For True Conditionalizers Weisberg's Paradox is a False Alarm

Franz Huber

Abstract: Weisberg (2009) introduces a phenomenon he terms *perceptual undermining*. He argues that it poses a problem for Jeffrey conditionalization (Jeffrey 1983), and Bayesian epistemology in general. This is *Weisberg's paradox*. Weisberg (2014) argues that perceptual undermining also poses a problem for ranking theory (Spohn 2012) and for Dempster-Shafer theory (Shafer 1976). In this note I argue that perceptual undermining does not pose a problem for any of these theories: for true conditionalizers Weisberg's paradox is a false alarm.

Keywords: perceptual undermining, Weisberg's paradox, Jeffrey conditionalization, Bayesian epistemology, ranking theory

1. Weisberg's Paradox

Weisberg's paradox consists in the inconsistency of four seemingly plausible constraints. It arises from the following example. Let D be the proposition that the sock really is red, and let F be the hypothesis that the lighting makes all socks look red. At time t_0 Sophia does not believe that the sock really is red and has a low degree of belief in D. Between t_0 and time t_1 she has a visual experience by looking at the sock. This visual experience between t_0 and t_1 causes her, among other things, to form the belief that the sock really is red at t_1; it leads to an increase in her degree of belief in D at t_1. At time t_2 she becomes certain that the lighting makes all socks look red; she assigns a maximal degree of belief to F at t_2. Since F is supposed to undermine the visual experience she has had between t_0 and t_1, this should make her drop her newly acquired belief that the sock really is red at t_2 again, and lower her degree of belief in D at t_2 back to what it was at t_0.

In a probabilistic setting this story is claimed to give rise to the following three constraints:

0. At time t_0 Sophia's degree of belief in D is low, say $\Pr_0(D) = .1$.

1a. At time t_1 Sophia's degree of belief in D is high, say $\Pr_1(D) = .9$.

2. At time t_2 Sophia's degree of belief in D is low again, $\Pr_2(D) = \Pr(D \mid F) = Pr_0(D) = .1$.

Sophia's degree of belief function at time t_2, Pr_2, comes from her degree of belief function at time t_1, Pr_1, by an application of strict conditionalization to the hypothesis F. Strict conditionalization is the following update rule:

Update Rule 1 (Strict Conditionalization) If $\mathrm{Pr}_0(\cdot)$: $\mathcal{A} \rightarrow \mathbb{R}$ is the ideal doxastic agent's probability measure at time t_0, and between t_0 and t_1 she becomes certain of the proposition $E \in \mathcal{A}$, Pr_0 (E) > 0, in the sense that $\mathrm{Pr}_1(E)$ = 1, and she does not become certain of any logically stronger proposition (and her probabilities are not directly affected in any other way such as forgetting etc.), then her probability measure at time t_1 should be $\mathrm{Pr}_1(\cdot)$:$\mathcal{A} \rightarrow \mathbb{R}$

$$\mathrm{Pr}_1(\cdot) = \mathrm{Pr}_E(\cdot) = \mathrm{Pr}_0(\cdot \mid E) = \mathrm{Pr}_0(\cdot \cap E) \,/\, \mathrm{Pr}_0(E)\,.$$

It is important to note that E is assumed to be the *total* evidence the ideal doxastic agent receives between time t_0 and time t_1.

Sophia's degree of belief function at time t_1, Pr_1, is assumed to come from her degree of belief function at time t_0, Pr_0, by an application of Jeffrey conditionalization to the evidential partition $\{D, \overline{D}\}$ with input parameters $\mathrm{Pr}_1(D) = .9$ and $\mathrm{Pr}_1(\overline{D}) = .1$. This assumption will turn out to be crucial. Hence it is stated as an independent constraint:

1b. Sophia's degree of belief function at time t_1, Pr_1, comes from her degree of belief function at time t_0, Pr_0, by an application of Jeffrey conditionalization to the evidential partition $\{D, \overline{D}\}$ with input parameters $\mathrm{Pr}_1(D) = .9$ and $\mathrm{Pr}_6(\overline{D}) = .1$.

