Agreement and Updating For Self-Locating Belief

Michael Caie

Suppose we have two ideally rational, perfectly introspective, individuals, who want to pool their information. Since both individuals are perfectly introspective, they are both in a position to characterize to the other exactly what they know. Suppose, then, that each tells the other what they know and each correspondingly incorporates this information into what they know. Given that such individuals are perfectly rational, we should expect that, having communicated in this manner, their resulting epistemic states should be such that any further communication would be otiose. Each will know exactly what the other knows, and will know that each knows exactly what the other knows and so on.[1]

Let us call an individual at a time an **agent**. And let us call a group of agents, whose epistemic states are of the sort that might result from this type of idealized information pooling, **epistemic confidants**.

In what follows, we will focus on agents whose epistemic and credal states are defined over *finite* algebras of *intensional* propositions. We'll also assume that any agent is such that there is a strongest proposition ϕ that they know, and that they know all and only those propositions entailed by ϕ. Obviously not every possible agent need satisfy these assumptions. However, even restricting our attention to such agents, the arguments that follow should be of interest. Moreover, we should, I think, expect similar considerations to apply to agents who don't satisfy these constraints. Restricting our attention in this manner, though, will help to simplify and streamline our discussion.

In what follows, I'll argue that a very plausible principle about how epistemic confidants may rationally update their credal states has some surprising consequences for how rational agents may update their credal states in light of self-locating information.

Let us call a proposition that may only vary its truth-value across worlds **uncentered**, and a proposition that may its truth-value, not only across worlds, but also across individuals or times **centered**.

We can represent an agent whose epistemic and credal states are defined over an algebra that includes only uncentered propositions as having epistemic and credal states that are

---

[1] A caveat. Strictly speaking this may be too quick. Lederman [2016] provides a convincing argument that introspection and ideal rationality are not sufficient to ensure that certain facts will be common knowledge amongst a group of agents given that such agents have communicated. (See §2 for a characterization of the notion of common knowledge.) Assuming that Lederman's arguments are sound (and I'm inclined to think they are), then *strictly speaking* we require certain additional idealizing assumptions to ensure that such common knowledge is achieved via communication. For present purposes, though, we can ignore this potential complication. For, given such additional assumptions, all of the arguments that follow may proceed mutatis mutandis.

defined over $\mathcal{P}(\mathbf{W})$, for some set of possible worlds $\mathbf{W}$.[2] In particular, we can take their epistemic state to be some $K \in \mathcal{P}(\mathbf{W})$ and their credal state to be a probability measure over $\mathcal{P}(\mathbf{W})$.

In order, however, to represent agents who have epistemic and credal states defined over algebras that include both centered and uncentered propositions, we need a more fine-grained space than a set of possible worlds. Let $\mathbf{W}$, then, be a finite set of worlds, and let $\mathbf{A}$ be a finite set of agents.[3] We call $\mathbf{C} \subseteq \mathbf{W} \times \mathbf{A}$ a set of **centered-worlds**. (As a useful convention, for any $\vec{z} \in \mathbf{C}$, we will let w_z be the world parameter of $\vec{z}$ and a_z the agent parameter of $\vec{z}$. Thus $\vec{z} = \langle w_z, a_z \rangle$.) Given an algebra $\mathcal{P}(\mathbf{C})$, then, we can characterize uncentered and centered propositions as follows:

> **Def.** We say that $\phi \in \mathcal{P}(\mathbf{C})$ is an **uncentered proposition** just in case for any $\vec{z}, \vec{q} \in \mathbf{C}$, such that $w_z = w_q$, $\vec{z} \in \phi$ just in case $\vec{q} \in \phi$.

> **Def.** We say that $\phi \in \mathcal{P}(\mathbf{C})$ is a **centered proposition** just in case it is not uncentered.

We can represent an agent whose epistemic and credal states are defined over an algebra that includes both centered and uncentered propositions, then, as having epistemic and credal states defined over $\mathcal{P}(\mathbf{C})$, for some set of centered worlds $\mathbf{C}$. In particular, we can take their epistemic state to be some $K \in \mathcal{P}(\mathbf{C})$ and their credal state to be a probability measure over $\mathcal{P}(\mathbf{C})$.

Let $\mathbf{X}$, then, be either a set of possible worlds or a set of centered worlds. Let $\mathcal{S}$ be the set of functions mapping elements in $[0, n]$, for variable $n \in \mathbb{R}$, to subsets of $\mathbf{X}$. We can think of each member of $\mathcal{S}$ as representing the evolution of some possible individual's epistemic states over time. Let $\mathcal{C}$ be the set of functions mapping elements of $[0, n]$, for variable $n \in \mathbb{R}$, to functions of the form $f : \mathcal{P}(\mathbf{X}) \to \mathbb{R}$. We can think of the value, for each $x \in [0, n]$, relative to an appropriate member of $\mathcal{C}$, as corresponding to a possible credal state defined over the algebra $\mathcal{P}(\mathbf{X})$. We can, therefore, think of each member of $\mathcal{C}$ as representing the evolution of some possible individual's credal states over time.

> **Def.** We say that an **update function** is a function $\mathbf{u} : \mathcal{S} \to \mathcal{C}$ such that if, for some $n \in \mathbb{R}$, the domain of $s \in \mathcal{S}$ is $[0, n]$, then the domain of $\mathbf{u}(s)$ is also $[0, n]$.

Consider, then, the following principle:

> **Permissible Agreement:** For any algebra of propositions $\mathcal{P}(\mathbf{X})$, there is some rational update function that is such that, necessarily, any set of epistemic confidants who update in line with this function will assign the same credence to any uncentered proposition.

---

[2]We need not assume that $\mathbf{W}$ is a set of maximally specific possible worlds. Instead, we can think of each member of $\mathbf{W}$ as corresponding to a class of such possibilities. The members of such classes will disagree about matters about which our agents are unable to make distinctions.

[3]One can think of the members of $\mathbf{A}$ as ordered pairs of non-time-bound individuals and times. To keep notational clutter to a minimum, however, we'll suppress this additional structure and work with a single parameter.

Permissible Agreement tells us, roughly, that there is at least some rational way of setting one's credences over time that will guarantee agreement about worldly matters amongst agents who have pooled their information in a certain idealized manner. This is, I take it, a very plausible principle. Nonetheless, in what follows, I'll argue that this principle is incompatible with some prima facie plausible principles about how agents should update their credences given self-locating, i.e., centered, information.[4]

The paper will proceed as follows.

In §1, I'll begin by providing a brief defense of Permissible Agreement.

In §2, I'll offer a precisification, restricted to agents whose credal states are defined only over uncentered propositions, of the intuitive idea of a group of epistemic confidants. I'll show why, given this precisification and some plausible auxilliary principles, Permissible Agreement holds.

In §3, I'll provide a precisification of the intuitive idea of a group of epistemic confidants that applies to agents whose credal states are defined over both centered and uncentered propositions. This precisification will be codified in a principle called Centered Confidants.

In §4, I'll draw out some surprising consequences of Permissible Agreement and Centered Confidants. First, I'll show that Centered Confidants and Permissible Agreement are jointly incompatible with an updating procedure that may be motivated by a prima facie plausible principle of indifference. Second, I'll show that Centered Confidants and Permissible Agreement are jointly incompatible with both the standard thirder and halfer solutions to the famous Sleeping Beauty problem.

In §5, I'll consider an argument against Centered Confidants. According to this argument, there are certain requirements, in addition to those imposed by Centered Confidants, that a group of agents must satisfy in order to be epistemic confidants. Given these additional constraints, we can block both the argument that Permissible Agreement is incompatible with a plausible principle of indifference, and the argument that Permissible Agreement is incompatible with the standard thirder solution to the Sleeping Beauty problem. I'll argue, however, that if one is to reject Centered Confidants in this manner, then one must impose some surprising and, I think, implausible restrictions on cases in which agents may, by communicating, transmit to one another evidential support for uncentered propositions. Seeing why this is so will provide further support for the adequacy of the precisification provided in §3.

Finally, in §6, I'll consider an argument against Permissible Agreement. I'll argue that there may be principled grounds for rejecting Permissible Agreement, if one rejects a plausible evidentialist constraint on rational updating functions in favor of a certain epistemic consequentialist constraint. Now, ultimately, I don't think that we have good grounds for rejecting Permissible Agreement. For I think that the relevant evidentialist constraint should

---

[4]One, though not the only, way of thinking about the argument that follows is as showing that some prima facie plausible principles concerning how an agent should update their credences given self-locating information has the implausible consequence that there may be evidence for uncentered propositions that is, in principle, incommunicable. For discussion related to the phenomenon of apparently incommunicable information in the context of self-locating beliefs see, for example, Leslie [1997], Bostrom [2000], Bostrom [2002], and Lewis [2004].

be preferred to the epistemic consequentialist constraint. However, the conflict between these two constraints that arises in the sorts of cases that we'll be considering is, I think, of some independent interest.

1 Agreement Amongst Epistemic Confidants

So far, the notion of a group of epistemic confidants has been left fairly imprecise. Nonetheless, even given our initial rough-and-ready characterization of what it is for a group to be epistemic confidants, **Permissible Agreement** is still, I think, extremely plausible.

To see this, first note that to reject **Permissible Agreement** is to claim that it is impossible for a group of agents to rationally settle on an updating policy that would allow them to take a shared view about worldly matters given the sort of full exchange of information that is characteristic of epistemic confidants. At least prima facie, though, it would seem to be a fairly plausible constraint on rationality that, in principle, it is possible for agents to guarantee that when they have fully exchanged information, then they may rationally share a view about worldly, non-self-locating, matters.

We can provide further support for this claim as follows. According to our initial characterization, a group of agents are epistemic confidants just in case their respective epistemic states are of the sort that would result from an idealized information pooling amongst perfectly introspective and rational agents. In such a case, the epistemic states of the members of the group are such that there is nothing informative that one member could tell any other member. In this sense, their epistemic states are, as it were, open to view to one another.

In such a case, there would seem to be a natural sense in which such agents have the same information about the world. But, given this, there would seem to be a certain evidential symmetry between the epistemic states of epistemic confidants. In particular, the following principle would seem to be quite plausible:

> **Confidant Symmetry:** Necessarily, if a_1 and a_2 are epistemic confidants, and K_{a_1} is the proposition that characterizes a_1's epistemic state, and K_{a_2} the proposition that characterizes a_2's epistemic state, then, for any uncentered proposition ϕ, the evidential support for ϕ that K_{a_1} provides a_1 is the same as the evidential support for ϕ that K_{a_2} provides a_2.

For, given that a_1 and a_2 are epistemic confidants, their epistemic states will be open to one another's view, so not only will a_1 and a_2 know what they know, but a_2 will also know what a_1 knows, and a_1 will know what a_2 knows. And a_1 will know that a_2 knows what a_1 knows, while a_2 will know that a_1 knows what a_2 knows. And so on. But given that a_1's epistemic state is open to view to a_2, in this manner, it would seem that, if ϕ is an uncentered proposition, then whatever evidential support for ϕ that K_{a_1} provides a_1, there should be mirroring support for ϕ that K_{a_2} provides a_2. And, since a_2's epistemic state is open to view to a_1, in this manner, it would seem that, if ϕ is an uncentered proposition, then whatever evidential support for ϕ that K_{a_2} provides a_2, there should be mirroring support for ϕ that K_{a_1} provides a_1. And so, it would seem that, if ϕ is an uncentered proposition, then the

evidential support for ϕ that K_{a_1} provides a_1 is the same as the evidential support for ϕ that K_{a_2} provides a_2.

More generally, we can see Confidant Symmetry as following from the plausible thought that, by communicating, agents may transmit to one another the evidential support that they have for uncentered propositions. Given that evidence is transmissible, in this manner, it follows that the epistemic states of any two epistemic confidants will provide the same support for any uncentered proposition. For the epistemic states of epistemic confidants are exactly the sorts of epistemic states that result when agents have communicated to one another *all* of their evidence concerning uncentered propositions.

Next, consider the following principle:

> **Weak Evidentialism:** There is some rational update function $\mathbf{u}$ such that, necessarily, for any agents a_1 and a_2, if K_{a_1} is a_1's epistemic state and K_{a_2} is a_2's epistemic state, and K_{a_1} provides a_1 with the same evidential support for a proposition ϕ as K_{a_2} provides a_2, then if a_1 and a_2 set their credences in line with $\mathbf{u}$, a_1 and a_2 will assign the same credence to ϕ.

This principle tells us, roughly, that there is at least one rational updating function that need not, in some possible case, require agents to assign distinct credences to some proposition for which they have the same evidential support. At least prima facie, this would seem to be quite plausible.

Together Confidant Symmetry and Weak Evidentialism entail Permissible Agreement. Since both Confidant Symmetry and Weak Evidentialism are quite plausible, we have good reason, then, to accept Permissible Agreement.

2   Epistemic Confidants and Uncentered Algebras

There is a natural way of precisifying the notion of a group of epistemic confidants for agents whose epistemic and credal states are defined over an algebra of uncentered propositions.

To this end, let us first introduce two epistemic properties. The first property is that of **common knowledge**.[5] We can characterize this inductively as follows:

> **Def.** We say that the members of $G \subseteq \mathbf{A}$ **mutually know$_0$** a proposition ϕ just in case, for every $i \in G$, i knows ϕ.

> **Def.** We say that the members of $G \subseteq \mathbf{A}$ **mutually know$_{n+1}$** a proposition ϕ just in case the members of G mutually know$_0$ that they mutually know$_n$ ϕ.

> **Def.** We say that the members of $G \subseteq \mathbf{A}$ **commonly know** a proposition ϕ just in case, for every n, the members of G **mutually know$_n$** ϕ.

