{W}$, and any n,

$$C(A \mid \langle \mathcal{E}(A) = n \rangle) = \frac{C(A \cap \langle \mathcal{E}(A) = n \rangle)}{C(\langle \mathcal{E}(A) = n \rangle)} = \frac{n \cdot \sum_{\mathscr{C}} \lambda_{\mathscr{C}} \cdot \|\langle \mathcal{E}(A) = n \rangle \cap \mathscr{C}\|}{\sum_{\mathscr{C}} \lambda_{\mathscr{C}} \cdot \|\langle \mathcal{E}(A) = n \rangle \cap \mathscr{C}\|} = n$$

So C defers to $\mathcal{E}$ locally.

But C does not defer to $\mathcal{E}$ globally, since, for any world w such that $C(w) > 0$, w is in a half-cycle, and $\langle \mathcal{E} = \mathcal{E}_w \rangle = \{w\}$, so $C(w \mid \langle \mathcal{E} = \mathcal{E}_w \rangle) = 1 \neq \mathcal{E}_w(w) = 1/2$. □

Theorem 2. *If $\mathcal{E}$ is half-cyclic, then C defers to $\mathcal{E}$ locally only if C is uniform over every cycle.*

Proof. For each half-cycle, use '$\mathscr{C}$' for the set of worlds in the cycle, and for any world w, let wR be the unique $x \neq w$ such that wRx. Then, take any world $w \in \mathscr{C}$. If $C(w) > 0$, then $C(\langle \mathcal{E}(wR) = 1/2 \rangle) = C(w \vee wR) > 0$. So $C(wR \mid \langle \mathcal{E}(wR) = 1/2 \rangle)$ is defined. Since C defers to $\mathcal{E}$ locally, $C(wR \mid \langle \mathcal{E}(wR) = 1/2 \rangle) = C(wR \mid w \vee wR)$ must be $1/2$. So $C(wR)$ must be equal to $C(w)$. The world w was arbitrary, so the credence of every world in the cycle must be the same as the credence of the unique distinct world it 'sees'. So every world in the cycle must have the same credence. The cycle was arbitrary, so if C defers locally to a half-cyclic $\mathcal{E}$, then C is uniform over every cycle. □

Lemma 2. *If C defers to $\mathcal{E}$ locally, then, for any two worlds $w, x \in \mathcal{W}$, if wRx and $C(w) > 0$, then $\mathcal{E}_w(x) = \mathcal{E}_x(x)$.*

Proof. Suppose, for *reductio*, that C defers to $\mathcal{E}$ locally and that, for some two worlds $w, x \in \mathcal{W}$, wRx, $C(w) > 0$, and $\mathcal{E}_w(x) \neq \mathcal{E}_x(x)$. Then, $x \notin \langle \mathcal{E}(x) = \mathcal{E}_w(x) \rangle$. Since $C(w) > 0$, $C(\langle \mathcal{E}(x) = \mathcal{E}_w(x) \rangle) > 0$. So $C(x \mid \langle \mathcal{E}(x) = \mathcal{E}_w(x) \rangle)$ is defined and equal to 0. Since C defers to $\mathcal{E}$ locally, $\mathcal{E}_w(x)$ must be 0, which contradicts our assumption that wRx. □

Lemma 3. *If C defers to $\mathcal{E}$ locally, then, for any two worlds $w, x \in \mathcal{W}$, if wRx and $C(w) > 0$, then $C(x) > 0$.*

Proof. By lemma 2, $\mathcal{E}_w(x) = \mathcal{E}_x(x) > 0$. Since $C(w) > 0$, $C(\langle \mathcal{E}(x) = \mathcal{E}_x(x) \rangle) > 0$. So $C(x \mid \langle \mathcal{E}(x) = \mathcal{E}_x(x) \rangle)$ is defined. Since C defers to $\mathcal{E}$ locally, $C(x \mid \langle \mathcal{E}(x) = \mathcal{E}_x(x) \rangle) = C(x)/C(\langle \mathcal{E}(x) = \mathcal{E}_x(x) \rangle)$ must be equal to $\mathcal{E}_x(x) > 0$, which requires that $C(x) > 0$. □

Lemma 4. *If C defers to $\mathcal{E}$ locally, then, for any world w in the support of C, and any world x, if wRx, then xR^+w. (R^+ is the transitive closure of R.)*

Proof. Take an arbitrary world w in the support of C. If w doesn't see any world besides itself, then the lemma is trivial. So suppose there's some x such that wRx. Let $A \equiv \{y \neq w \mid xR^+y\}$. Since $C(w) > 0$, lemmas 3 and 1 tell us that xRx, so $x \in A$, and $\mathcal{E}_w(A) > 0$. Since C defers locally to $\mathcal{E}$, $C(A \mid \langle \mathcal{E}(A) = \mathcal{E}_w(A)\rangle) = \mathcal{E}_w(A) > 0$. So $A \cap \langle \mathcal{E}(A) = \mathcal{E}_w(A)\rangle$ must be non-empty. But this is only possible if there are some worlds $y \in A$ such that $\mathcal{E}_y(A) = \mathcal{E}_w(A) < 1$. But the only way a world $y \in A$ could have $\mathcal{E}_y(A) < 1$ is if $\mathcal{E}_y(w) > 0$—by the definition of A, any world other than w that y sees would itself be in A. So there's some world $y \in A$ such that yRw. But if $y \in A$ then xR^+y. And if xR^+y and yRw, then xR^+w. □