Jeffrey conditionalization is the following update rule:

Update Rule 2 (Jeffrey Conditionalization, Jeffrey 1983) *If* $\mathrm{Pr}_0(\cdot)$:$\mathcal{A} \rightarrow \mathbb{R}$ *is the ideal doxastic agent's probability measure at time* t_0, *and between* t_0 *and* t_1 *her probabilities on the evidential partition* $\{E_i \in \mathcal{A} : i \in I\}$ *directly change to* $p_i \in \mathbb{R}$, *where* $\sum_{i \in I} p_i = 1$, *and* $p = 0$ *if* $\mathrm{Pr}_0(E_i) = 0$, *and her positive probabilities do not directly change on any finer partition (and her probabilities are not directly affected in any other way such as forgetting etc.), then her probability measure at time* t_1 *should be* $\mathrm{Pr}_1(\cdot)$:$\mathcal{A} \rightarrow \mathbb{R}$

$$\mathrm{Pr}_1(\cdot) = \mathrm{Pr}_{E_i \rightarrow p_i}(\cdot) = \sum_{i \in I} \mathrm{Pr}_0(\cdot \mid E_i) \cdot p_i.$$

It is important to note that the evidential partition $\{E_i \in \mathcal{A} : i \in I\}$ and the input parameters $p_i \in \mathbb{R}$ are assumed to be a *complete* description of all doxastically relevant events that happen between t_0 and t_1. In our example it is important to note that constraint (1b) amounts to the assumption that the only doxastically relevant effect of the visual experience between t_0 and t_1 is that Sophia's degrees of belief in D and $\overline{D}$ change to $\mathrm{Pr}_1(D) = .9$ and $\mathrm{Pr}_1(\overline{D}) = .1$. In particular, it follows from constraint (1b) that F is not directly affected, in any doxastically relevant way, by Sophia's visual experience between t_0 and t_1. Among other things, this means that, at t_1, Sophia does not also form a belief about how she came to be more confident in D (or, if she has a belief at t_1 about how she came to be more confident in D, then it is the same belief that she had at t_0, *before* she had the visual experience). In order to simplify matters, let us assume, as is reasonable, that, at t_0, Sophia has no particular belief about what will happen between t_0 and t_1. Then constraint (1b) implies that, at t_1, Sophia remains agnostic as to whether it was by vision, or by some other form of perception, or by testimony, or by clairvoyance that she became more confident in D.

In other words, unless she already does so at t_0, Sophia does not believe at t_1 that the experience she undergoes between t_0 and t_1 is a *visual* experience. For all she believes at t_1, the experience she undergoes between t_0 and t_1 may not even be a *perceptual* experience. Indeed, given our reasonable assumption, for all she believes at t_1, what happens between t_0 and t_1 may not even be an *experience* of hers. *We* know this from the way the story was told, but, according to constraint (1b), *she* does not.[1]

So far, so good. Now the allegedly bad news. Jeffrey conditionalization is *rigid*. This implies that Jeffrey conditionalization preserves probabilistic independence of the members of the evidential partition. If a proposition A is independent of the evidential proposition D according to Sophia's degree of belief function at time t_0, $\Pr_0(A \mid D) = \Pr_0(A)$, then A is also independent of D according to her degree of belief function at a time t_1, $\Pr_1(A \mid D) = \Pr_1(A)$.

Weisberg (2009; 2014) thinks that this poses a problem for Jeffrey conditionalization, and Bayesian epistemology in general. He does so, because there is the following fourth constraint:

3. F is independent of D according to Sophia's degree of belief function at time t_0, $\Pr_0(F \mid D) = \Pr_0(F)$.

Weisberg (2009; 2014) is, of course, free to stipulate any constraints he wants. However, in order to evaluate the joint plausibility of his constraints, we need to understand what this condition says, and what it does not say. The condition says: at t_0, the degree to which Sophia believes that the lighting makes all socks look red has no bearing on the degree to which she believes that the

[1] If, contrary to what constraint (1b) implies, Sophia had even the slightest of hunches about what might have caused the changes in her degrees of belief in D and in $\overline{D}$, the evidential partition would include other propositions besides D and $\overline{D}$. In this case Weisberg's paradox would not arise.

sock really is red. The condition does *not* say: at t_0, whether or not the lighting makes all socks look red has no bearing on whether or not the sock really is red.

Together with the rigidity of Jeffrey conditionalization this fourth constraint implies that F is independent of D according to Sophia's degree of belief function at time t_1, $\mathrm{Pr}_1(F \mid D) = \mathrm{Pr}_1(F)$. However, if F is independent of D according to Sophia's degree of belief function at time t_1, $\mathrm{Pr}_1(F \mid D) = \mathrm{Pr}_1(F)$, then D is independent of F according to her degree of belief function at time t_1, $\mathrm{Pr}_1(D \mid F) = \mathrm{Pr}_1(D)$. In this case her degree of belief in D at time t_2 equals her degree of belief in D at time t_1, $Pr_2(D) = \mathrm{Pr}_1(D) = .9$. This contradicts the third constraint according to which $\mathrm{Pr}_2(D) = .1$.