The second property is that of being **third-personally introspective**. We can characterize this as follows:

---

[5]See, for example, Lewis [1969], Aumann [1976], and Barwise [1988] for discussion and formal characterizations of common knowledge.

Def. We say that an agent i is **third-personally introspective** just in case, for any proposition ϕ, (i) if i knows ϕ, then i knows that i knows ϕ, and (ii) if i does not know ϕ, then i knows that i does not know that ϕ.

For agents whose epistemic and credal states are defined over an algebra of uncentered propositions, I suggest that the following provides a natural precisifcation of the intuitive idea of a group of agents being epistemic confidants:

Uncentered Confidants: A group of agents G whose epistemic states are defined over a set of possible worlds propositions are **epistemic confidants** just in case it is common knowledge amongst G what each member of G's epistemic state is, and it is common knowledge amongst G that each member of G is third-personally introspective.

To see this, first note that being third-personally introspective is the strongest introspective property that can be instantiated by an agent whose epistemic and credal states are defined over an algebra of uncentered propositions.[6]

Second, note that if a group G of perfectly introspective and rational agents fully exchange information, then, not only will each $i \in G$ know what each $i \in G$ knows, but, in virtue of the act of exchanging information, each $i \in G$ will know that each $i \in G$ knows what each $i \in G$ knows, and each $i \in G$ will know that each $i \in G$ knows that each $i \in G$ knows what each $i \in G$ knows, and so on. Thus, if a group G of perfectly introspective and rational agents fully exchange information, then it will be common knowledge amongst the members of G what each member of G's epistemic state is. Moreover, this is the strongest sense in which the epistemic states of a group of agents whose epistemic and credal states are defined over an algebra of uncentered propositions may be open to one another's view.

And, finally, note that if it is common knowledge amongst a group of agents what each of their epistemic states is, and each of the agents is third-personally introspective, then it will be common knowledge amongst the group that each agent is third-personally introspective.

We can now show that, given this precisification and some plausible auxillary assumptions, Permissible Agreement will hold, at last as restricted to agents whose epistemic and credal states are defined over algebras of uncentered propositions.

To this end, let us introduce the notion of an **agreement frame**. Let $\mathbf{A}$, then, be a finite set of agents, and $\mathbf{W}$ a finite set of possible worlds. And let $\star$ be an arbitrary object that we stipulate to be distinct from all other objects that appear in our frames. We let $(K_i)_{i \in \mathbf{A}}$ be an indexed family of functions such that $K_i : \mathbf{W} \to \mathcal{P}(\mathbf{W}) \cup \{\star\}$. We assume that, for each $w \in \mathbf{W}$, $K_i(w) \neq \emptyset$. If i exists at w, then $K_i(w) \in \mathcal{P}(\mathbf{W})$. In this case, $K_i(w)$ represents the strongest proposition that i knows at a world w. We'll call this set i's **epistemic state** at w. If i doesn't exist at w, then we let $K_i(w) = \star$. As noted earlier, we'll assume that any agent we are modeling knows a proposition ϕ at a world w just in case $K_i(w) \subseteq \phi$. Note that since, for each proposition ϕ, $\star \not\subseteq \phi$, we have, as desired, that if an agent doesn't exist at a world w, then they do not know any proposition ϕ at w.

---

[6]As we'll see, this isn't true for agents whose epistemic and credal states are defined over algebras that include centered propositions.

Def. We say that $K_i(\cdot)$ is **reflexive** at w just in case $w \in K_i(w)$.

Epistemic states are, of course, subject to the following constraint:

Factivity: If an agent i knows a proposition ϕ, then i is not mistaken in virtue of believing ϕ.[7]

To ensure that the family of functions $(K_i)_{i\in\mathbf{A}}$ satisfy **Factivity**, then, we stipulate that each $K_i(\cdot)$ is reflexive relative to each $w \in W$.

Def. We say that $\mathcal{F} = \langle \mathbf{W}, \mathbf{A}, (K_i)_{i\in\mathbf{A}} \rangle$ is an **agreement frame**.

An agreement frame represents various possible epistemic states of a set of agents $\mathbf{A}$ defined over the algebra $\mathcal{P}(\mathbf{W})$. Note that since $(K_i)_{i\in\mathbf{A}}$ encodes information about each member of $\mathbf{A}$'s epistemic state at each $w \in \mathbf{W}$, agents whose epistemic states are defined over an algebra $\mathcal{P}(\mathbf{W})$ of an agreement frame thereby have epistemic states that at least partially concern the epistemic states of the members of $\mathbf{A}$. And so, since the agents represented in an agreement frame, $\mathbf{A}$, have epistemic states defined over the algebra of this frame, $\mathcal{P}(\mathbf{W})$, it follows that their epistemic states at least partially concern one another's epistemic states.

We can represent the properties of common knowledge and third-personal introspection in an agreement frame as follows.

Def. Let $\mathsf{M}_G^0(\phi)$ be the proposition that is true at w just in case, for every $i \in G$, $K_i(w) \subseteq \phi$.

Def. Let $\mathsf{M}_G^{n+1}(\phi)$ be the proposition that is true at w just in case $\mathsf{M}_G^0(\mathsf{M}^n(\phi))$ is true at w.

Def. Let $\mathsf{Com}_G(\phi)$ be the proposition that is true at w just in case $\mathsf{M}_G^n(\phi)$ is true at w, for all n.

Def. Let $\mathsf{Hom}[K_i(\cdot)]$ be the proposition that is true at some world w just in case $K_i(w) \in \mathcal{P}(\mathbf{W})$ and, for every $w' \in K_i(w)$, $K_i(w) = K_i(w')$.

Then we can show that (i) ϕ is common knowledge amongst G at w just in case $\mathsf{Com}_G(\phi)$ is true at w and (ii) an agent $i \in \mathbf{A}$ is third-personally introspective at a world w just in case $\mathsf{Hom}[K_i(\cdot)]$ holds at w.[8]

Now let $[K_i = \phi]$ be the proposition that is true at a world w' just in case $K_i(w') = \phi$. Given **Uncentered Confidants**, we have:

---

[7]In the case in which ϕ is an uncentered proposition, we can take **Factivity** to say: If an agent i knows ϕ, then ϕ is true. However, if ϕ is a centered proposition, we can't say that such a proposition is true or false simpliciter. Instead, such a proposition will be true relative to some agents and false relative to others. Our formulation, then, is meant to apply when ϕ is an uncentered proposition and when ϕ is a centered proposition. Where ϕ is a centered proposition, i will be mistaken in virtue of believing ϕ just in case ϕ is false relative to i.

[8]See Caie [2015] for the justification of these claims.

Uncentered Representation: A group of agents $G \subseteq \mathbf{A}$ whose epistemic states are defined over a set of possible worlds propositions are epistemic confidants at a world w just in case $\mathsf{Com}_G([K_i = K_i(w)])$ and $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$ hold at w, for each $i \in G$.[9]

In addition, we have the following theorem:

Uncentered Agreement Theorem: Let $\mathcal{F}$ be an agreement frame and $Pr(\cdot)$ a probability function on $\mathcal{P}(\mathbf{W})$. If, at some $w \in \mathbf{W}$, for each $i \in G \subseteq \mathbf{A}$, $\mathsf{Com}_G([K_i = K_i(w)])$ and $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$ hold, then, for each $i, j \in G$, $Pr(\cdot|K_i(w)) = Pr(\cdot|K_j(w))$.[10]

The key fact underlying this theorem is that if $\mathsf{Com}_G([K_i = K_i(w)])$ and $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$ are true at w, for each $i \in G \subseteq \mathbf{A}$, then it follows that, for each $i, j \in G$, $K_i(w) = K_j(w)$. Given this, the result is obvious.

Now, amongst the class of update functions, there will be some that correspond, in a natural way, to probability functions.

Def. We say that an update function $\mathbf{pr}(\cdot)$ is **determined by a probability function** $Pr(\cdot)$ just in case, for every $s \in \mathcal{S}$, if the domain of s is $[0, n]$, then $\mathbf{pr}(s)(x) = Pr(\cdot|s(x))$, for each $x \in [0, n]$.

It is natural to think that amongst the class of rational update functions for agents with credences defined over some algebra $\mathcal{P}(\mathbf{W})$, at least some of these functions will be determined by a probability function over $\mathcal{P}(\mathbf{W})$. Thus:

Weak Uncentered Bayesianism: Given some algebra $\mathcal{P}(\mathbf{W})$, there is, amongst the class of rational updating functions defined over this algebra, some function that is determined by a probability function.

We're now in a position to see why, given Uncentered Confidants, Permissible Agreement will hold as restricted to agents whose credences are defined only over algebras of uncentered propositions. For, given Uncentered Confidants, we have Uncentered Representation. And, given Uncentered Representation and Weak Uncentered Bayesianism, it follows, given the Uncentered Agreement Theorem, that Permissible Agreement holds as restricted to agents whose credences are defined only over algebras of uncentered propositions.

To see this, note that, given Uncentered Representation and the Uncentered Agreement Theorem, it follows that any epistemic confidants whose epistemic and credal states are

---

[9]Strictly speaking we could simplify this by saying that a group of agents $G \subseteq \mathbf{A}$, whose epistemic states are representable in an agreement frame, are epistemic confidants at a world w just in case $\mathsf{Com}_G([K_i = K_i(w)])$ holds at w, for each $i \in G$. For given that $\mathsf{Com}_G([K_i = K_i(w)])$ holds at w, for each $i \in G$, it follows that $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$ hold at w, for each $i \in G$.

[10]For the proof of this result, see the Second Agreement Theorem in Caie [2015].

defined over an algebra of uncentered propositions and who set their credences by conditionalizing their respective epistemic states on the same probability function will assign the same credence to every proposition over which their credal states are defined. And, given **Weak Uncentered Bayesianism**, we have that there is some rational updating function that is determined by probability function. And so we have that there is some rational updating function that guarantees that epistemic confidants whose epistemic and credal states are defined over an algebra of uncentered propositions will assign the same credence to every proposition over which their credal states are defined.

3   Epistemic Confidants and Centered Algebras

In this section, I'll suggest a natural precisification of the notion of a group of epistemic confidants that applies to agents who have epistemic and credal states defined over algebras that include both centered and uncentered propositions.

To this end, let us introduce, in addition to the epistemic properties of common knowledge and third-personal introspection, the property of being **first-personally introspective**. We can characterize this as follows:

> **Def.** We say that an agent i is **first-personally introspective** just in case, for any proposition ϕ, (i) if i knows ϕ, then i knows the centered proposition that they know ϕ and (ii) if i does not know ϕ, then i knows the centered proposition that they do not know ϕ.

For agents whose epistemic and credal states are defined over an algebra that includes both centered and uncentered propositions, I suggest that the following provides a natural precisifcation of the intuitive idea of a group of agents being epistemic confidants:

> **Centered Confidants:** A group of agents G whose epistemic states are defined over a set of centered worlds propositions are **epistemic confidants** just in case it is common knowledge amongst G what each member of G's epistemic state is, and it is common knowledge amongst G that each member of G is both first-personally and third-personally introspective.

To see this, first note that being both first-personally and third-personally introspective is the strongest introspective property that can be instantiated by an agent whose epistemic and credal states are defined over an algebra that includes both centered and uncentered propositions.

Second, note that if a group G of perfectly introspective and rational agents fully exchange information, then, not only will each $i \in G$ know what each $i \in G$ knows, but, in virtue of the act of exchanging information, each $i \in G$ will know that each $i \in G$ knows what each $i \in G$ knows, and each $i \in G$ will know that each $i \in G$ knows that each $i \in G$ knows what each $i \in G$ knows, and so on. Thus, if a group G of perfectly introspective and rational agents fully exchange information, then it will be common knowledge amongst the members of G what each member of G's epistemic state is.

Moreover, I claim that this is the strongest sense in which the epistemic states of a group of agents whose credences are defined over an algebra of both centered and uncentered propositions may be open to one another's view that is common amongst cases of idealized information pooling. A full defense of this claim will, however, have to wait until §5.

Finally, note that if it is common knowledge amongst a group of agents what each of their epistemic states is, and each of the agents is both first-personally and third-personally introspective, then it will be common knowledge amongst the group that each agent is both first-personally and third-personally introspective.

We can now show that, given this precisification, Permissible Agreement is compatible with both of the following constraints on update functions, as well as their negations:[11]

> **Strong Centered Bayesianism:** Given an algebra $\mathcal{P}(\mathbf{C})$, every rational updating function defined over this algebra is determined by a probability function.

> **Strong Centered Anti-Bayesianism:** Given an algebra $\mathcal{P}(\mathbf{C})$, no rational updating function defined over this algebra is determined by a probability function.

To this end, let us introduce the notion of a **centered agreement frame**. Let $\mathbf{A}$, again, be a finite set of agents, and $\mathbf{W}$ a finite set of possible worlds. Let $\mathbf{C} \subseteq \mathbf{W} \times \mathbf{A}$ be a set of centered worlds.[12] And let $\star$ be an arbitrary object that we stipulate to be distinct from all other objects that appear in our frames.

We now let $(K_i)_{i \in \mathbf{A}}$ be an indexed family of functions such that $K_i : \mathbf{C} \to \mathcal{P}(\mathbf{C}) \cup \{\star\}$. We assume that, for every $\vec{z}, \vec{q} \in \mathbf{C}$, $K_i(\vec{z}) = K_i(\vec{q})$, if $w_z = w_q$. If i exists at w_q, then $K_i(\vec{q}) \in \mathcal{P}(\mathbf{C})$. In this case, $K_i(\vec{q})$ represents the strongest proposition that i knows at the world w_q. As before, we assume that any agent we are modeling knows all and only those propositions that are entailed by this proposition. If i doesn't exist at w_q, then we let $K_i(\vec{q}) = \star$.