Lemma 5. *If C defers to $\mathcal{E}$ locally and for some pair of distinct worlds w, x in the support of C, wRx and xRw, then, for any world y in the support of C, $yRw \leftrightarrow yRx$.*

Proof. Suppose (for *reductio*) that C defers locally to $\mathcal{E}$, $C(w) > 0$, wRx, and xRw, yet there's some world y in the support of C such that y sees one of w or x without seeing the other. Without loss of generality, suppose that yRw and $\neg yRx$. By lemma 2, $\mathcal{E}_y(w) = \mathcal{E}_w(w)$. And since $\mathcal{E}_y(x) = 0$, $\mathcal{E}_y(w \vee x) = \mathcal{E}_w(w)$. But $\mathcal{E}_w(w \vee x) \neq \mathcal{E}_w(w)$ and $\mathcal{E}_x(w \vee x) \neq \mathcal{E}_w(w)$. So $y \in \langle \mathcal{E}(w \vee x) = \mathcal{E}_w(w)\rangle$ but $w, x \notin \langle \mathcal{E}(w \vee x) = \mathcal{E}_w(w)\rangle$. So

$$C(w \vee x \mid \langle \mathcal{E}(w \vee x) = \mathcal{E}_w(w)\rangle) = 0$$

and C doesn't defer to $\mathcal{E}$ locally. Contradiction. □

Lemma 6. *If C defers to $\mathcal{E}$ locally, $C(w) > 0$, wRx, xRw, and xRy, then wRy, yRw, and yRx, too (i.e., if w and x see each other and x sees y, then each of w, x, and y see all of w, x, and y).*

Proof. Assume C defers to $\mathcal{E}$ locally, $C(w) > 0$, wRx, xRw, and xRy. We will first show that either wRy or yRw. Suppose for *reductio* that $\neg wRy$ and $\neg yRw$. Then, $\mathcal{E}_w(w \vee y) = \mathcal{E}_w(w)$ and $\mathcal{E}_y(w \vee y) = \mathcal{E}_y(y)$. But $\mathcal{E}_x(w \vee y) = \mathcal{E}_x(w) + \mathcal{E}_x(y) = \mathcal{E}_w(w) + \mathcal{E}_y(y)$ (by lemma 2). So $x \in \langle \mathcal{E}(w \vee y) = \mathcal{E}_x(w \vee y)\rangle$, but $w, y \notin \langle \mathcal{E}(w \vee y) = \mathcal{E}_x(w \vee y)\rangle$. So

$$C(w \vee y \mid \langle \mathcal{E}(w \vee y) = \mathcal{E}_x(w \vee y)\rangle) = 0$$

And C doesn't defer to $\mathcal{E}$ locally. Contradiction.

So either (a) yRw or (b) wRy.

Suppose first that (a) yRw. Then, by lemma 5, yRx, too. So xRy and yRx. So, by lemma 5, $wRx \leftrightarrow wRy$. Since wRx, wRy. So each of w, x, and y sees all of w, x, and y.

Next, suppose that (b) wRy. Then, wRx, xRw, xRy, and wRy. Suppose for *reductio* that $\neg yRw$ and $\neg yRx$. And let $Rw \equiv \{z \neq w, x, y \mid zRw \wedge \neg zRy\}$. Let $Rwy \equiv \{z \neq w, x, y \mid zRw \wedge zRy\}$. And let $Ry \equiv \{z \neq w, x, y \mid \neg zRw \wedge zRy\}$. By lemma 5, every world in Rw and Rwy sees x, and no world in Ry sees x (if it did, it would also see w by lemma 5, and so it would be in Rwy, not Ry). Similarly, lemma 5 tells us that any world which sees x is either in Rw or Rwy.

There are two cases to consider. Either (b1) $\mathcal{E}_y(y) = \mathcal{E}_w(w \vee x)$ or (b2) $\mathcal{E}_y(y) \neq \mathcal{E}_w(w \vee x)$. Start with (b1). Then, since $\neg yRw$, $\mathcal{E}_y(w \vee y) = \mathcal{E}_y(y)$, whereas for any $z \in Rw$, $\mathcal{E}_z(w \vee y) = \mathcal{E}_z(w) = \mathcal{E}_w(w)$ by lemma 2. Since $\mathcal{E}_y(y) = \mathcal{E}_w(w \vee x) > \mathcal{E}_w(w)$, for any $z \in Rw$, $\mathcal{E}_y(y) > \mathcal{E}_z(w \vee y)$. So $Rw \cap \langle \mathcal{E}(w \vee y) = \mathcal{E}_y(y)\rangle = \varnothing$. Likewise, for any $z \in Rwy$, $\mathcal{E}_z(w \vee y) = \mathcal{E}_z(w) + \mathcal{E}_z(y) = \mathcal{E}_w(w) + \mathcal{E}_y(y)$ by lemma 2. So $Rwy \cap \langle \mathcal{E}(w \vee y) = \mathcal{E}_y(y)\rangle = \varnothing$. So $\langle \mathcal{E}(w \vee y) = \mathcal{E}_y(y)\rangle = Ry \cup \{y\}$. So

$$C(w \vee y \mid \langle \mathcal{E}(w \vee y) = \mathcal{E}_y(y)\rangle) = \frac{C(y)}{C(y) + C(Ry)} \tag{2}$$