Weisberg's paradox consists in this inconsistency. In the next section I will defend Jeffrey conditionalization and argue that Weisberg's paradox is resolved once we notice the implications of constraints (0-1) and (3) – in particular, (1b) – which make constraint (2) utterly implausible.

2. Jeffrey Conditionalization Defended

Weisberg's paradox consists in the inconsistency of the four seemingly plausible constraints (0-3), where Jeffrey conditionalization is part of constraint (1). For Weisberg (2009; 2014) the culprit is constraint (1) with Jeffrey conditionalization. I want to defend Jeffrey conditionalization.

Constraints (0-1) imply that what Sophia has experienced between t_0 and t_1 when looking at the sock results in an increase in her degree of belief in the proposition D that the sock really is red: $\mathrm{Pr}_1(D) > \mathrm{Pr}_0(D)$. However, in conjunction with the allegedly plausible constraint (3) constraints (0-1) also imply that what Sophia has experienced between t_0 and t_1 makes her hold onto her degree of belief in the hypothesis F that the lighting makes all socks look red: $\mathrm{Pr}_1(F) = \mathrm{Pr}_0(F)$. I want to use this consequence of constraints (0-1) and (3) to motivate my defense of Jeffrey conditionalization.

Roughly speaking, the former inequality $\mathrm{Pr}_1(D) > \mathrm{Pr}_0(D)$ says that at t_1 Sophia thinks that what she has experienced between t_0 and t_1 – or better, as Sophia may not even believe that it was an experience: what has happened between t_0 and t_1 – is related to D. The difference between $\mathrm{Pr}_1(D)$ and $\mathrm{Pr}_0(D)$, in whichever way it is measured, reflects how likely she thinks, at t_1, that what has happened between t_0 and t_1 is related to D. The latter equation $\mathrm{Pr}_1(F) = \mathrm{Pr}_0(F)$ says, roughly, that at t_1 she thinks that what has happened between t_0 and t_1 has nothing to do with F. The difference between $\mathrm{Pr}_1(F)$ and $\mathrm{Pr}_0(F)$, in whichever way it is measured, is nil, and that is how likely she thinks it, at t_1, that what has happened between t_0 and t_1 is related to F. In particular, this latter equation says, roughly, that, at t_1, Sophia thinks that F is not a potential underminer for what has happened between t_0 and t_1.

Less roughly speaking, on the one hand constraints (0-1) say that Sophia's belief in D is directly affected by what happens between t_0 and t_1 in such a way that she ends up being more confident in D at t_1 than she was at t_0. On the other hand constraints (0-1) say that Sophia's belief in F is *not* directly affected by what happens between t_0 and t_1. Constraints (0-1) and the allegedly plausible stipulation (3) add to this that Sophia's belief in F is not indirectly affected by what happens between t_0 and t_1 either. Together constraints (0-1) and (3) thus say that Sophia's belief in F is neither directly nor indirectly affected by what happens between t_0 and t_1. In particular, constraints (0-1) and (3) say that, at t_1, Sophia thinks that F is not a potential underminer for what has happened between t_0 and t_1.[2]

In other words, constraints (0-1) and (3) imply that at t_1 Sophia thinks that what has happened between t_0 and t_1 may be due the fact that the sock really is red, D, but is definitely not due to the fact that the lighting makes all socks look red, F. Given this consequence of constraints (0-1) and (3), constraint (2) clearly should be rejected: constraint (2) says that, at t_1, Sophia thinks that F is a potential underminer for what has happened between t_0 and t_1, whereas it follows from constraints (0-1) and (3) that, at t_1, she thinks that F is not a potential underminer for what has happened between t_0 and t_1.

Here is a different way of putting things. Suppose we modeled the change of Sophia's degree of belief function from t_0 to t_1 by an application of strict conditionalization to the appearance proposition D^* that the sock appears to be red, $\Pr_1(\cdot) = \Pr_0(\cdot \mid D^*)$. In this case Sophia would believe that the apparent color of the sock is relevant to the actual color of the sock, $\Pr_1(D) = \Pr_0(D \mid D^*) > \Pr_0(D)$. More importantly, however, she would presumably also believe that the apparent color of the sock is relevant to the hypothesis that the lighting makes all socks look red, $\Pr_1(F) = \Pr_0(F \mid D^*) > \Pr_0(F)$. Consequently we would not have the consequence $\Pr_1(F) = \Pr_0(F \mid D^*) = \Pr_0(F)$.