> **Def.** We say that $K_i(\cdot)$ is **center-reflexive** at $\vec{q}$ just in case $\langle w_q, i \rangle \in K_i(\vec{q})$.

The key difference between an agreement frame and a centered agreement frame is the formal constraint that we need to impose on the family of functions $(K_i)_{i \in \mathbf{A}}$ in order to ensure that the states that they represent satisfy Factivity. In the case of agreement frames, the formal property that ensures that the states represented in such a frame satisfy Factivity

---

[11]Both Strong Centered Bayesianism and Strong Centered Anti-Bayesianism are controversial theses, at least when we are dealing with agents whose credal states are defined over algebras that include centered propositions. See, for example, Meacham [2008] for a view that endorses Strong Centered Anti-Bayesianism. For an endorsement of Strong Centered Bayesianism see Dorr and Arntzenius [2015]. I take it, then, that it's a mark in favor of Centered Confidants, considered as precisification of the notion of a group of epistemic confidants, that Permissible Agreement, so construed, is compatible with each of these claims as well as their negations.

[12]We'll assume that for each $a \in \mathbf{A}$, there is some $\vec{x} \in \mathbf{C}$ such that $a_x = a$. Thus $\mathbf{A} = \{a_x : \text{for some } \vec{x} \in \mathbf{C}\}$. Similarly, we assume that, for every $w \in \mathbf{W}$, there is some $\vec{x} \in \mathbf{C}$ such that $w_x = w$. We'll assume, further that, for each $w \in \mathbf{W}$ and each $a \in \mathbf{A}$, a exists in w just in case there is some $\vec{x} \in \mathbf{C}$ such that $w_x = w$ and $a_x = a$.

is reflexivity. In the case of centered agreement frames, however, the formal property that ensures that the states represented in such a frame satisfy Factivity is center-reflexivity.

To see why reflexivity is not required in order to ensure that Factivity holds for the states represented by the family of functions $(K_i)_{i \in \mathbf{A}}$ in a centered agreement frame, suppose we have an agent i who is omniscient at some world w. Such an agent, then, knows exactly which world they are in and, in addition, which agent they are. We have, then, that, for each $\vec{z} \in \mathbf{C}$, if $w_z = w$, then $K_i(\vec{z}) = \{\langle w_z, i \rangle\}$. Now consider some $\vec{q}$ such that $w_q = w$ and $a_q = j$, for some $j \neq i$. We have, then, that $\vec{q} \notin K_i(\vec{q})$. Thus, given that i is omniscient, it follows that $K_i(\cdot)$ is not reflexive at $\vec{q}$. However, surely the fact that i is omniscient at w_q does not entail that i is mistaken in virtue of some proposition that they believe. Thus, it is not necessary that $K_i(\cdot)$ be reflexive at some point in order for the state that it represents at the world of that point to satisfy Factivity.

Note, though, that, while $K_i(\cdot)$ fails to be reflexive at $\vec{q}$, it is center-reflexive. To see why center-reflexivity is the formal property corresponding to Factivity for the sorts of epistemic states represented in a centered agreement frame, note that we can think, more generally, of the *doxastic* state of the sort of agent that is represented in a centered agreement frame as being the smallest set of world-agent pairs such that the agent believes that amongst this set there is a pair whose world parameter is the actual world and whose agent parameter is them. If, then, we take some function $B_i(\vec{q})$ to represent the doxastic state, so construed, of an agent i at world w_q, it is clear that such a state will not be mistaken just in case $\langle w_q, i \rangle \in B_i(\vec{q})$. Thus, we can see that the family of functions $(K_i)_{i \in \mathbf{A}}$ representing the possible epistemic states of the agents in $\mathbf{A}$ in a centered agreement frame will satisfy Factivity just in case each such function is center-reflexive relative to each $\vec{q} \in \mathbf{C}$.

To ensure that the family of functions $(K_i)_{i \in \mathbf{A}}$ satisfy Factivity, then, we stipulate that each $K_i(\cdot)$ is center-reflexive relative to each $\vec{q} \in \mathbf{C}$.

Def. We say that $\mathcal{C} = \langle \mathbf{C}, \mathbf{A}, (K_i)_{i \in \mathbf{A}}, \rangle$ is a **centered agreement frame**.

A centered agreement frame represents various possible epistemic states of a set of agents $\mathbf{A}$ defined over the algebra $\mathcal{P}(\mathbf{C})$. Note that since $(K_i)_{i \in \mathbf{A}}$ encodes information about each member of $\mathbf{A}$'s epistemic state at each $\vec{z} \in \mathbf{C}$, agents whose epistemic states are defined over an algebra $\mathcal{P}(\mathbf{C})$ of a centered agreement frame thereby have epistemic states that at least partially concern the epistemic states of the members of $\mathbf{A}$. And so, since the agents represented in a centered agreement frame, $\mathbf{A}$, have epistemic states defined over the algebra of this frame, $\mathcal{P}(\mathbf{C})$, it follows that their epistemic states at least partially concern one another's epistemic states.

We can represent the properties of common knowledge and third-personal introspection in a centered agreement frame in much the same manner as we represented these properties in an agreement frame.

Def. Let $\mathsf{M}_G^0(\phi)$ be the proposition that is true at $\vec{q}$ just in case, for every $i \in G$, $K_i(\vec{q}) \subseteq \phi$.

Def. Let $\mathsf{M}_G^{n+1}(\phi)$ be the proposition that is true at $\vec{q}$ just in case $\mathsf{M}_G^0(\mathsf{M}^n(\phi))$ is true at $\vec{q}$.

Def. Let $\mathsf{Com}_G(\phi)$ be the proposition that is true at $\vec{q}$ just in case $\mathsf{M}_G^n(\phi)$ is true at $\vec{q}$, for all n.

Def. Let $\mathsf{Hom}[K_i(\cdot)]$ be the proposition that is true at some centered world $\vec{q}$ just in case $K_i(\vec{q}) \in \mathcal{P}(\mathbf{C})$ and, for every $\vec{z} \in K_i(\vec{q})$, $K_i(\vec{q}) = K_i(\vec{z})$.

Then we can show that (i) ϕ is common knowledge amongst G at w_z just in case $\mathsf{Com}_G(\phi)$ is true at $\vec{z}$ and (ii) an agent $i \in \mathbf{A}$ is third-personally introspective at w_z just in case $\mathsf{Hom}[K_i(\cdot)]$ holds at $\vec{z}$.[13]

In addition, we can also represent the property of being first-personally introspective in a centered agreement frame.

Def. Let $\mathsf{CHom}[K_i(\cdot)]$ be the proposition that is true at some centered world $\vec{q}$ just in case $K_i(\vec{q}) \in \mathcal{P}(\mathbf{C})$ and $K_i(\vec{q}) = K_{a_z}(\vec{z})$, for all $\vec{z} \in K_i(\vec{q})$.

Then we can show that an agent $i \in \mathbf{A}$ is first-personally introspective at w_z just in case $\mathsf{CHom}[K_i(\cdot)]$ holds at $\vec{z}$.[14]

Now let $[K_i = \phi]$ be the proposition that is true at a centered world $\vec{z}$ just in case $K_i(\vec{z}) = \phi$. Given **Centered Confidants**, we have:

Centered Representation: A group of agents $G \subseteq \mathbf{A}$ whose epistemic states are defined over a set of centered worlds propositions are epistemic confidants at a world w_z just in case $\mathsf{Com}_G([K_i = K_i(\vec{z})])$, $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$ and $\mathsf{Com}_G(\mathsf{CHom}[K_i(\cdot)])$ hold at $\vec{z}$, for each $i \in G$.

In addition, we have the following theorem:

Centered Agreement Theorem: Let $\mathcal{C}$ be a centered agreement frame. Then there exists some probability function $Pr(\cdot)$ defined over $\mathcal{P}(\mathbf{C})$, such that, for any $\vec{q} \in \mathbf{C}$, if every $i \in G \subseteq \mathbf{A}$ is such that the following propositions are true at $\vec{q}$: $\mathsf{Com}_G([K_i = K_i(\vec{q})])$, $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$, $\mathsf{Com}_G(\mathsf{CHom}[K_i(\cdot)])$, then, for each uncentered proposition ψ and each $i, j \in G$, $Pr(\psi|K_i(\vec{q})) = Pr(\psi|K_j(\vec{q}))$.[15]

Given **Centered Confidants**, then, we can see that **Permissible Agreement** is compatible **Strong Centered Bayesianism**. For, given **Centered Confidants**, we have **Centered Representation**. And, given **Centered Representation**, it follows from the **Centered Agreement Theorem** that, given an algebra $\mathcal{P}(\mathbf{C})$, there is some update function that is determined by a probability function over this algebra, that ensures that any epistemic confidants, with epistemic and credal states defined over $\mathcal{P}(\mathbf{C})$, who update in line with this function will assign the same credence to every uncentered proposition over which their credal states are defined.

In addition, we can also see that, given **Centered Confidants**, **Permissible Agreement** is compatible with **Strong Centered Anti-Bayesianism**. For, let $\mathbf{pr}(\cdot)$ be an update function that

---

[13]Again, see Caie [2015] for the justification of these claims.
[14]See Caie [2015].
[15]See the Second Centered Agreement Theorem in Caie [2015].

is determined by some probability function $Pr(\cdot)$ that ensures that any epistemic confidants, so construed, who update in line with $\mathbf{pr}(\cdot)$ will assign the same credence to any uncentered proposition. We can then let $\mathbf{u}(\cdot)$ be a function that is exactly like $\mathbf{pr}(\cdot)$ with respect to the epistemic states of epistemic confidants, but that differs arbitrarily for other such states. All such functions will ensure agreement amongst epistemic confidants, but many such functions will not be determined by any probability function.

And, of course, given that there are update functions that ensure agreement amongst epistemic confidants, as construed given **Centered Confidants**, that are determined by probability functions and ones that are not determined by any probability function, it follows that, given **Centered Confidants**, **Permissible Agreement** is compatible with both the rejection of **Strong Centered Bayesianism** and the rejection of **Strong Centered Anti-Bayesianism**.

4 Centered Updating

I've argued that the intuitive principle of **Permissible Agreement** is extremely plausible. I've further argued that **Centered Confidants** provides a plausible precisification of the notion of a group of epistemic confidants, for agents whose epistemic and credal states are defined over algebras that include both centered and uncentered propositions. We have, then, good prima facie grounds for endorsing **Permissible Agreement**, understood given the precisification of the notion of a group of epistemic confidants provided by **Centered Confidants**. This is further supported, I take it, by the fact that, given **Centered Confidants**, **Permissible Agreement** is compatible with both **Strong Centered Bayesianism** and **Strong Centered Anti-Bayesianism**.

I'll now show, however, that, given **Centered Confidants**, **Permissible Agreement** has some surprising consequences for how an agent may rationally update their credences in light of centered information. In §4.1, I'll show that, given **Centered Confidants**, **Permissible Agreement** is incompatible with a constraint on updating that may be motivated by a prima facie plausible principle of indifference. In §4.2, I'll argue that, given **Centered Confidants**, **Permissible Agreement** is incompatible with both the thirder and halfer solutions to the Sleeping Beauty problem.

4.1 Agreement and Center-Indifference

Consider the following class of probability functions:

> **Def.** We say that a probability function $Pr(\cdot)$ defined on $\mathcal{P}(\mathbf{C})$ is **center-indifferent** just in case, if $w_x = w_q$, then $Pr(\vec{x}) = Pr(\vec{q})$.

It is, I think, tempting to suppose that the class of rational update functions will be determined by a (perhaps proper) subclass of the class of center-indifferent probability functions. Thus:

> **Mandatory Center-Indifference:** For any finite algebra $\mathcal{P}(\mathbf{C})$, any rational update function $\mathbf{u}$ is such that, for every $s \in \mathcal{S}$, there is some class of center-indifferent probability functions $\mathbf{Pr}$ defined on $\mathcal{P}(\mathbf{C})$, such that if the domain of s is $[0, n]$, then, for each $x \in [0, n]$, $\mathbf{u}(s)(x) = Pr(\cdot | s(x))$, for some $Pr(\cdot) \in \mathbf{Pr}$.

One might, for example, motivate **Mandatory Center-Indifference** by appeal to the thought that an agent should always set their credences by conditionalizing on a probability function that is as unbiased as possible about which agent one is and which world one is in. In particular, then, an agent should always conditionalize on a probability function that doesn't make it any more likely that one is some agent i in a world w than that one is some agent j in a world v. But this requires that an agent always conditionalize on a center-indifferent probability function. In particular, it requires that an agent always conditionalize on the unique probability function such that $Pr(\vec{q}) = Pr(\vec{z})$, for each $\vec{q}, \vec{z} \in \mathbf{C}$. Or one might allow that an agent may update by conditionalizing on some probability function that gives greater probability to the world being one way rather than another, but one might nonetheless maintain that one should only update by conditionalizing on a probability function that doesn't give greater probability to one being one agent in some world w rather than some other agent in w. But, again, this requires that an agent always conditionalize on a center-indifferent probability function.