And, by lemma 2, $\langle \mathcal{E}(y) = \mathcal{E}_y(y) \rangle = Rwy \cup Ry \cup \{w, x, y\}$, so

$$(3) \qquad C(y \mid \langle \mathcal{E}(y) = \mathcal{E}_y(y) \rangle) = \frac{C(y)}{C(y) + C(Ry) + C(Rwy) + C(w) + C(x)}$$

But, since C defers to $\mathcal{E}$ locally, (2) and (3) together imply that either $C(y) = 0$ or else $C(Rwy) = C(w) = C(x) = 0$. Either possibility contradicts our assumption that $C(w) > 0$ and wRy (since lemma 3 then implies that $C(y) > 0$). So case (b1) leads to a contradiction.

Next consider case (b2) $\mathcal{E}_y(y) \neq \mathcal{E}_w(w \vee x)$. Then, $C(Rw)$ must be zero. For suppose $C(Rw) > 0$. Then, by lemma 2, $Rw \subseteq \langle \mathcal{E}(w \vee x \vee y) = \mathcal{E}_w(w \vee x) \rangle$. But $\mathcal{E}_w(w \vee x \vee y) \neq \mathcal{E}_w(w \vee x)$, $\mathcal{E}_x(w \vee x \vee y) \neq \mathcal{E}_w(w \vee x)$, and $\mathcal{E}_y(w \vee x \vee y) = \mathcal{E}_y(y) \neq \mathcal{E}_w(w \vee x)$. So

$$C(w \vee y \vee x \mid \langle \mathcal{E}(w \vee x \vee y) = \mathcal{E}_w(w \vee x) \rangle)$$

is defined and equal to 0, and not $\mathcal{E}_w(w \vee x)$. So C doesn't defer to $\mathcal{E}$ locally. Contradiction. So it must be that $C(Rw) = 0$. But now, $C(\langle \mathcal{E}(w \vee x \vee y) = \mathcal{E}_y(y) \rangle) = C(Ry) + C(y)$. Neither w nor x nor any world in Rwy is in $\langle \mathcal{E}(w \vee x \vee y) = \mathcal{E}_y(y) \rangle$, since all of them give a credence of $\mathcal{E}_w(w) + \mathcal{E}_x(x) + \mathcal{E}_y(y)$ to $w \vee x \vee y$ (by lemma 2). And no world in Rw has positive credence. So

$$(4) \qquad C(w \vee x \vee y \mid \langle \mathcal{E}(w \vee x \vee y) = \mathcal{E}_y(y) \rangle) = \frac{C(y)}{C(Ry) + C(y)}$$

And, by lemma 2, $\langle \mathcal{E}(y) = \mathcal{E}_y(y) \rangle = Ry \cup Rwy \cup \{w, x, y\}$. So

$$(5) \qquad C(y \mid \langle \mathcal{E}(y) = \mathcal{E}_y(y) \rangle) = \frac{C(y)}{C(Ry) + C(y) + C(Rwy) + C(w) + C(x)}$$

But, since C defers to $\mathcal{E}$ locally, (4) and (5) together imply that either $C(y) = 0$ or else $C(Rwy) = C(w) = C(x) = 0$. Either possibility contradicts our assumption that $C(w) > 0$ and wRy (since lemma 3 then implies that $C(y) > 0$). So case (b2) also leads to a contradiction.

So our assumption that $\neg yRw$ and $\neg yRx$ has led to a contradiction. So it must be that either yRw or yRx. If yRw, then, by lemma 5, yRx, too. And if yRx, then, by lemma 5, yRw, too. So either way, wRx, wRy, xRw, xRy, yRw, and yRx. So we have that each of w, x, and y sees all of w, x, and y.

So, whether (a) or (b), each of w, x, and y sees all of w, x, and y. □

Lemma 7. *If C defers to $\mathcal{E}$ locally, $C(y) > 0$, wRx, xRw, and yRx, then xRy, wRy, and yRw, too (i.e., if w and x see each other and y sees x, then each of w, x, and y see all of w, x, and y).*

Proof. Suppose C defers to $\mathcal{E}$ locally, $C(y) > 0$, wRx, xRw, and yRx. Then, by lemma 5, yRw, too. Now, let $Ry \equiv \{z \neq w, x, y \mid zRy \wedge \neg zRw\}$, let $Rw \equiv \{z \neq w, x, y \mid \neg zRy \wedge zRw\}$, and let $Rwy \equiv \{z \neq w, x, y \mid zRy \wedge zRw\}$. By lemma 5, every world in Rwy and Rw sees x, and no world in Ry sees x (if it did, it would also see w by lemma 5, and it would be in Rwy, not Ry). Similarly, lemma 5 tells us that any world which sees x is either in Rwy or Rw.