[2] The conjunction of constraints (0-1) and (3) is logically strictly stronger than its consequence $\Pr_0(F) = \Pr_1(F)$. The latter equation can be true if constraints (0-1) or (3) are false, because the various effects on Sophia's belief in F may "cancel out." For instance, what happens between t_0 and t_1 may affect Sophia's belief in the four propositions $X \cap F$, $X \cap \overline{F}$, $\overline{X} \cap F$, $\overline{X} \cap \overline{F}$,for some proposition X, even though Sophia's degrees of belief in $X \cap F$ and $\overline{X} \cap F$ may sum to the same number at t_0 and at t_1. It is the conjunction of constraints (0-1) and (3), not its consequence, that says that, at t_1, Sophia thinks that F is not a potential underminer for what has happened between t_0 and t_1.

Now in Weisberg's example there is no appearance proposition D^* that Sophia becomes certain of between t_0 and t_1. Instead of an application of strict conditionalization to the appearance proposition D^* we have an application of Jeffrey conditionalization to the evidential partition $\{D, \overline{D}\}$ that is caused by some visual experience d^*. Instead of the appearance proposition D^* that Sophia becomes certain of, the doxastically relevant effects of the visual experience d^* are now described more indirectly: they correspond to the differences between Sophia's degrees of belief in D and $\overline{D}$ at t_0 and her degrees of belief in D and $\overline{D}$ at t_1. The visual experience d^* is non-propositional evidence that is reflected in the different shapes of Sophia's two degree of belief functions at t_0 and at t_1 on the evidential partition $\{D, \overline{D}\}$. Most importantly, in contrast to the above case where Sophia becomes certain of an *appearance* proposition, and thus learns something about what drives the change in her degrees of belief, constraint (1b) *ex*cludes information about what drives the change in Sophia's degrees of belief in D and $\overline{D}$. For all Sophia believes, the change in her degrees of belief in D and $\overline{D}$ may be due to a lapse of rationality.

Given these consequences it becomes clear that, in the presence of constraints (0-1) and (3) – in particular, constraint (1b) – constraint (2) is utterly implausible. If Sophia does not believe that the change in her degrees of belief in D and $\overline{D}$ between t_0 and t_1 has been caused by her vision, why should she believe that F can undo this doxastically relevant effect of what has happened between t_0 an t_1? The way the story is told makes it clear to *us* that it was her vision that caused her increase in confidence in D. However, as long as this information is not also made available to *her*,[3] there is no reason whatsoever to assume that her becoming certain of F should have any effect at all on her degree of belief in D.

To be sure, there are potential underminers G for Sophia's visual experience that takes place between t_0 and t_1, and constraint (2) can be satisfied. However, if constraint (2) is satisfied for some potential underminer G, then this potential underminer G must violate constraint (3) or constraint (1). Constraint (3) is violated by G if G is not independent of D according to $\Pr_0$. Constraint (1) is violated by G if Jeffrey conditionalization is applied to an evidential partition $\{E_i: i \in I\}$ whose members are not all logically independent of the potential

3 One way of making this information available to Sophia is by using Shafer (1985)'s notion of a *protocol*. Halpern (2003, ch. 6) uses protocols to solve Freund (1965)'s puzzle. Spohn (2012, sct. 9.3) uses protocols to solve the puzzle of the three prisoners (Mosteller 1965, problem 13) that is also known as the Monty Hall problem. Halpern (ms) uses protocols to solve the Doomsday Argument and the Sleeping Beauty Problem. A different proposal is discussed in the Appendix.

underminer G, and which must thus be more fine-grained than $\{D, \overline{D}\}$, say, $\{D \cap F, \overline{D} \cap F, D \cap \overline{F}, \overline{D} \cap \overline{F}\}$.

If constraint (3) is violated by G, but constraint (1) holds, then G is indirectly affected by the visual experience that takes place between t_0 and t_1, and Jeffrey conditionalization governs this indirect way of being affected. In this case G figures in the output of Jeffrey conditionalization, which tells Sophia, among other things, what to believe about G. If constraint (1) is violated by G, then G is directly affected by the visual experience that takes place between t_0 and t_1, and Jeffrey conditionalization has to be applied to an evidential partition some of whose members are logically dependent on G. In this case G figures in the input of Jeffrey conditionalization, and Sophia has to specify her new degrees of belief for all these members, including those that logically depend on G, before Jeffrey conditionalization can be applied.

Either way there is no problem for Jeffrey conditionalization. We just have to realize that, while a change of Sophia's degrees of belief from t_0 to t_1 can be undermined, it can only be undermined by a hypothesis G that is doxastically affected by what drives the former change. This can happen indirectly by G not being probabilistically independent of the members of the evidential partition and thus violating the allegedly plausible stipulation (3). Or it can happen directly by G not being logically independent of all members of the evidential partition to which Jeffrey conditionalization is applied. Which is determined by experience, not by methodology.