While **Mandatory Center-Indifference** is prima facie plausible, it is, however, incompatible with **Permissible Agreement**, given **Centered Confidants**. To see this, consider the following centered agreement frame:

> **Frame 1:** Let $\mathcal{C}_1$ be such that $\mathbf{W} = \{w_1, w_2\}$, $\mathbf{A} = \{a_1, a_2, a_3\}$, and $\mathbf{C} = \mathbf{W} \times \mathbf{A}$. We let $K_{a_1}(\cdot)$ be such that, if $w_z = w_1$ or $w_z = w_2$, then $K_{a_1}(\vec{z}) = \{\langle w_1, a_1 \rangle,$ $\langle w_2, a_1 \rangle, \langle w_2, a_2 \rangle\}$. We let $K_{a_2}(\cdot)$ be such that, if $w_z = w_1$, then $K_{a_2}(\vec{z}) = \{\langle w_1, a_2 \rangle\}$; if $w_z = w_2$, then $K_{a_2}(\vec{z}) = \{\langle w_1, a_1 \rangle, \langle w_2, a_1 \rangle, \langle w_2, a_2 \rangle\}$. And, finally, we let $K_{a_3}(\cdot)$ be such that, if $w_z = w_1$ or $w_z = w_2$, then $K_{a_3}(\vec{z}) = \{\langle w_1, a_3 \rangle, \langle w_2, a_3 \rangle\}$.

One can verify, somewhat tediously, that, letting $G = \{a_1, a_3\}$, $\mathsf{Com}_G([K_i = K_i(\vec{z})])$, $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$ and $\mathsf{Com}_G(\mathsf{CHom}[K_i(\cdot)])$ hold at each $\vec{z} \in \mathbf{C}$, for each $i \in G$. And so, given **Centered Confidants**, we have that a_1 and a_3 are epistemic confidants at both w_1 and w_2. We can visualize this frame as in Figure 1.
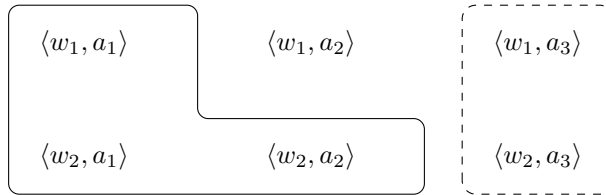

Figure 1: $\mathcal{C}_1$

Here the set of centered-worlds enclosed in the solid line corresponds to a_1's (invariant) epistemic state, while the set of centered-worlds enclosed in the dashed line corresponds to a_3's (invariant) epistemic state.

Now let $I^c(\cdot)$ be an arbitrary center-indifferent probability function defined over $\mathcal{P}(\mathbf{C})$, let $\phi = \{\langle w_2, a_1 \rangle, \langle w_2, a_2 \rangle, \langle w_2, a_3 \rangle\}$, and let K_{a_1} and K_{a_3} be a_1 and a_3's invariant epistemic

states within this frame. We can show that $I^c(\phi|K_{a_1}) \neq I^c(\phi|K_{a_3})$.[16] Since ϕ is an uncentered proposition and $I^c(\cdot)$ is an *arbitrary* center-indifferent probability function, it follows that conditionalizing on any center-indifferent probability function will lead a_1 and a_3 to assign different credences to some uncentered proposition. This shows that, given Mandatory Center-Indifference and Centered Confidants, there is no rational update function defined over $\mathcal{P}(\mathbf{C})$ that guarantees that epistemic confidants who set their credences by updating in line with that function will assign the same credence to each uncentered proposition. And so, it follows that Mandatory Center-Indifference is incompatible with Permissible Agreement and Centered Confidants.

4.2   Agreement and Sleeping Beauty

There are a number of vexing issues about how rational agents should update their credences in light of self-locating information. One puzzling case that has received a great deal of attention is the so-called Sleeping Beauty problem. In this section, I'll argue that two of the standard treatments of this puzzle are incompatible with the conjunction of Permissible Agreement and Centered Confidants.

> **Sleeping Beauty:** Sam is subject to the following experiment. On Sunday night he will be put to sleep and a fair coin will be flipped. If the coin lands heads, then he will be awoken once on Monday. If the coin lands tails, however, he will be awoken twice – once on Monday and once on Tuesday. Each of these awakenings will be indistinguishable. Shortly after each awakening he will be informed what day it is, and then he will be given a drug that erases his memory. Sam is informed of all of these facts.[17]

Let Heads be the uncentered proposition that the coin came up heads, and let Tails be the uncentered proposition that the coin came up tails. There are two questions that we'll want to consider. First, what credence should Sam have in Heads on Monday when he is woken up? Second, what credence should Sam have in Heads on Monday after being told that it is Monday?

Here are two ways of answering these questions that have some prima facie appeal.

> **Beauty Half:** On Monday, upon awakening, Sam should give Heads credence $1/2$, and, upon learning that it is Monday, he should give Heads credence $2/3$.[18]

---

[16]Let $I^c(\langle w_1, a_1 \rangle) = I^c(\langle w_1, a_2 \rangle) = I^c(\langle w_1, a_3 \rangle) = x$, $I^c(\langle w_2, a_1 \rangle) = I^c(\langle w_2, a_2 \rangle) = I^c(\langle w_2, a_3 \rangle) = y$. Then $I^c(\phi|K_{a_1}) = \frac{2y}{x+2y}$ and $I^c(\phi|K_{a_3}) = \frac{y}{x+y}$. But, in general, we have $\frac{2y}{x+2y} \neq \frac{y}{x+y}$ if $x, y \neq 0$. One way of seeing this is to note that if $\frac{2y}{x+y} = \frac{y}{x+y}$, then we have that $\frac{x}{x+2y} = \frac{x}{x+y}$. But given that $y \neq 0$, we have $(x+2y) \neq (x+y)$. But, in general, we have $\frac{m}{q} \neq \frac{m}{r}$ if $m, q, r \neq 0$ and $q \neq r$.

[17]This problem was introduced into the philosophical literature in Elga [2000]. Related puzzles appear earlier in Aumann et al. [1997] and Piccione and Rubenstein [1997]. The literature on this puzzle is now vast. For treatments of this and related puzzles and further references see, e.g., Elga [2000], Lewis [2001], Dorr [2002], Halpern [2004], Meacham [2008], Titelbaum [2008], Briggs [2010] and Moss [2012].

[18]For a defense of this position see Lewis [2001]. A related line of thought is also advocated in Lewis [2004].

Beauty Third: On Monday, upon awakening, Sam should give Heads credence 1/3, and, upon learning that it is Monday, he should give Heads credence 1/2.[19]

I'll now argue that if one endorses either Beauty Half or Beauty Third, then one is committed to rejecting either Permissible Agreement or Centered Confidants. The argument for this will proceed as follows. First, I'll introduce a case called Duplication. I'll argue that if one endorses either Beauty Half or Beauty Third, then one should endorse the corresponding positions for Duplication: Duplication Half, Duplication Third. I'll then introduce a variant on this case called Duplication With Confidants that involves two agents who, given Centered Confidants, are epistemic confidants. I'll argue that if one endorses either Duplication Half or Duplication Third, then one should endorse the corresponding positions for Duplication With Confidants: Confidant Half, Confidant Third. However, we can show that, given Centered Confidants, if one endorses either Confidant Half or Confidant Third, then one must reject Permissible Agreement. Since if one endorses either Beauty Half or Beauty Third, one should endorse either Confidant Half or Confidant Third, it follows that if one endorses either Beauty Half or Beauty Third, then one must reject either Centered Confidants or Permissible Agreement.

Duplication: Beth is subject to the following experiment. On Sunday night she will be put to sleep in a red room and a coin will be flipped. If the coin lands heads, then she will be awoken on Monday. First, her eyes will be closed. She will then open her eyes to see that she is in the red room. If, however, the coin lands tails, then a perfect duplicate of Beth, Beth$_d$, will be placed in a blue room. Both Beth and Beth$_d$ will be woken up on Monday. First, their eyes will be closed. They will then open their eyes and each will discover which room they are in, and so, who they are. Beth is informed of all of these facts.

Again, let Heads and Tails be the uncentered proposition that the coin came up heads or tails, respectively. Corresponding to Beauty Half and Beauty Third, we have the following two prima facie plausible accounts of how Beth should update her credences in Heads throughout this case:

Duplication Half: On Monday upon awakening but before opening her eyes Beth should give Heads credence 1/2, and upon opening her eyes and learning that she is Beth in the red room she should give Heads credence 2/3.

Duplication Third: On Monday upon awakening but before opening her eyes Beth should give Heads credence 1/3, and upon opening her eyes and learning that she is Beth in the red room she should give Heads credence 1/2.

I submit that, if one endorses Beauty Half then one should endorse Duplication Half, while if one endorses Beauty Third then one should endorse Duplication Third. To see why this is so, note that these two cases have the following isomorphic structure. First, an individual

---

[19] For a defense of this position see Elga [2000]. See also Dorr [2002].

16

is apprised that a fair coin will be flipped, and that if the coin lands heads then a future waking experience of theirs will be unique, while if the coin lands tails then that experience will be realized twice. Second, there is a waking event in which, given knowledge of the coin set-up, an agent loses track of which agent they are. In the case of Sleeping Beauty, Sam, upon awakening, loses track of whether he is Sam-on-Monday or Sam-on-Tuesday. In the case of Duplication, Beth, upon awakening, loses track of whether she is Beth or $Beth_d$. In each case, given knowledge of the coin set up, together with their waking experience, the agent is uncertain of who they are, both unconditionally and conditional on the proposition that the appropriate coin came up tails, but each is certain of who they are, conditional on the proposition that the appropriate coin came up heads. Second, there is an event in which an agent reorients themselves. In the case of Sleeping Beauty, Sam, upon being told which day it is, determines that he is Sam-on-Monday and not Sam-on-Tuesday. In the case of Duplication, Beth, upon opening her eyes, determines that she is Beth and not $Beth_d$.

Given these structural parallels, it seems to me that whatever epistemic bearing the events of Sam's losing track of himself and then reorienting himself have for Sam concerning the proposition that the coin in Sleeping Beauty came up heads, the events of Beth's losing track of herself and then reorienting herself should have exactly the same epistemic bearing for Beth concerning the proposition that the coin in Duplication came up heads. Thus, if one endorses Beauty Half or Beauty Third, then one should correspondingly endorse Duplication Half or Duplication Third.

Consider, finally, the following variant on Duplication:

> **Duplication With Confidants:** Esme and Adam are subject to the following experiment. On Sunday night Esme will be put to sleep in a red room, and a coin will be flipped. If the coin lands heads, then she will be awoken on Monday. First, her eyes will be closed. She will then open her eyes to see that she is in the red room. If, however, the coin lands tails, then a perfect duplicate of Esme, $Esme_d$, will be placed in a blue room. Both Esme and $Esme_d$ will be woken up on Monday. First, their eyes will be closed. They will then open their eyes and each will discover which room they are in, and so, who they are. Throughout this process Adam will remain awake in a third room. He'll have no knowledge of the outcome of the coin toss. Adam and Esme, however, are perfect at first-personally introspecting their own epistemic states and at predicting one another's epistemic states. Both upon awakening, then, and when Esme opens her eyes, Adam and Esme will be such that it is common knowledge amongst them what each of their epistemic states is and that each is both first-personally and third-personally introspective. Esme and Adam are informed of all of these facts.

We can represent Adam and Esme's epistemic states, both upon Esme's awakening, and upon her opening her eyes, in the following centered agreement frame. Here we let w_h be the world in which the coin comes up heads and w_t the world in which the coin comes up tails. We let e be Esme, e^d $Esme_d$, and a Adam. And we let m_0 be the moment that Esme first awakens and m_1 the moment that she opens her eyes.

Frame 2: Let $\mathcal{C}_2$ be such that $\mathbf{W} = \{w_h, w_t\}$, $\mathbf{A} = \{e_{m_0}, e_{m_1}, e_{m_0}^d, e_{m_1}^d, a_{m_0}, a_{m_1}\}$, and $\mathbf{C} = \{\langle w_h, e_{m_0}\rangle, \langle w_h, e_{m_1}\rangle, \langle w_h, a_{m_0}\rangle, \langle w_h, a_{m_1}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_t, e_{m_1}\rangle, \langle w_t, e_{m_0}^d\rangle, \langle w_t, e_{m_1}^d\rangle, \langle w_t, a_{m_0}\rangle, \langle w_t, a_{m_1}\rangle\}$. We let $K_{e_{m_0}}(\cdot)$ be such that, for every $\vec{x} \in \mathbf{C}$, $K_{e_{m_0}}(\vec{x}) = \{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_h, e_{m_0}^d\rangle\}$; $K_{e_{m_1}}(\cdot)$ be such that, for every $\vec{x} \in \mathbf{C}$, $K_{e_{m_1}}(\vec{x}) = \{\langle w_h, e_{m_1}\rangle, \langle w_t, e_{m_1}\rangle\}$; $K_{a_{m_0}}(\cdot)$ be such that, for every $\vec{x} \in \mathbf{C}$, $K_{a_{m_0}}(\vec{x}) = \{\langle w_h, a_{m_0}\rangle, \langle w_t, a_{m_0}\rangle\}$, and $K_{a_{m_1}}(\cdot)$ be such that, for every $\vec{x} \in \mathbf{C}$, $K_{a_{m_1}}(\vec{x}) = \{\langle w_h, a_{m_1}\rangle, \langle w_t, a_{m_1}\rangle\}$. We, further, let $K_{e_{m_0}^d}(\cdot)$ be such that, for each $\vec{x} \in \mathbf{C}$ such that $w_x = w_h$, $K_{e_{m_0}^d}(\vec{x}) = \star$, while for each $\vec{x} \in \mathbf{C}$ such that $w_x = w_t$, $K_{e_{m_0}^d}(\vec{x}) = K_{e_{m_0}}(\vec{x}) = \{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_t, e_{m_0}^d\rangle\}$. And we let $K_{e_{m_1}^d}(\cdot)$ be such that, for each $\vec{x} \in \mathbf{C}$ such that $w_x = w_h$, $K_{e_{m_0}^d}(\vec{x}) = \star$, while for each $\vec{x} \in \mathbf{C}$ such that $w_x = w_t$, $K_{e_{m_1}^d}(\vec{x}) = \{\langle w_t, e_{m_1}^d\rangle\}$.