Now, suppose for *reductio* that $\neg wRy$ and $\neg xRy$. Then, by lemma 2, $\mathcal{E}_y(y \vee w \vee x) = \mathcal{E}_y(y) + \mathcal{E}_w(w) + \mathcal{E}_x(x)$, whereas $\mathcal{E}_w(y \vee w \vee x) = \mathcal{E}_x(y \vee w \vee x) = \mathcal{E}_w(w) + \mathcal{E}_x(x) = \mathcal{E}_w(w \vee x)$. So $y \notin \langle \mathcal{E}(y \vee w \vee x) = \mathcal{E}_w(w \vee x) \rangle$. For any $z \in Ry$, $\mathcal{E}_z(y \vee w \vee x) = \mathcal{E}_y(y)$. And for any $z \in Rwy$, $\mathcal{E}_z(y \vee w \vee x) = \mathcal{E}_y(y) + \mathcal{E}_w(w) + \mathcal{E}_z(z) \neq \mathcal{E}_w(w \vee x)$. So $\langle \mathcal{E}(y \vee w \vee x) = \mathcal{E}_w(w \vee x) \rangle =$

$Rw \cup \{w, x\}$. So

$$C(y \vee w \vee x \mid \langle \mathcal{E}(y \vee w \vee x) = \mathcal{E}_w(w \vee x) \rangle) = \frac{C(w) + C(x)}{C(w) + C(x) + C(Rw)} \tag{6}$$

Since any world which sees either w or x sees the other (by lemma 5), for any world in $z \in Rwy \cup Rw \cup \{y, w, x\}$, $\mathcal{E}_z(w \vee x) = \mathcal{E}_w(w) + \mathcal{E}_x(x) = \mathcal{E}_w(w \vee x)$, by lemma 2. No other worlds see either w or x. So $\langle \mathcal{E}(w \vee x) = \mathcal{E}_w(w \vee x) \rangle = Rwy \cup Rw \cup \{y, w, x\}$. So

$$C(w \vee x \mid \langle \mathcal{E}(w \vee x) = \mathcal{E}_w(w \vee x) \rangle) = \frac{C(w) + C(x)}{C(w) + C(x) + C(y) + C(Rwy) + C(Rw)} \tag{7}$$

But, since C defers locally to $\mathcal{E}$, equations (6) and (7) together imply that either $C(w) = C(x) = 0$ or else $C(y) = C(Rwy) = 0$. But either possibility contradicts our assumption that $C(y) > 0$ and yRx (since lemma 3 then implies that $C(x) > 0$). So our assumption has led to a contradiction.

So either wRy or xRy. But then each of w, x, and y sees all of w, x, and y by lemma 6. □

Definition 1. *An* ***S5 cluster*** *is a non-empty set $\mathscr{S} \subseteq \mathcal{W}$ such that, for every $w \in \mathscr{S}$, $\{x \mid wRx\} = \mathscr{S}$. An S5 cluster $\mathscr{S}$ is* ***immodest*** *iff $\mathcal{E}_w = \mathcal{E}_x$ for every $w, x \in \mathscr{S}$. Else, $\mathscr{S}$ is modest.*

Lemma 8. *If C defers to $\mathcal{E}$ locally, then, for any world w in the support of C, if there are two distinct worlds $x, y \in wR^+ \equiv \{z \mid wR^+z\}$ such that xRy and yRx, then wR^+ is an S5 cluster.*

Proof. Suppose C defers locally to $\mathcal{E}$, $C(w) > 0$, and for two distinct worlds $x, y \in \{z \mid wR^+z\}$, xRy and yRx. Let $wR^+ \equiv \{z \mid wR^+z\}$. Fix an enumeration of the worlds in $wR^+ \setminus \{x, y\}$ such that, if z_i comes before z_j in the enumeration, then the shortest R-chain[5] from x to z_i is shorter than the shortest R-chain from x to z_j. (We will eventually show that xRz and zRx, for every $z \in wR^+$, but for now we only assume that there is some finite R-chain from x to each $z \in wR^+$.)

Base Case: take z_1. Since z_1 begins the enumeration, either xRz_1 or else z_1Rx. By lemmas 6 and 7, every world in $\{x, y, z_1\}$ sees every other world in $\{x, y, z_1\}$. *Inductive Step*: Assume that every world in $wR^k \equiv \{x, y, z_1, z_2, \ldots, z_k\}$ sees every other world in wR^k. Take z_{k+1}. Given our choice of enumeration, there is some $u \in wR^k$ such that either uRz_{k+1} or $z_{k+1}Ru$. Take any other world $v \in wR^k$ ($v \neq u$). By the inductive hypothesis, uRv and vRu. So, by lemmas 6 and 7, $z_{k+1}Rv$, $z_{k+1}Ru$, uRz_{k+1}, and vRz_{k+1}. v was arbitrary, so z_{k+1} sees and is seen by every world in wR^k. So every world in $wR^{k+1} \equiv wR^k \cup \{z_{k+1}\}$ sees every other world in wR^{k+1}.