Note that, for each $\vec{x} \in \mathbf{C}$ and each $i \in G = \{e_{m_0}, a_{m_0}\}$, $\mathsf{Com}_G([K_i = K_i(\vec{x})])$, $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$ and $\mathsf{Com}_G(\mathsf{CHom}[K_i(\cdot)])$ hold. Given **Centered Confidants**, then, at each point in this frame, Esme and Adam are epistemic confidants upon Esme's awakening. Similarly, for each $\vec{x} \in \mathbf{C}$ and each $i \in G = \{e_{m_1}, a_{m_1}\}$, $\mathsf{Com}_G([K_i = K_i(\vec{x})])$, $\mathsf{Com}_G(\mathsf{Hom}[K_i(\cdot)])$, and $\mathsf{Com}_G(\mathsf{CHom}[K_i(\cdot)])$ hold. And so, given **Centered Confidants**, at each point in this frame, Esme and Adam are epistemic confidants at the moment that Esme opens her eyes.

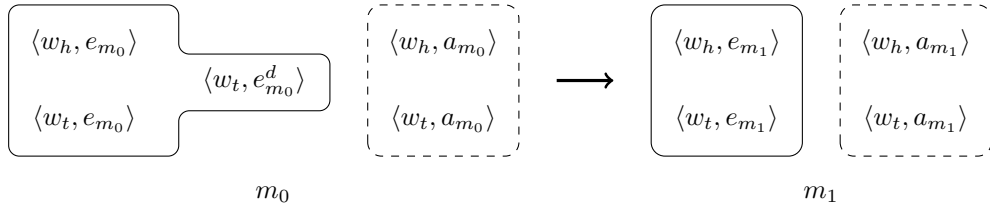We can represent some of the key features of this frame as in Figure 2.

Figure 2: $\mathcal{C}_2$

Here the cells to the left of the arrow represent Esme and Adam's epistemic states upon Esme's awakening, while the cells to the right of the arrow represent their epistemic states at the moment that Esme opens her eyes. The solid-lined cells correspond to Esme's epistemic states at these different moments, while the dashed-lined cells correspond to Adam's epistemic states at these different moments.

Corresponding to **Duplication Half** and **Duplication Third**, we have the following two prima facie plausible accounts of how Esme should update her credences in **Heads** throughout this case:

> **Confidant Half:** On Monday upon awakening but before opening her eyes Esme should give **Heads** credence $1/2$, and upon opening her eyes and learning that she is Esme in the red room she should give **Heads** credence $2/3$.

Confidant Third: On Monday upon awakening but before opening her eyes Esme should give Heads credence 1/3, and upon opening her eyes and learning that she is Esme in the red room she should give Heads credence 1/2.

I submit that if one endorses Duplication Half or Duplication Third, then one should endorse Confidant Half or Confidant Third. Here's one way of seeing this. In Duplication With Confidants, the series of possible events that might occur to Esme and how these are related to the toss of a fair coin are exactly as they are for Beth in Duplication. The only differences are that we've stipulated that Esme is perfect at first-personally introspecting her own epistemic state, and that there is another rational individual with whom Esme is an epistemic confidant, who will be awake while Esme is asleep and during her awakening, but who has no special knowledge of the outcome of the coin toss. Neither of these features of the case, though, would seem to make a difference to what credence Esme should have in Heads upon awakening or upon finding out that she is in the red room. For insofar as Duplication Half and Duplication Third are plausible, they should, be equally plausible given the stipulation that the agent in question is perfectly introspective. And, while in Duplication there is no additional agent such as Adam, whose epistemic state Beth is aware of and who is similarly aware of Beth's epistemic state, but who possesses no information about the outcome of the coin-toss, still it seems that Beth might imaginatively entertain that there is some such individual, at various points in this process, when considering what credences are rational. For Duplication allows for the existence of such an individual without this changing the basic set-up of the case. But it seems clear that when Beth considers her position in Duplication under this imaginative hypothesis, this should make no difference to what credences she should view as rational. And so it seems that it shouldn't make any difference to which credences would be rational, if there were, in fact, such an individual. But, then, the fact that in Esme's case there is some epistemic confidant who is present throughout the process but who has no special knowledge of the coin-toss, should make no difference to what credence it is rational for her to assign to Heads either upon awakening or upon finding out that she is in the red room.

Any plausible justification for Duplication Half, then, should provide an equally compelling justification for Confidant Half, while any plausible justification for Duplication Third should provide an equally compelling justification for Confidant Third.

Note, though, that whatever one says about Esme's credences in this case, it seems clear that throughout the process described in Duplication With Confidants, Adam should have credence 1/2 in Heads. For it seems patently obvious that Adam receives no information throughout this process that is relevant to the outcome of the coin toss. He simply sits in a room aware of the passage of time and the changes in his and Esme's mental states. But none of this would seem to be the sort of information that could make it rational for him to change his credence in Heads.

Given, this, though, it follows that if one endorses either Confidant Half or Confidant Third, then, given Centered Confidants, one must say that every rational update function is such that there are some cases in which epistemic confidants must assign different credences to some uncentered proposition given that they update in line with that function.

To see this, let $K_e^{m_0} = \{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_h, e_{m_0}^d\rangle\}$, $K_a^{m_0} = \{\langle w_h, a_{m_0}\rangle, \langle w_t, a_{m_0}\rangle\}$, $K_e^{m_1} = \{\langle w_h, e_{m_1}\rangle, \langle w_t, e_{m_1}\rangle\}$ and $K_a^{m_1} = \{\langle w_h, a_{m_1}\rangle, \langle w_t, a_{m_1}\rangle\}$. $K_e^{m_0}$ and $K_a^{m_0}$, then, represent Esme and Adam's epistemic states when Esme is first awoken, while $K_e^{m_1}$ and $K_a^{m_1}$ represent Esme and Adam's epistemic states when Esme finds herself in the red room. Note that, given Centered Confidants, $K_e^{m_0}$ and $K_a^{m_0}$ ensure that Adam and Esme are epistemic confidants upon Esme's awakening, while $K_e^{m_1}$ and $K_a^{m_1}$ ensure that Adam and Esme are epistemic confidants when Esme opens her eyes.

Now, if $s \in \mathcal{S}$ is such that, for some subintervals $[l, r)$ $[r, q]$, of $[0, n]$, if $x \in [l, r)$, then $s(x) = K_e^{m_0}$ and if $x \in [r, q]$, then $s(x) = K_e^{m_1}$, then, given Confidant Half, we have that if $\mathbf{u}$ is a rational update function, then, if $x \in [l, r)$, then $\mathbf{u}(s)(x) = Cr_e^{m_0}(\cdot)$, such that $Cr_e^{m_0}(\mathsf{Heads}) = 1/2$, while if $x \in [r, q]$, then $\mathbf{u}(s)(x) = Cr_e^{m_1}(\cdot)$, such that $Cr_e^{m_0}(\mathsf{Heads}) = 2/3$. And, if $s \in \mathcal{S}$ is such that, for some subintervals $[l, r)$ $[r, q]$, of $[0, n]$, if $x \in [l, r)$, then $s(x) = K_a^{m_0}$ and if $x \in [r, q]$, then $s(x) = K_a^{m_1}$, then, given Confidant Half, we have that if $\mathbf{u}$ is a rational update function, then, if $x \in [l, r)$, then $\mathbf{u}(s)(x) = Cr_a^{m_0}(\cdot)$, such that $Cr_a^{m_0}(\mathsf{Heads}) = 1/2$, while if $x \in [r, q]$, then $\mathbf{u}(s)(x) = Cr_a^{m_1}(\cdot)$, such that $Cr_a^{m_1}(\mathsf{Heads}) = 1/2$.

And so, given that agents with epistemic states $K_e^{m_1}$ and $K_a^{m_1}$ are epistemic confidants, it follows, given Confidant Half, that for any rational updating function, there is at least one possible case in which epistemic confidants who update in line with that function must assign different credences to some uncentered propositions. In particular, we have that there are cases in which, one epistemic confidant must assign credence $1/2$ to Heads, while another epistemic confidant must assign credence $2/3$ to Heads.

If, on the other hand, $s \in \mathcal{S}$ is such that, for some subintervals, $[l, r)$, $[r, q]$, of $[0, n]$, if $x \in [l, r)$, then $s(x) = K_e^{m_0}$ and if $x \in [r, q]$, then $s(x) = K_e^{m_1}$, then, given Confidant Third, we have that if $\mathbf{u}$ is a rational update function, then, if $x \in [l, r)$, then $\mathbf{u}(s)(x) = Cr_e^{m_0}(\cdot)$, such that $Cr_e^{m_0}(\mathsf{Heads}) = 1/3$, while if $x \in [r, q]$, then $\mathbf{u}(s)(x) = Cr_e^{m_1}(\cdot)$, such that $Cr_e^{m_0}(\mathsf{Heads}) = 1/2$. And, if $s \in \mathcal{S}$ is such that, for some subintervals $[l, r)$ $[r, q]$, of $[0, n]$, if $x \in [l, r)$, then $s(x) = K_a^{m_0}$ and if $x \in [r, q]$, then $s(x) = K_a^{m_1}$, then, given Confidant Third, we have that if $\mathbf{u}$ is a rational update function, then, if $x \in [l, r)$, then $\mathbf{u}(s)(x) = Cr_a^{m_0}(\cdot)$, such that $Cr_a^{m_0}(\mathsf{Heads}) = 1/2$, while if $x \in [r, q]$, then $\mathbf{u}(s)(x) = Cr_a^{m_1}(\cdot)$, such that $Cr_a^{m_1}(\mathsf{Heads}) = 1/2$.

And so, given that agents with epistemic states $K_e^{m_0}$ and $K_a^{m_0}$ are epistemic confidants, it follows, given Confidant Third, that for any rational updating function, there is at least one possible case in which epistemic confidants who update in line with that function must assign different credences to some uncentered propositions. In particular, we have that there are cases in which, one epistemic confidant must assign credence $1/2$ to Heads, while another epistemic confidant must assign credence $1/3$ to Heads.

Both Confidant Half and Confidant Third, then, are incompatible with Permissible Agreement, given Centered Confidants and some plausible auxiliary assumptions. We've argued, though, that Confidant Half stands or falls with Beauty Half, while Confidant Third stands or falls with Beauty Third. We have, then, that both Beauty Half and Beauty Third are incompatible with Permissible Agreement, given Centered Confidants and some plausible auxiliary

assumptions.

Let me close this section by considering a third account of how Sam ought to update his credences in Sleeping Beauty:

> **Beauty Double-Half:** On Monday upon awakening Sam should give Heads credence 1/2, and upon learning that it is Monday he should continue to give Heads credence 1/2.[20]

Interestingly, the preceding argument that showed that both Beauty Half and Beauty Third are incompatible with Permissible Agreement and Centered Confidants, cannot be used to show that Beauty Double-Half is similarly incompatible with these two principles. For consider the corresponding account of how Esme ought to update her credences in Duplication With Confidants:

> **Confidant Double-Half:** On Monday, upon awakening but before opening her eyes, Esme should give Heads credence 1/2, and, upon opening her eyes and learning that she is Esme in the red room, she should continue to give Heads credence 1/2.

If one endorses this claim, then, given the plausible claim that Adam should also have credence half in Heads both when Esme awakes and when she discovers which room she is in, it follows that, given any rational update function, Esme and Adam will assign the same

---

[20]For a defense of this position see Meacham [2008] and Halpern [2004]. A related account is developed in Halpern and Tuttle [1993]. It's perhaps worth noting that while these accounts of updating involve an endorsement of Strong Centered Anti-Bayesianism, the proponent of Beauty Double-Half need not endorse this principle.

The reason that these particular accounts endorse Strong Centered Anti-Bayesianism is that they assume a certain modeling of Sam's epistemic dynamics. In particular, it is assumed that upon awakening, Sam's epistemic state will consist of three possibilities: either (i) the coin came up heads and it's Monday, or (ii) the coin came up tails and it's Monday, or (iii) the coin came up heads and it's Tuesday, and that the result of being told that it is Monday is that his epistemic state is trimmed down to the following two possibilities: either (i) the coin came up heads and it's Monday, or (ii) the coin came up tails and it's Monday. And given this modeling of the dynamics of Sam's epistemic state, it follows that if Sam is to rationally assign credence 1/2 to Heads at both times, then he must not conditionalize on the same prior at both of these times.

This, however, is not an entirely realistic way of modeling the dynamics of Sam's epistemic state. For some time will have elapsed between when Sam wakes and when he is learns that it is Monday and Sam will, presumably, be aware that some time has passed. And so, Sam's epistemic state after being told that it isn't Monday won't be a subset of his epistemic state upon awakening. Instead, it will be a set that is like a subset of his epistemic state upon awakening but with the temporal elements of the centers moved forward. But given this dynamics, there's no problem with finding a prior credal function such that by conditioning on this function Sam will assign Heads credence 1/2 both upon awakening and after learning that it is Monday. Whether everything that the standard double-halfer wants to say can be accommodated in this manner is a subtle question that I won't try to answer. But the basic thought that in Sam's case he should not change his credence in Heads between the time that he wakes and the time that he learns that the present day is Monday is one that can be accommodated by one who thinks that rationality involves conditioning some prior on one's epistemic state at that time and that which prior one should conditionalize on will not change over time.

credence to Heads at both of these times. Confidant Double-Half, then, is compatible with Permissible Agreement and Centered Confidants. And so we don't have similar grounds for maintaining that Beauty Double-Half is incompatible with Permissible Agreement and Centered Confidants as we so with Beauty Half and Beauty Third.

5   Rejecting **Centered Confidants**?

We've said that a group of agents are **epistemic confidants** just in case their respective epistemic states are of the sort that would result from idealized communication amongst perfectly introspective and rational agents. Given this intuitive concept, I've argued that Permissible Agreement is quite plausible. Furthermore, we've seen that, given a putative pre-cisification of this intuitive concept, provided by Centered Confidants, Permissible Agreement is incompatible with some prima facie plausible principles concerning how agents should set their credences in light of self-locating information. At least at first sight, this conflict would seem to give us some reason to reject these prima facie plausible updating principles.

One way to resist the preceding arguments is to reject the putative precisification of the intuitive concept of a group of agents being epistemic confidants provided by Centered Confidants. In particular, one may block the preceding arguments by maintaining that, for agents with epistemic and credal states defined over algebras that include both centered and uncentered propositions, there is some constraint on a group of such agents being epistemic confidants whose satisfaction is not ensured by the conditions imposed by Centered Confidants.

In this section, I'll consider one such constraint that has at least some prima facie appeal.[21] Given this constraint, both the argument that Permissible Agreement is incompatible with Mandatory Center-Indifference and the argument that Permissible Agreement is incompatible with Beauty Third can be shown to fail. I'll argue, however, that if one wants to endorse this additional constraint, then one must maintain that, in certain situations in which two agents are communicating, but one agent is uncertain of their identity, neither agent can transmit to the other the evidential support that they have for uncentered propositions. I'll argue that this is, at the very least, a surprising claim. Ultimately, I think it's one that we should reject.

I noted earlier that, for agents whose epistemic states are defined over algebras that include both centered and uncentered propositions, there are two ways in which an individual may have knowledge of their own epistemic state. As with agents whose epistemic states are defined over algebras that include only uncentered propositions, such an agent may have third-personal knowledge of their own epistemic state. In addition, though, such an agent may have first-personal knowledge of their own epistemic state.

A tempting thought, which we'll now explore, is that, for agents whose epistemic states are defined over an algebra that includes both centered and uncentered propositions, in characterizing what it is for a group of such agents G to be epistemic confidants, we must not only require that each agent have first-personal knowledge of their own epistemic state,

---

[21]Thanks to Harvey Lederman for suggesting an additional constraint along these lines.

i.e., that there is some proposition K such that they know first-personally that their epistemic state is characterized by K, but, in addition, we must require that each agent know, in an appropriately first-personal way, that the members of G know that their epistemic state is characterized by K, and that each agent know, in an appropriately first-personal way, that the members of G know that the members of G know that their epistemic state is characterized by K, and so on.

To see how this will work, it will be useful to focus on how such states may be represented in a centered agreement frame. Given that an agent a is first-personally introspective and their epistemic state is characterized by K, it follows that they will know the centered proposition that their epistemic state is characterized by K. This centered proposition may be represented in a centered agreement frame by $\{\vec{z} : K_{a_z}(\vec{z}) = K\}$. And so, given that a is first-personally introspective and their epistemic state is characterized by K, we have $K_a(\{\vec{z} : K_{a_z}(\vec{z}) = K\})$.

We can represent the centered proposition that everyone in G knows that one's epistemic state is characterized by K, in a centered agreement frame, by $\{\vec{z} : \mathsf{M}_G^0[K_{a_z}(\vec{z}) = K]\}$. If, then, an agent a knows, first-personally, that everyone in G knows that their epistemic state is characterized by K, we will have $K_a(\{\vec{z} : \mathsf{M}_G^0[K_{a_z}(\vec{z}) = K]\})$. We can, similarly, represent the centered proposition that everyone in G knows that everyone in G knows that one's epistemic state is characterized by K, in a centered agreement frame, by $\{\vec{z} : \mathsf{M}_G^1[K_{a_z}(\vec{z}) = K]\}$, and so on. Finally, we can represent the centered proposition that it is commonly known amongst G that one's epistemic state is characterized by K, in a centered agreement frame, by $\{\vec{z} : \mathsf{Com}_G[K_{a_z}(\vec{z}) = K]\}$. If, then, an agent a knows, first-personally, that it is commonly known in G that their epistemic state is characterized by K, then we will have $K_a(\{\vec{z} : \mathsf{Com}_G[K_{a_z}(\vec{z}) = K]\})$. Note that if $K_a(\{\vec{z} : \mathsf{Com}_G[K_{a_z}(\vec{z}) = K]\})$ obtains at some point, then so will $K_a(\{\vec{z} : \mathsf{M}_G^n[K_{a_z}(\vec{z}) = K]\})$, for each $n \in \mathbb{N}$.

> **Def.** We say that an agent a is **first-personally G-introspective** just in case, if a's epistemic state is K, then a first-personally knows that it is commonly known in G that their epistemic state is characterized by K.

Consider now the following claim:[22]

---

[22]A few points that are perhaps worth highlighting. First, Group Introspection only provides a necessary condition on a group of agents being epistemic confidants. The proponent of Group Introspection will, I take it, accept that the conditions imposed by Centered Confidants are also necessary for any group of agents to be epistemic confidants. They will, however, deny that these conditions are sufficient.

Second, insofar as one is attracted to Group Introspection, there are further conditions involving first-personal higher-order knowledge of one's own and others epistemic states that one may naturally be inclined to require of a group of epistemic confidants. For example, given Group Introspection, it is also natural to require of a group of epistemic confidants that each $a \in G$ be such that, if their epistemic state is characterized by K, then, for any n iterations of $\mathsf{M}_G^0 K_{a_z}$, $K_a(\{\vec{z} : (\mathsf{M}_G^0 K_{a_z})^n\{\vec{x} : K_{a_x}(\vec{x}) = K\}\})$ holds. That is, if a set of agents G are epistemic confidants, then if some member a's epistemic state is characterized by K, then a will know, in a first-personal way, that each member of G knows that they know that their epistemic state is characterized by K, and a will know, in a first-personal way, that each member of G knows that they know that each member of G knows that they know that their epistemic state is characterized by K, and so on.

Group Introspection: If a group of agents G are epistemic confidants, then each $a \in G$ must be first-personally G-introspective.

The thought behind this principle, I take it, is that if a group of agents G have successfully communicated in the manner characteristic of epistemic confidants, then the epistemic states of the members of this group will be open to one another's view. But, so the thought goes, if such epistemic states are open to one another's view, then, given that some member of G's epistemic state is characterized by K, not only must they know, in a first-personal way, that their epistemic state is characterized by K, but, in addition, they must know, in a first-personal way, that each member of G knows that their epistemic state is characterized by K, and so on.

Group Introspection is, I think, at least prima facie plausible. It is, however, incompatible with Centered Confidants. For there are cases in which a group of agents G satisfy the conditions imposed by Centered Confidants, but at least some members of G fail to be first-personally G-introspective. Indeed, both the case that we appealed to in order to show that, given Centered Confidants, Permissible Agreement is incompatible with Mandatory Center-Indifference, and the case that we appealed to in order to show that, given Centered Confidants, Permissible Agreement is incompatible with Beauty Third have exactly this feature.

To see this, consider again Frame 2. Recall that $K_e^{m_0} = \{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_h, e_{m_0}^d\rangle\}$ represents Esme's epistemic state when she is first awoken, while $K_a^{m_0} = \{\langle w_h, a_{m_0}\rangle, \langle w_t, a_{m_0}\rangle\}$ represents Adam's epistemic state when Esme is first-awoken. Now, given these epistemic states, it follows from Centered Confidants that Esme and Adam are epistemic confidants. However, letting $G = \{e_{m_0}, a_{m_0}\}$, we can show that Esme, upon awakening, is not first-personally G-introspective. For $K_{e_{m_0}}(\{\vec{z} : \mathsf{M}_G^0[K_{a_z}(\vec{z}) = K_e^{m_0}]\})$ fails to hold at each point in this frame. The reason for this is that at no point in the frame does Esme, upon her awakening, know the centered proposition that Adam, upon Esme's awakening, knows that her epistemic state is given by $\{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_h, e_{m_0}^d\rangle\}$. For the centered proposition $\{\vec{z} : K_{a_{m_0}}[K_{a_z}(\vec{z}) = K_e^{m_0}]\}$ holds only at $\langle w_h, e_{m_0}\rangle$ and at $\langle w_t, e_{m_0}\rangle$. But, given that Esme's epistemic state is $K_e^{m_0}$, she does not know $\{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle\}$. Thus, Esme, upon her awakening, does not know $\{\vec{z} : K_{a_{m_0}}[K_{a_z}(\vec{z}) = K_e^{m_0}]\}$. And so, since, $\{\vec{z} : \mathsf{M}_G^0[K_{a_z}(\vec{z}) = K_e^{m_0}]\}$ entails $\{\vec{z} : K_{a_{m_0}}[K_{a_z}(\vec{z}) = K_e^{m_0}]\}$, it follows that Esme, upon her awakening, does not know $\{\vec{z} : \mathsf{M}_G^0[K_{a_z}(\vec{z}) = K_e^{m_0}]\}$.

The problem here is that Esme, upon awakening, does not know who she is. As far as she is concerned she might either be Esme or the duplicate of Esme, if such a duplicate exists. But while she knows that Adam knows that Esme's epistemic state is given by $\{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_h, e_{m_0}^d\rangle\}$, she also knows that Adam does not know that the duplicate's epistemic state is given by $\{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_h, e_{m_0}^d\rangle\}$. For Adam does not know that the duplicate exists.[23] So, since Esme is uncertain whether she is Esme or the duplicate,

---

The points that I'll make about Group Introspection, in what follows, also apply, mutatis mutandis, to proposals that would require additional first-personal higher-order knowledge conditions such as this. I'll leave the details, however, to the interested reader.

[23] Of course, if the duplicate exists, then her epistemic state is given by, $\{\langle w_h, e_{m_0}\rangle, \langle w_t, e_{m_0}\rangle, \langle w_h, e_{m_0}^d\rangle\}$. And Adam knows this. But if the duplicate doesn't exist, then she doesn't have any epistemic state. And

if such a duplicate exists, she is similarly uncertain whether Adam knows that *her* epistemic state is given by $\{\langle w_h, e_{m_0} \rangle, \langle w_t, e_{m_0} \rangle, \langle w_h, e_{m_0}^d \rangle\}$. Thus, while Esme knows first-personally what her epistemic state is, she does not know first-personally that Adam knows what her epistemic state is.

Exactly similar considerations show that, while the agents represented in Frame 1 are epistemic confidants, given Centered Confidants, they do not satisfy the constraints imposed by Group Introspection.

If, then, one rejects Centered Confidants by endorsing Group Introspection, one may block both the argument that Permissible Agreement is incompatible with Mandatory Center-Indifference and the argument that Permissible Agreement is incompatible with Beauty Third.[24]

While Group Introspection is prima facie plausible, there are reasons to reject this principle. For there are certain cases in which agents are able to communicate with one another, but because one agent is uncertain about their identity, having communicated, they fail to satisfy the conditions imposed by Group Introspection. I'll argue that, if one wants to endorse Group Introspection, then one must maintain that, in this sort of case, such agents are unable to transmit to one another the evidential support that they have for uncentered propositions. Since this seems to me at least prima facie implausible, we have some reason to reject Group Introspection. At the very least, we can see that Group Introspection leads to some surprising consequences about evidential transmission.

To see why all of this is so, consider the following variant on Duplication.

> **Duplication With Communication:** Esme and Adam are subject to the following experiment. On Sunday night Esme will be put to sleep in a red room, and a coin will be flipped. If the coin lands heads, then she will be awoken on Monday. First, her eyes will be closed. She will then open her eyes to see that she is in the red room. If, however, the coin lands tails, then a perfect duplicate of Esme, Esme$_d$, will be placed in a blue room. Both Esme and Esme$_d$ will be woken up on Monday. First, their eyes will be closed. They will then open their eyes and each will discover which room they are in, and so, who they are. Throughout this process Adam will remain awake in a third room. He'll have no knowledge of the outcome of the coin toss. Adam and Esme, however, are perfect at first-personally introspecting their own epistemic states and are perfectly rational.
>
> In addition, prior to opening her eyes, Esme and Adam are able to communicate their epistemic states to one another as follows. If the coin comes up heads then Esme and Adam will be outfitted with earbuds. A computer will record each of their reports on their epistemic states and relay the information in an automated voice to the other. If, on the other hand, the coin comes up tails, Esme, Esme$_d$, and Adam will all be outfitted with earbuds. A computer will record each of their

---

Adam knows this.

[24]It's worth noting, though, that endorsing Group Introspection will not suffice to block the argument that Permissible Agreement is incompatible with Beauty Half. For, letting $G = \{e_{m_1}, a_{m_1}\}$, given the epistemic states represented in Frame 2, both Esme, upon opening her eyes, and Adam, upon Esme's opening her eyes, are such that they are both first-personally G-introspective.

25

reports on their epistemic states. Since Esme and Esme$_d$ are perfect duplicates they are guaranteed to issue the same report. The computer will then issue Adam's report to Esme and Esme$_d$, and will issue Esme and Esme$_d$'s univocal report to Adam. Adam, of course, will not be told whether the report is a univocal report from Esme and Esme$_d$ or a single report from Esme. Esme and Adam are informed of all of these facts.