So every world in wR^+ sees every other world in wR^+. They cannot see any other worlds, else those worlds would also be in wR^+. So wR^+ is an S5 cluster. □

Lemma 9. *If C defers to $\mathcal{E}$ locally, wRx and $\neg xRw$, then either every world which sees w also sees x, or else $\mathcal{E}_w(w) = \mathcal{E}_x(x)$.*

Proof. Suppose that for some world, u, uRw and $\neg uRx$. Then, $C(w \vee x \mid \langle \mathcal{E}(w \vee x) = \mathcal{E}_u(w \vee x) \rangle) = \mathcal{E}_u(w \vee x) = \mathcal{E}_u(w)$. By lemma 2, $\mathcal{E}_u(w) = \mathcal{E}_w(w)$. But $\mathcal{E}_w(w \vee x) \neq \mathcal{E}_w(w)$, so $w \notin \langle \mathcal{E}(w \vee x) = \mathcal{E}_u(w \vee x) \rangle$. So in order for $C(w \vee x \mid \langle \mathcal{E}(w \vee x) = \mathcal{E}_u(w \vee x) \rangle)$ to not be 0, it

5. An R-chain is any sequence of worlds such that, for any two adjacent worlds in the sequence, w_i and w_{i+1}, either $w_i R w_{i+1}$ or else $w_{i+1} R w$.

must be that $x \in \langle \mathcal{E}(w \vee x) = \mathcal{E}_u(w \vee x)\rangle$. So $\mathcal{E}_x(w \vee x) = \mathcal{E}_x(x) = \mathcal{E}_u(w \vee x) = \mathcal{E}_w(w)$. So $\mathcal{E}_x(x) = \mathcal{E}_w(w)$. □

Lemma 10. *Suppose that C defers locally to $\mathcal{E}$. Then, for any world w in the support of C and any world $x \in \mathcal{W}$: if wRx and $\neg xRw$, then $\mathcal{E}_x(x) = \mathcal{E}_w(w)$.*

Proof. Suppose that $C(w) > 0$. If there's no world $x \neq w$ such that wRx, then the lemma is trivially satisfied. So suppose there's some $x \neq w$ such that wRx and $\neg xRw$. Suppose further (for *reductio*) that every world which sees w also sees x. Let $Rw \equiv \{y \neq w \mid yRw\}$. By Lemma 2, for every $y \in Rw$, $\mathcal{E}_y(w) = \mathcal{E}_w(w)$. Then,

$$C(w \mid \langle \mathcal{E}(w) = \mathcal{E}_w(w)\rangle) = \frac{C(w)}{C(w) + C(Rw)}$$

$$\text{and} \qquad C(w \vee x \mid \langle \mathcal{E}(w \vee x) = \mathcal{E}_w(w \vee x)\rangle) = \frac{C(w)}{C(w) + C(Rw)}$$

(The second equality follows because every world which sees w also sees x and so, by lemma 2, for every $y \in Rw$, $\mathcal{E}_y(w \vee x) = \mathcal{E}_w(w) + \mathcal{E}_x(x) = \mathcal{E}_w(w \vee x)$. And any world which sees only x must give a probability of only $\mathcal{E}_x(x)$ to $w \vee x$.) Because C defers to $\mathcal{E}$ locally, it then must be that $\mathcal{E}_w(w) = \mathcal{E}_w(w \vee x)$, which contradicts our assumption that wRx. So our assumption that every world which sees w also sees x has led to a contradiction. So lemma 9 tells us that $\mathcal{E}_w(w) = \mathcal{E}_x(x)$. □

Lemma 11. *If C defers to $\mathcal{E}$ locally, then, for every w in the support of C which is not in an S5 cluster, any every world x, if wRx, then $\mathcal{E}_x(x) = \mathcal{E}_w(w) = \mathcal{E}_w(x)$.*

Proof. Suppose C defers to $\mathcal{E}$ locally, and take a world w not in an S5 cluster. Take any world x such that wRx. Since w is not in an S5 cluster, lemma 8 tells us that $\neg xRw$. So lemma 10 tells us that $\mathcal{E}_x(x) = \mathcal{E}_w(w)$. And lemma 2 tells us that $\mathcal{E}_w(x) = \mathcal{E}_x(x)$. □

Lemma 12. *If $\mathscr{S}$ is an S5 cluster, $C(w) > 0$ for some $w \in \mathscr{S}$, and C defers to $\mathcal{E}$ locally, then $\mathscr{S}$ is an immodest S5 cluster.*

Proof. Suppose for *reductio* that $\mathscr{S}$ is modest, that $C(w) > 0$ for some $w \in \mathscr{S}$, and that C defers locally to $\mathcal{E}$. By lemma 3, C gives positive credence to every world in $\mathscr{S}$. Since $\mathscr{S}$ is modest, there are $x, y, z \in \mathscr{S}$ such that $\mathcal{E}_x(z) \neq \mathcal{E}_y(z)$. Now, either $\mathcal{E}_z(z) \neq \mathcal{E}_x(z)$ or $\mathcal{E}_z(z) \neq \mathcal{E}_y(z)$. Either way, there are two worlds $u, z \in \mathscr{S}$ such that $\mathcal{E}_u(z) \neq \mathcal{E}_z(z)$. However, since u and z are both in $\mathscr{S}$, $\mathcal{E}_u(z) \neq 0$. But since $z \notin \langle \mathcal{E}(z) = \mathcal{E}_u(z)\rangle$, $C(z \mid \langle \mathcal{E}(z) = \mathcal{E}_u(z)\rangle)$ must be zero, if defined. So it must not be defined. So $C(u)$ must be zero. Contradiction. □