Just as in Duplication With Confidants, in Duplication With Communication, prior to Esme's opening her eyes, but after she and Adam have communicated, she and Adam fail to satisfy the conditions imposed by Group Introspection on their being epistemic confidants. For, the reason that, in Duplication With Confidants, Esme and Adam fail to satisfy the conditions imposed by Group Introspection on their being epistemic confidants, prior to Esme's opening her eyes, is that, in this case, prior to opening her eyes, Esme does not know whether she is Esme or the duplicate of Esme, if such a duplicate exists. But, in Duplication With Communication, it's also the case that Esme fails to know which individual she is after having communicated with Adam but prior to opening her eyes. For nothing that Adam could tell her would indicate to her whether she is Esme or the duplicate, if such an individual exists. Thus, since Esme doesn't know whether she is Esme or the duplicate, if ϕ characterizes her epistemic state after communicating with Adam, but prior to opening her eyes, she also will fail to know the centered proposition that Adam knows that her epistemic state is characterized by ϕ.

What will Adam and Esme's epistemic states look like, in Duplication With Communication, after they have communicated but before Esme has opened her eyes? I suggest that, mutatis matandis, they will look just like Esme and Adam's epistemic states in Duplication With Confidants prior to opening their eyes. Each will know the set-up of the case. Esme will be first-personally introspective and know that either she is Esme, after communication, but before opening her eyes, in a world where the coin came up heads, or she is Esme, after communication, but before opening her eyes, in a world where the coin came up tails, or she is Esme$_d$, after communication, but before opening her eyes, in a world where the coin came up tails. And Adam will be first-personally introspective and know that either he is Adam, after communication, but before Esme opens her eyes, in a world where the coin came up heads, or that he is Adam, after communication, but before Esme opens her eyes, in a world where the coin came up tails. Furthermore, it will be common-knowledge between the two that these are their respective epistemic states.

The question that we want to focus on is whether, in Duplication With Communication, after communicating, but prior to Esme's opening her eyes, Esme and Adam are epistemic confidants. Given the preceding, we can see that the proponent of Centered Confidants will answer this question affirmatively, while the proponent of Group Introspection will answer this question negatively.

The crucial issue for adjudicating between these two answers, I suggest, is whether Esme and Adam, by communicating with one another, in the manner specified in Duplication With Communication, are able to transmit to one another the evidential support that they have for uncentered propositions, in the manner in which such evidence may, quite plausibly, be

transmitted in normal communicative situations. For, the key principle concerning epistemic confidants that we appealed to in justifying **Permissible Agreement** was **Confidant Symmetry**. According to this principle, if a group of agents G are epistemic confidants, then, for any $a_1, a_2 \in G$, if ϕ is an uncentered proposition, then a_1's epistemic state provides a_1 with the same evidential support for ϕ, as a_2's epistemic state provides a_2. If, then, Esme and Adam, in communicating with one another, in the manner specified in **Duplication With Communication**, are able to transmit to one another the evidential support that they have for uncentered propositions, then we should, as **Centered Confidants** tells us, count them as epistemic confidants. And if one wants to endorse **Group Introspection**, and so maintain that Esme and Adam are not epistemic confidants, then one should maintain that, in **Duplication With Communication**, Esme and Adam are unable to transmit to one another the evidential support that they have for uncentered propositions.

It would be quite surprising, I think, if Esme and Adam are unable to transmit to one another the evidential support that they have for uncentered propositions, in this sort of case. For consider, again, the situation in **Duplication With Communication**. To make things vivid, we might imagine that Esme and Adam converse for awhile. Each tells the other what they know, and the other responds in kind, and so on. They continue to talk until they've reached a fixed-point where, given each agent's reasoning and their background knowledge, there is nothing more that one could tell the other that would be informative. To make things more vivid, we can imagine that the computer is set-up so that this conversation proceeds at a normal pace.[25]

On the face of it, this would certainly seem to be a case in which agents are exchanging information. And, on the face of it, it would seem that, by the time they have reached the fixed-point where there is nothing else that either can tell the other, they will have passed on to the other whatever evidential support they might have for uncentered propositions. But, if we are to endorse **Group Introspection**, then we must deny that this is so.

Now perhaps this is ultimately the lesson to be drawn from this case. That this is so, however, seems to me to be very much non-obvious. I conclude, then, that while one might reject **Centered Confidants** by endorsing **Group Introspection**, and so block the arguments that **Permissible Agreement** is incompatible with **Mandatory Center-Indifference** and **Beauty Third**, doing so requires one to impose some surprising restrictions on cases in which agents may transmit to one another evidential support for uncentered propositions.

If, however, we grant—what seems to me to be quite plausible—that, in this case, Esme and Adam are able to exchange information in such a manner that each agent's evidential support for uncentered propositions is transmitted to the other, then, **Centered Confidants** would seem to get this case exactly right. I'm, at least, inclined to take this as further evidence that **Centered Confidants** provides a reasonable precisification of the intuitive idea of a group of agents being epistemic confidants.

More generally, granting that Esme and Adam are able to exchange information in such

---

[25]Of course, if Esme$_d$ exists, then she too will be responding, at the exact same time, in the exact same manner as Esme. But this need not prevent Esme and Adam from conversing, via the computerized medium, in a perfectly natural fluid manner.

a manner that each agent's evidential support for uncentered propositions is transmitted to the other, we can see from this case that the tempting thought that motivated **Group Introspection** is misguided. The tempting thought was that if two or more perfectly introspective agents have epistemic states that are open to one another's view in the manner ensured by successful communication, then, if ϕ characterizes one of these agents' epistemic states, then not only will this agent know, first-personally, that their epistemic state is characterized by ϕ, but they will also know, first-personally, that each member of the group knows that their epistemic state is characterized by ϕ, and so on. Granting, though, that Esme and Adam are able to exchange information in such a manner that each agent's evidential support for uncentered propositions is transmitted to the other, we can see that there are cases of successful communication—in the sense that the agents manage to transmit to one another the evidential support that they have for uncentered propositions—in which one agent may fail to know, in a first-personal way, that the other knows what their epistemic state is. Granting, then, that Esme and Adam are able to transmit to one another their evidential support for uncentered propositions, it follows that the sense in which the epistemic states of a group will be open to one another's view, given that they have successfully communicated in the manner characteristic of epistemic confidants, does not require that each agent know, first-personally, that the other knows what their epistemic state is. It is enough, if their epistemic states are open to view in the manner characterized by the conditions imposed by **Centered Confidants**.

6 Rejecting **Permissible Agreement**?

The other way to resist the preceding arguments is to reject **Permissible Agreement**. In justifying **Permissible Agreement**, we appealed to two principles: **Confidant Symmetry** and **Weak Evidentialism**. One natural way of motivating the rejection of **Permissible Agreement**, then, is to reject **Confidant Symmetry**. Another is to reject **Weak Evidentialism**. In this section, I'll consider an argument against **Permissible Agreement** that targets **Weak Evidentialism**.

Before doing so, though, let me briefly say something about **Confidant Symmetry**. According to this principle, the epistemic states, K_{a_1}, K_{a_2}, of any two epistemic confidants, a_1, a_2, are such that K_{a_1} provides a_1 with the same evidential support for any uncentered proposition ϕ, as K_{a_2} provides a_2. To reject this claim, one must maintain that, on any way of making precise the intuitive idea of a group of epistemic confidants—that is, a group of agents with the sorts of epistemic states that might result from an idealized information pooling in which each agent discloses their epistemic state to the others—there may in principle be differences between the evidential support that such agents' epistemic states provide for some uncentered proposition.

Now perhaps this is the lesson to draw from the preceding discussion. It is, however, quite a radical conclusion. For, it seems to me that evidence for an uncentered proposition ϕ is just a consideration that counts in favor ϕ. But a consideration that counts in favor of an uncentered proposition ϕ would seem to be the sort of thing that is, at least in principle, transmissible. And so, it would seem that evidence for uncentered propositions is always, at least in principle, transmissible. Given this, though, **Confidant Symmetry** should seem quite

plausible.

Granting **Confidant Symmetry**, then, I want to now consider an argument against **Permissible Agreement** that instead focuses on **Weak Evidentialism**. To this end, I'll show that, given **Centered Confidants**, **Permissible Agreement** is incompatible with a prima facie plausible consequentialist constraint on rational update functions called **Credal Rule Consequentialism**. I'll argue, in addition, that this incompatibility can be used to target **Weak Evidentialism**. For we can show that, given the conjunction of **Centered Confidants** and **Confidant Symmetry**, **Weak Evidentialism** is incompatible with **Credal Rule Consequentialism**. Given **Centered Confidants** and **Confidant Symmetry**, then, we can see one way in which one might be motivated to reject **Permissible Agreement**. For **Permissible Agreement** is motivated by a prima facie plausible evidentialist constraint on rational update functions. But given plausible principles concerning epistemic confidants, we can see that this evidentialist constraint is incompatible with another prima facie plausible consequentialist constraint on update functions. If, then, one accepts these plausible principles concerning epistemic confidants and the consequentialist constraint on rational update functions, one will have reason to reject the evidentialist constraint on rational update functions and, so, **Permissible Agreement**.

Now I should note that I'm not particularly inclined to accept **Credal Rule Consequentialism**. For I think that **Weak Evidentialism** is more plausible that **Credal Rule Consequentialism**, and I'm inclined to accept both **Centered Confidants** and **Confidant Symmetry**. I won't, however, be especially concerned to argue for this point here. Instead, my goal is simply to develop a picture of epistemic rationality that is, I think, of some interest, and that might lead one to reject **Permissible Agreement**.

We begin with some assumptions about epistemic value. Given some $\phi \in \mathcal{P}(\mathbf{C})$, it is natural to think that there will be some way of measuring how good, from an epistemic perspective, a credence $x \in \mathbb{R}$ is, relative to some $\vec{z} \in \mathbf{C}$. Now, amongst those who think that there is some way of measuring the epistemic goodness of credences, there is controversy about how exactly this value should be measured.[26] The main points that I want to make, however, are independent of most of these controversies. To make these points, though, it will be helpful to have a particular measure in mind. Let us suppose, then, as many find plausible, that the epistemic goodness of credences may be measured by the so-called negative Brier score.[27] Thus:

> **Brier Utility:** For each $\vec{z} \in \mathbf{C}$, let $\vec{z}(\cdot) : \mathcal{P}(\mathbf{C}) \to \{0, 1\}$ be the characteristic function of $\vec{z}$.[28] Given some $\phi \in \mathcal{P}(\mathbf{C})$, the epistemic utility of credence x, relative to $\vec{z}$, is given by:

$$\mathcal{B}(\phi, x, \vec{z}) = -(x - \vec{z}(\phi))^2$$

---

[26] For discussion of how epistemic utility should be measured see, for example, Joyce [1998], Joyce [2009], Leitgeb and Pettigrew [2010] and Levinstein [2012].

[27] For discussion of the merits of the Brier score see Joyce [1998] and Joyce [2009]. For further discussion see Predd et al. [2009] and Pettigrew [2016]. For an argument for **Beauty Third** that appeals to the Brier score see Kierland and Monton [2005].

[28] This is the function mapping members of $\mathcal{P}(\mathbf{C})$ to the value 1 just in case they are true relative to $\vec{z}$.

Given Brier Utility, there is a natural way of measuring how good the epistemic consequences are for a policy of giving ϕ credence x, given an epistemic state K. Thus:

Def. For each $K \in \mathcal{P}(\mathbf{C})$, let $\mathbf{C}_K =_{df} \{\vec{z} : K_{a_z}(\vec{z}) = K\}$.

Def. Given some $\phi \in \mathcal{P}(\mathbf{C})$. We let:

$$\mathcal{V}(\phi, x, K) = \frac{1}{|\mathbf{C}_K|} \sum_{\vec{z} \in \mathbf{C}_K} \mathcal{B}(\phi, x, \vec{z})$$

So defined, $\mathcal{V}(\phi, x, K)$ serves to measure the average epistemic utility that results from an agent with an epistemic state K having credence x in ϕ, where, in determining the average, each such agent's epistemic utility is weighted equally.[29] Since, quite plausibly, in assessing the epistemic consequences for a policy of giving ϕ credence x, given an epistemic state K, each possible agent with epistemic state K should be given equal weighting, $\mathcal{V}(\phi, x, K)$ provides a natural measure of how good the epistemic consequences of such a policy are.

Consider now the following consequentialist constraint on rational update functions:

Credal Rule Consequentialism: Necessarily, any rational update function $\mathbf{u}$ is such that, if there is some x such that $\mathcal{V}(\phi, x, K) > \mathcal{V}(\phi, y, K)$, for all $y \neq x$, then, if a has epistemic state K and updates in line with $\mathbf{u}$, then a will assign credence x to ϕ.

Given Brier Utility, this principle tells us that rational updating policies require an agent to assign a credence x to some proposition ϕ given an epistemic state K, when the policy of assigning credence x to ϕ, given K, leads to the better overall epistemic consequences than the policy of assigning credence y to ϕ, given K, for any $y \neq x$.

We now note the following fact:

$\mathcal{V}$-**Maximization:** In Duplication With Confidants, $\mathcal{V}(\mathsf{Heads}, \cdot, K_e^{m_0})$ is uniquely maximized at $1/3$, while $\mathcal{V}(\mathsf{Heads}, \cdot, K_a^{m_0})$ is uniquely maximized at $1/2$.[30]

---

[29] Or at least it does in cases in which the truth of ϕ does not depend in any constitutive way on the agent's credence in ϕ. For cases in which such dependence arises, see Caie [2013]. Such dependence, however, is absent in the sorts of cases we're considering, so we can safely ignore this complication.