Lemma 13. *If C defers to $\mathcal{E}$ locally and C invests all its credence in S5 clusters, then C defers to $\mathcal{E}$ globally.*

Proof. Suppose that C defers to $\mathcal{E}$ locally and invests all of its credence in S5 clusters. Take any S5 cluster $\mathscr{S}$, and any worlds $w, x \in \mathscr{S}$. By lemma 12, $\mathcal{E}_w = \mathcal{E}_x$, so $x \in \langle \mathcal{E} = \mathcal{E}_w\rangle$. x was arbitrary, so $\mathscr{S} \subseteq \langle \mathcal{E} = \mathcal{E}_w\rangle$. Moreover, for any world $z \notin \mathscr{S}$, either $\mathcal{E}_z \neq \mathcal{E}_w$ or else $C(z) = 0$. For, if $\mathcal{E}_z = \mathcal{E}_w$, then $\mathcal{E}_z(z) = 0$, so $\neg zRz$, so $C(z) = 0$ by lemma 1. So $C(\mathscr{S}) = C(\langle \mathcal{E} = \mathcal{E}_w\rangle)$.

Take any $A \subseteq \mathscr{S}$. Then, $\mathcal{E}_w(A) > 0$; whereas, for any $z \notin \mathscr{S}$, either $\mathcal{E}_z(A) = 0$ or else $C(z) = 0$. (For suppose that $\mathcal{E}_z(A) > 0$. Then, z sees some world in $\mathscr{S}$, but since $z \notin \mathscr{S}$, no world in $\mathscr{S}$ sees z. So z is not in an S5 cluster. So $C(z) = 0$.) Either way, $C(\langle \mathcal{E}(A) = \mathcal{E}_w(A)\rangle) =$

$C(\mathscr{S})$. So, for any $w \in \mathscr{S}$, and any $A \subseteq \mathscr{S}$, $C(\mathscr{S}) = C(\langle \mathcal{E} = \mathcal{E}_w \rangle)$ and $C(\langle \mathcal{E}(A) = \mathcal{E}_w(A) \rangle) = C(\mathscr{S})$. So $C(\langle \mathcal{E} = \mathcal{E}_w \rangle) = C(\langle \mathcal{E}(A) = \mathcal{E}_w(A) \rangle)$. So

$$C(A \mid \langle \mathcal{E} = \mathcal{E}_w \rangle) = C(A \mid \langle \mathcal{E}(A) = \mathcal{E}_w(A) \rangle) = \mathcal{E}_w(A)$$

Moreover, for any $B \subseteq \mathcal{W}$, there will be some $B \cap \mathscr{S} \subseteq \mathscr{S}$, and $C(B \mid \mathscr{S}) = C(B \cap \mathscr{S} \mid \mathscr{S}) = C(B \cap \mathscr{S} \mid \langle \mathcal{E} = \mathcal{E}_w \rangle) = \mathcal{E}_w(B \cap \mathscr{S}) = \mathcal{E}_w(B)$. $\mathscr{S}$ was arbitrary, so the same holds for any S5 cluster in the support of C. So C defers to $\mathcal{E}$ globally. □

Theorem 3. *If $\mathcal{E}$ is not half-cyclic, then C defers to $\mathcal{E}$ locally iff C defers to $\mathcal{E}$ globally.*

Proof. Because global deference implies local deference, it is enough to show that, if $\mathcal{E}$ is not half-cyclic and C defers to $\mathcal{E}$ locally, then C defers to $\mathcal{E}$ globally. So suppose, for *reductio*, that $\mathcal{E}$ is not half-cyclic and that C defers to $\mathcal{E}$ locally without deferring to $\mathcal{E}$ globally.

We will first show that, for any w in the support of C which is not in an S5 cluster, $\mathcal{E}_w(w) \neq 1/2$. Suppose the negation: for some w such that $C(w) > 0$ and w is not in an S5 cluster, $\mathcal{E}_w(w) = 1/2$. Since $\mathcal{E}_w(w) = 1/2$, there's some x_1 such that wRx_1. By lemma 11, $\mathcal{E}_{x_1}(x_1) = 1/2$ and $\mathcal{E}_w(x_1) = 1/2$. So w sees just itself and one other world, and $\mathcal{E}_w(w) = 1/2$. This is the base case. *Inductive Step*: suppose that, for some $x_i \in wR^+$, x_i sees only itself and x_{i+1}, and $\mathcal{E}_{x_i}(x_i) = \mathcal{E}_{x_i}(x_{i+1}) = 1/2$. Then, lemma 11 tells us that $\mathcal{E}_{x_{i+1}}(x_{i+1}) = 1/2$. So there's some x_{i+2} such that $x_{i+1}Rx_{i+2}$. So lemma 11 tells us that $\mathcal{E}_{x_{i+2}}(x_{i+2}) = 1/2$ and $\mathcal{E}_{x_{i+1}}(x_{i+2}) = 1/2$. So x_{i+1} sees only itself and x_{i+2}, and $\mathcal{E}_{x_{i+1}}(x_{i+1}) = \mathcal{E}_{x_{i+1}}(x_{i+2}) = 1/2$. Completing the induction: every $x \in wR^+$ sees itself and one other world, and $\mathcal{E}_x(x) = 1/2$. Since wRx_1, lemma 4 assures us that $x_1 R^+ w$, so this sequence of worlds must loop back on itself, and we have a half-cycle. Contradiction.

If every w in the support of C were in an S5 cluster, then lemma 13 tells us that C would defer to $\mathcal{E}$ globally, contradicting our assumption. So it must be that there is some world u in the support of C which is *not* in an S5 cluster. We've just learnt that every $w \in uR^+$ is such that $\mathcal{E}_w(w) \neq 1/2$. Moreover, it must be that $\mathcal{E}_w(w) < 1/2$. For if $\mathcal{E}_w(w)$ were greater than 1/2, it would either be 1 or between 1/2 and 1. If $\mathcal{E}_w(w) = 1$, then w sees only itself, and is an S5 cluster. Contradiction. If $\mathcal{E}_w(w) > 1/2$ but $\mathcal{E}_w(w) \neq 1$, then there would be some x such that wRx and $\mathcal{E}_w(x) \neq \mathcal{E}_w(w)$, contradicting lemma 11.

So, for every $w \in uR^+$, $\mathcal{E}_w(w) < 1/2$. So, for every $w \in uR^+$, there's some world $x \neq w$ such that wRx and $\mathcal{E}_w(x) = \mathcal{E}_x(x) < 1/2$, by lemma 11. So there must be at least three worlds in $wR \equiv \{x \mid wRx\}$—w itself and at least two other worlds. And since no world in wR sees w besides w itself (otherwise, lemma 8 tells us that uR^+ would be an S5 cluster), lemma 11 tells us that every world in wR gives itself precisely the same credence, so $\mathcal{E}_w(w)$ must be $1/N$, where N is the number of worlds in wR. Moreover, by lemma 2, every world in wR gives every world in wR that it sees a credence of $1/N$.

There now must be a unique world in wR—call it 'x_1'—such that $x_1 \neq w$ and x_1 sees every world in $A_1 \equiv wR \setminus \{w\}$. There must be one such world, otherwise no world in A_1 would give a credence of $\mathcal{E}_w(A_1) = (N-1)/N$ to A_1, and $C(A_1 \mid \langle \mathcal{E}(A_1) = (N-1)/N \rangle)$ would be defined but equal to 0, not $(N-1)/N$, so C wouldn't defer to $\mathcal{E}$ locally. Moreover, this world must be unique. For if there were two worlds in A_1 which saw every world in A_1, then they would see each other, and uR^+ would be an S5 cluster by lemma 8. So x_1 is the unique world in A_1 which sees every world in A_1.

Now, let $RA_1 \equiv \{z \notin wR \mid \forall x \in A_1 \; zRx\}$ be the set of all worlds besides w and x_1 which see every world in A_1. Then, every $z \in RA_1$ gives a credence of $1/N$ to each world in A_1, by lemma 2. So, for every $z \in RA_1$, $\mathcal{E}_z(A_1) = (N-1)/N$. w and x_1 both give a credence of $(N-1)/N$ to A_1. And no other worlds not in A_1 give a credence of $(N-1)/N$ to A_1, since all of those worlds see strictly fewer than $N-1$ of the worlds in A_1, and so by lemma 2 give a credence of less than $(N-1)/N$ to A_1. So $\langle \mathcal{E}(A_1) = (N-1)/N \rangle = RA_1 \cup \{w, x_1\}$. Since C defers to $\mathcal{E}$ locally,

$$C(A_1 \mid \langle \mathcal{E}(A_1) = (N-1)/N \rangle) = \frac{C(x_1)}{C(w) + C(x_1) + C(RA_1)} = \frac{N-1}{N}$$

So

$$C(x_1) = (N-1) \cdot C(w) + (N-1) \cdot C(RA_1)$$

Since $N > 2$, $C(x_1) > C(w)$.

All of the above reasoning iterates. Turn to x_1, and let $x_1 R \equiv \{z \mid x_1 R z\}$ be the set of worlds which x_1 sees. No world in $x_1 R$ can see x_1 besides x_1 itself (else, lemma 8 tells us that we'd have an S5 cluster). Since $\mathcal{E}_{x_1}(x_1) = 1/N$, lemma 10 tells us that every other world in $x_1 R$ gives itself the credence $1/N$, and so lemma 2 tells us that every world in $x_1 R$ gives every world in $x_1 R$ that it sees a credence of $1/N$.

Now, there must be a unique world—call it 'x_2'—such that $x_2 \neq x_1$, $x_2 \in x_1 R$, and x_2 sees every world in $A_2 \equiv x_1 R \backslash \{x_1\}$. There must be one such world, else no world in A_2 would give a credence of $\mathcal{E}_{x_1}(A_2) = (N-1)/N$ to A_2, and $C(A_2 \mid \langle \mathcal{E}(A_2) = (N-1)/N \rangle)$ would be defined but equal to 0, not $(N-1)/N$, and C wouldn't defer to $\mathcal{E}$ locally. Moreover, this world must be unique, else there would be two worlds in A_2 which saw every world in A_2, so they would see each other, and by lemma 8, uR^+ would be an S5 cluster. So x_2 is the unique world in A_2 which sees every world in A_2.