[30] **Justification:** First, we'll justify the claim that, for all $x \in \mathbb{R}$, such that $x \neq 1/3$, $\mathcal{V}(\mathsf{Heads}, 1/3, K_e^{m_0}) > \mathcal{V}(\mathsf{Heads}, x, K_e^{m_0})$. To see this, note that $\mathcal{V}(\mathsf{Heads}, y, K_e^{m_0}) = 1 - [1/3 \sum_{\vec{z} \in \mathbf{C}_{K_e^{m_0}}} [(y - \vec{z}(\mathsf{Heads})^2]]$. Now, since we have that $\mathbf{C}_{K_e^{m_0}} = K_e^{m_0} = \{\langle w_h, e_{m_0} \rangle, \langle w_t, e_{m_0} \rangle, \langle w_h, e_{m_0}^d \rangle\}$, and that $\mathsf{Heads} = \{\vec{z} : w_z = w_h\}$, it follows that $\mathcal{V}(\mathsf{Heads}, y, K_e^{m_0}) = 1 - [1/3[(y-1)^2 + (y-0)^2 + (y-0)^2]] = 1 - [1/3[3y^2 - 2y + 1]]$. And since $1/3[3y^2 - 2y + 1]$ attains a unique minimum value at $y = 1/3$, we have that $1 - [1/3[3y^2 - 2y + 1]]$ attains a unique maximum value at $y = 1/3$.

Next, we'll justify the claim that, for all $x \in \mathbb{R}$, such that $x \neq 1/2$, $\mathcal{V}(\mathsf{Heads}, 1/2, K_a^{m_0}) > \mathcal{V}(\mathsf{Heads}, x, K_e^{m_0})$. To see this note that $\mathcal{V}(\mathsf{Heads}, y, K_a^{m_0}) = 1 - [1/2 \sum_{\vec{z} \in \mathbf{C}_{K_a^{m_0}}} [(y - \vec{z}(\mathsf{Heads})^2]]$. Now, since we have that $\mathbf{C}_{K_a^{m_0}} = K_a^{m_0} = \{\langle w_h, a_{m_0} \rangle, \langle w_t, a_{m_0} \rangle\}$, and that $\mathsf{Heads} = \{\vec{z} : w_z = w_h\}$, it follows that $\mathcal{V}(\mathsf{Heads}, y, K_a^{m_0}) = 1 - [1/2[(y-1)^2 + (y-0)^2]] = 1 - [1/2[2y^2 - 2y + 1]]$. And since $1/2[2y^2 - 2y + 1]$ attains a unique minimum value at $y = 1/2$, we have that $1 - [1/2[2y^2 - 2y + 1]]$ attains a unique maximum value at $y = 1/2$.

Given this, we can show:

> **Incompatibility A:** Permissible Agreement is incompatible with Credal Rule Consequentialism, given Centered Confidants.

To see that Incompatibility A follows from $\mathcal{V}$-Maximization, assume that Centered Confidants holds. Given this, we have that, in Duplication With Confidants, upon Esme's awakening, Esme and Adam are epistemic confidants. And so, given Permissible Agreement, it follows that there is some rationally permissible update function, such that, if Esme and Adam, upon Esme's awakening, both update in line with this function, then they will assign the same credence to Heads. However, given $\mathcal{V}$-Maximization, it follows from Credal Rule Consequentialism, that for any rational update function, if Esme and Adam, upon Esme's awakening, both update in line with this function, then they will assign different credences to Heads. Thus, given Centered Confidants, it follows that Permissible Agreement and Credal Rule Consequentialism are incompatible.

Given Centered Confidants and Credal Rule Consequentialism, then, we have reason to reject Permissible Agreement. Since both Centered Confidants and Credal Rule Consequentialism have some plausibility, there is at least some reason to be given for rejecting Permissible Agreement.

To get a better sense of the picture of epistemic rationality that might, in this manner, motivate the rejection of Permissible Agreement, note that the following is also a consequence of $\mathcal{V}$-Maximization.

> **Incompatibility B:** Weak Evidentialism is incompatible with Credal Rule Consequentialism, given Centered Confidants and Confidant Symmetry.

To see this, assume that Centered Confidants and Confidant Symmetry both hold. Given Centered Confidants we have that, in Duplication With Confidants, upon Esme's awakening, Esme and Adam are epistemic confidants. And so, given Confidant Symmetry, we have that, for each uncentered proposition ϕ, Esme's epistemic state provides Esme with the same evidential support for ϕ, that Adam's epistemic state provides for him. Given Weak Evidentialism, then, we have that there is some rationally permissible update function, such that, if Esme and Adam, upon Esme's awakening, both update in line with this function, then they will assign the same credence to Heads. However, given $\mathcal{V}$-Maximization, it follows from Credal Rule Consequentialism, that for any rational update function, if Esme and Adam, upon Esme's awakening, both update in line with this function, then they will assign different credences to Heads. Thus, given the assumption that Centered Confidants and Confidant Symmetry both hold, we can see that Weak Evidentialism and Credal Rule Consequentialism are incompatible.

Given Centered Confidants and Confidant Symmetry, then, we can articulate the following rationale for rejecting Permissible Agreement. First, as the proof of Incompatibility B shows, given these two plausible principles concerning epistemic confidants, it follows that there may be cases in which the epistemic states, K_{a_1}, K_{a_2}, of two agents, a_1, a_2, are such that K_{a_1} provides a_1 with the same evidential support for some uncentered proposition ϕ, as

K_{a_2} provides a_2, and yet the policy that leads to the best overall epistemic outcomes for agents with epistemic state K_{a_1} mandates credence x in ϕ, while the policy that leads to the best overall epistemic outcomes for agents with epistemic state K_{a_2}, mandates credence y in ϕ, where $x \neq y$. Given this, one can't endorse both Weak Evidentialism and Credal Rule Consequentialism. But, according to this line of thought, Permissible Agreement is plausible only insofar as it is motivated by Weak Evidentialism. If, then, one thinks that Credal Rule Consequentialism should be endorsed over Weak Evidentialism, then one will want to reject Permissible Agreement. Permissible Agreement, according to this line of thought, fails because, in certain cases, epistemic consequentialist considerations require agents to assign different credences to some proposition ϕ even when those agents' respective epistemic states provide them with the same evidential support for ϕ.[31]

As I noted, I'm not inclined to think that this argument provides particularly strong grounds for rejecting Permissible Agreement. For, given a conflict between Weak Evidentialism and Credal Rule Consequentialism, I'm inclined to reject Credal Rule Consequentialism. Nonetheless, I think this argument has some independent interest. Let me close this section, then, by noting one interesting consequence of this argument.

Recently, a number of authors have pointed out that, in certain cases, the credences that are motivated by epistemic consequentialist considerations need not match up with those that are motivated by evidentialist considerations. For example, there are possible situations where, if an agent has low credence in some obvious truth, then their other credences will be extremely accurate, while if they fail to have a low credence in this obvious truth, then their other credences will not be as accurate.[32] In this sort of case, evidentialist considerations demand that one have high credence in the obvious truth, while epistemic consequentialist considerations demand that one have low credence in this truth. There are also cases in which it can be shown that an agent will maximize their epistemic utility by having a particular set of probabilistically incoherent credences.[33] Assuming, as seems plausible, that evidential support relations are probabilisitically coherent, we have, then, another sort of case in which an agent's evidence demands one set of credences while epistemic consequentialist considerations demand another.

Now, in both of these sorts of cases, the conflict between evidentialist and epistemic consequentialist considerations arises because there is some trade-off whereby the agent sacrifices how epistemically good their credence is in one proposition ϕ, in order to secure greater epistemic utility with respect to some other proposition(s) ψ. Given this, a natural question that arises is whether evidentialist and epistemic consequentialist considerations will converge, when we ignore cases in which such trade-offs arise?

An interesting feature of the above argument is that it may be leveraged to provide a negative answer to this question. For the argument for Incompatibility B shows that, given

---

[31]It's worth noting that while an appeal to Credal Rule Consequentialism may be used by proponents of Beauty Third to block the preceding arguments, the same is not true for the proponent of Beauty Half. For, it's easy to show that $\mathcal{V}(\mathsf{Heads}, \cdot, K_e^{m_1})$ and $\mathcal{V}(\mathsf{Heads}, \cdot, K_a^{m_1})$ are both uniquely maximized at $1/2$.

[32]For discussion of so-called epistemic bribes see: Jenkins [2007], Berker [2013], Greaves [2013] and Carr [n.d.].

[33]See Caie [2013] for this sort of case.

Centered Confidants and Confidant Symmetry, in Duplication with Confidants, evidentialist and epistemic consequentialist considerations offer conflicting verdicts either about what credence Esme, upon awakening, ought to have in Heads, or what credence Adam, upon Esme's awakening, ought to have in Heads. However, insofar as there is a conflict between evidentialist and epistemic consequentialist considerations in this sort of case, it isn't one that depends on any epistemic trade-off involving Heads and some other proposition. And so, given Centered Confidants and Confidant Symmetry, it follows that there are cases in which evidentialist and epistemic consequentialist considerations fail to converge that don't depend on the possibility of an agent sacrificing how epistemically good their credence is in one proposition ϕ, in order to secure greater epistemic utility with respect to some other proposition(s) ψ.

7   Conclusion

I've argued that we have good reason to endorse both Centered Confidants and Permissible Agreement. Given this, we have good reason to reject some prima facie plausible principles concerning how agents should update their credences in light of self-locating information. First, we have good reason to reject the prima facie plausible indifference principle Mandatory Center-Indifference. Second, we have good reason to reject both Beauty Third and Beauty Half.

   To resist these arguments one must reject either Centered Confidants or Permissible Agreement. I've argued that rejecting either of these leads to some surprising conclusions. Perhaps some of these surprising conclusions are the ones that we should ultimately draw given the preceding arguments. For my part, however, I'm inclined to instead accept the surprising constraints on rational updating that Centered Confidants or Permissible Agreement impose.

References

Robert Aumann. Agreeing to disagree. *The Annals of Statistics*, pages 1236–1239, 1976.

Robert Aumann, Sergiu Hart, and Motty Perry. The forgetful passenger. *Games and Economic Behavior*, 20(1):117–120, 1997.

Jon Barwise. Three views of common knowledge. In *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 365–379. Morgan Kaufmann Publishers Inc., 1988.

Selim Berker. Epistemic teleology and the seperateness of propositions. *The Philosophical Review*, 122:337–393, 2013.

Nick Bostrom. Observer-relative chances in anthropic reasoning? *Erkenntnis*, 52:93–108, 2000.

Nick Bostrom. *Anthropic Bias.* Routledge, 2002.

Rachael Briggs. Putting a value on beauty. In Tamar Szabo Gendler and John Hawthorne, editors, *Oxford Studies in Epistemology*, volume 3. Oxford University Press, 2010.

Michael Caie. Rational probabilistic incoherence. *The Philosophical Review*, 122(4):527–575, 2013.

Michael Caie. Agreement theorems for self-locating belief. (ms), 2015.

Jennifer Carr. Epistemic utility theory and the aim of belief. MS, n.d.

Cian Dorr. Sleeping beauty: In defence of elga. *Analysis*, 62(276):292–296, 2002.

Cian Dorr and Frank Arntzenius. Self-locating priors and cosmological measures. (ms), 2015.

Adam Elga. Self-locating belief and the sleeping beauty problem. *Analysis*, 60(266):143–147, 2000.

Hilary Greaves. Epistemic decision theory. *Mind*, 122(488):915–952, 2013.

Joseph Halpern. Sleeping beauty reconsidered: Conditioning and reflection in asynchronous systems. In *Proceedings of the Twentieth Conference on Uncertainty in AI*, pages 226–234, 2004.

Joseph Halpern and Mark Tuttle. Knowledge, probability and adversaries. *Journal of the ACM*, 40(4):917–960, 1993.

C.S. Jenkins. Entitlement and rationality. *Synthese*, 157:25–45, 2007.

James Joyce. A non-pragmatic vindication of probabilism. *Philosophy of Science*, 65:575–603, 1998.

James Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber and C. Schmidt-Petri, editors, *Degrees of Belief*. Synthese Library, 2009.

Brian Kierland and Brandley Monton. Minimizing inaccurracy for self-locating beliefs. *Philosophy and Phenomenological Research*, 70(2):384–395, 2005.

Harvey Lederman. Uncommon knowledge. (ms), 2016.

Hannes Leitgeb and Richard Pettigrew. An objective justification of bayesianism I: Measuring inaccuracy. *Philosophy of Science*, 77, 2010.

John Leslie. Observer-relative chances and the doomsday argument. *Inquiry*, 40:427–436, 1997.

Benjamin Levinstein. Leitgeb and Pettigrew on accuracy and updating. *Philosophy of Science*, 79(3):413–424, 2012.

David Lewis. *Convention*. Cambridge University Press, 1969.

David Lewis. Sleeping beauty: Reply to Elga. *Analysis*, 61(271):171–176, 2001.

David Lewis. How many lives has Schrödinger's cat? *Australasian Journal of Philosophy*, 82(1):3–22, 2004.

Christopher Meacham. Sleeping beauty and the dynamics of de se beliefs. *Philosophical Studies*, 138(2):245–269, 2008.

Sarah Moss. Updating as communication. *Philosophy and Phenomenological Research*, 85 (2):225–248, 2012.

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

Michele Piccione and Ariel Rubenstein. On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior*, 20(1):3–24, 1997.

Joel Predd, Sanjeev Kulkarni, Robert Seiringer, Elliott H. Lieb, Daniel Osherson, and H. Vincent Poor. Probabilistic coherence and proper scoring rules. *IEEE Transactions on Information Theory*, 55(10):4786–4792, 2009.

Michael Titelbaum. The relevance of self-locating beliefs. *The Philosophical Review*, 117(4): 555–606, 2008.