As before, $\langle \mathcal{E}(A_2) = (N-1)/N \rangle = RA_2 \cup \{x_1, x_2\}$, where $RA_2 \equiv \{z \notin x_1 R \mid \forall x \in A_2 \; zRx\}$. And, since C defers to $\mathcal{E}$ locally,

$$C(A_2 \mid \langle \mathcal{E}(A_2) = (N-1)/N \rangle) = \frac{C(x_2)}{C(x_1) + C(x_2) + C(RA_2)} = \frac{N-1}{N}$$

So

$$C(x_2) = (N-1) \cdot C(x_1) + (N-1) \cdot C(RA_2)$$

Since $N > 2$, $C(x_2) > C(x_1)$.

Proceeding in this way generates an infinite sequence of worlds, $w, x_1, x_2, \ldots$ such that

$$C(w) < C(x_1) < C(x_2) < \ldots$$

Since $C(w) > 0$, $C(w) > 1/M$ for some M. Then, $C(w \vee x_1 \vee \cdots \vee x_{M-1}) = C(w) + C(x_1) + \cdots + C(x_{M-1}) > M \cdot C(w) > 1$. So C isn't a probability. Contradiction.

So our assumption that $\mathcal{E}$ is not half-cyclic and C defers to $\mathcal{E}$ locally without deferring to $\mathcal{E}$ globally has led to a contradiction. So, if $\mathcal{E}$ is not half-cyclic, C defers to $\mathcal{E}$ locally iff C defers to $\mathcal{E}$ globally. □

REFERENCES

Arntzenius, Frank. 2003. "Some Problems for Conditionalization and Reflection." In *The Journal of Philosophy*, **100** (7): 356–370. [1]

Briggs, R. A. 2009. "Distorted Reflection." In *The Philosophical Review*, **118** (1): 59–85. [1]

Christensen, David. 2010. "Rational Reflection." In *Philosophical Perspectives*, **24**: 121–140. [1], [2]

Dorst, Kevin. 2020. "Evidence: A Guide for the Uncertain." In *Philosophy and Phenomenological Research*, **100** (3): 586–632. [1]

Dorst, Kevin, Levinstein, Benjamin A., Salow, Bernhard, Husic, Brooke E., & Fitelson, Branden. 2021. "Deference Done Better." In *Philosophical Perspectives*, **35** (1): 99–150. [1]

Elga, Adam. 2013. "The puzzle of the unmarked clock and the new rational reflection principle." In *Philosophical Studies*, **164**: 127–139. [2]

Gaifman, Haim. 1988. "A Theory of Higher Order Probabilities." In *Causation, Chance, and Credence: Proceedings of the Irvine Conference on Probability and Causation*, edited by Brian Skyrms & William L. Harper, Dordrecht: Kluwer Academic Publishers, volume 1, 191–220. [1], [2], [3], [4]

Hall, Ned. 1994. "Correcting the Guide to Objective Chance." In *Mind*, **103** (412): 505–517. [1]

Hall, Ned. 2004. "Two Mistakes About Credence and Chance." In *Australasian Journal of Philosophy*, **82** (1): 93–111. [1]

Hall, Ned & Arntzenius, Frank. 2003. "On What We Know About Chance." In *The British Journal for the Philosophy of Science*, **54** (2): 171–179. [1]

Ismael, Jenann. 2008. "Raid! Dissolving the Big, Bad Bug." In *Noûs*, **42** (2): 292–307. [1]

Ismael, Jenann. 2015. "In Defense of IP: A Response to Pettigrew." In *Noûs*, **49** (1): 197–200. [1]

Lasonen-Aarnio, Maria. 2015. "New Rational Reflection and Internalism about Rationality." In *Oxford Studies in Epistemology*, edited by Tamar Szabó Gendler & John Hawthorne, Oxford: Oxford University Press, volume 5, chapter 5. [1], [2]

Levinstein, Ben. forthcoming. "Accuracy, Deference, and Chance." In *The Philosophical Review*. [1]

Lewis, David K. 1980. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*, edited by Richard C. Jeffrey, Berkeley: University of California Press, volume II, 263–293. [1], [7]

Lewis, David K. 1994. "Humean Supervenience Debugged." In *Mind*, **103** (412): 473–490. [1]

Pettigrew, Richard. 2012. "Accuracy, Chance, and the Principal Principle." In *The Philosophical Review*, **121** (2): 241–275. [1]

Stalnaker, Robert C. 2019. "Rational Reflection and the Notorious Unmarked Clock." In *Knowledge and Conditionals: Essays on the Structure of Inquiry*, Oxford: Oxford University Press. [1], [2] Talbott, W. J. 1991. "Two Principles of Bayesian Epistemology." In *Philosophical Studies*, **62**: 135–150. [1]

Thau, Michael. 1994. "Undermining and Admissibility." In *Mind*, **103** (412): 491–503. [1]