For these reasons I conclude that Weisberg's paradox may affect the applicability of Jeffrey conditionalization, but not its validity. Weisberg's paradox does not undermine Jeffrey's rule of conditionalization: for true conditionalizers Weisberg's paradox is a false alarm. Parallel considerations show that perceptual undermining does not pose a problem for ranking theory or Dempster-Shafer theory either.

3. Appendix

One way of making the information that it was her vision that caused her increase in confidence in D available to Sophia is by using Shafer (1985)'s notion of a *protocol*. Protocols have proved to be a powerful tool in solving paradoxes. For instance, Halpern (2003, ch. 6) uses protocols to solve Freund (1965)'s puzzle. Spohn (2012, sct. 9.3) uses protocols to solve the puzzle of the three prisoners (Mosteller 1965, problem 13) that is also known as the Monty Hall problem. Halpern (ms) uses protocols to solve the Doomsday Argument and the Sleeping Beauty Problem. It is only natural that protocols can also be used to solve Weisberg's Paradox.

A different proposal for making the information that it was Sophia's vision that caused her increase in confidence in D available to Sophia is presented in Gallow (2014). Gallow (2014) first proposes the following update rule:

Update Rule 3 (Gallow Conditionalization I, Gallow 2014) If $Pr_0(\cdot):\mathcal{A} \to \mathbb{R}$ is the ideal doxastic agent's probability measure at time t_0, and between t_0 and t_1 she receives total evidence of the form $\{(T_i, E_i): i \in I\}$, where the $T_i \in \mathcal{A}$ form a partition of the underlying set of possible worlds W, and her probabilities are not directly affected in any other way such as forgetting etc., then her probability measure at time t_1 should be $Pr_1(\cdot):\mathcal{A} \to \mathbb{R}$,

$$Pr_1(\cdot) = \sum_{i \in I} Pr_0(\cdot \mid T_i \cap E_i) \cdot Pr_0(T_i).$$

The interpretation of a pair (T_i, E_i) is that the ideal doxastic agent's evidence is E_i, provided T_i is the case. For instance, Sophia's total evidence may be that the sock really is red, provided the lighting does not make all socks look red and everything else is normal as well; and nothing otherwise: $\{(N, E), (\neg N, W)\}$.

Gallow conditionalization I is an instance of Jeffrey conditionalization on the evidential partition $\{T_i \cap E_i, T_i \cap \overline{E_i}: i \in I\}$ with input parameters $Pr_1(T_i \cap E_i) = Pr_0(T)$ and $Pr_1(T_i \cap \overline{E_i}) = 0$. Unfortunately it does not allow the ideal doxastic agent to ever change her confidence in any of the theories T_i, i.e. $Pr_1(T_i) = Pr_0(T_i)$.

For this reason Gallow (2014) generalizes his first update rule to the following second update rule:

Update Rule 4 (Gallow Conditionalization II, Gallow 2014) *If* $Pr_0(\cdot):\mathcal{A} \to \mathbb{R}$ *is the ideal doxastic agent's probability measure at time* t_0*, and between* t_0 *and* t_1 *she receives total evidence of the form* $\{(T_i, E_i): i \in I\}$*, where the* $T_i \in \mathcal{A}$ *form a partition of the underlying set of possible worlds W, and her probabilities are not directly affected in any other way such as forgetting etc., then her probability measure at time* t_1 *should be* $Pr_1(\cdot):\mathcal{A} \to \mathbb{R}$,

$$Pr_1(\cdot) = \sum_{i \in I} Pr_0(\cdot \mid T_i \cap E_i) \cdot Pr_0(T_i) \cdot \Delta_i,$$

where Δ_i *is a non-negative real number representing the degree to which the total evidence* $\{(T_i, E_i): i \in I\}$ *(dis)confirms theory* T_i.

Contrary to what Gallow (2014, 21) claims his second update rule does not generalize Jeffrey conditionalization. Indeed, quite the opposite is the case. Gallow conditionalization II is also just an instance of Jeffrey conditionaliaztion on the evidential partition $\{T_i \cap E_i, T_i \cap \overline{E_i}: i \in I\}$ with input parameters $Pr_1(T_i \cap E_i) = Pr_0(T_i) \cdot \Delta_i$ and $Pr_1(T_i \cap \overline{E_i}) = 0$. (As Gallow (2014, 25ff) shows, the sum

$\sum_{i \in I} \mathrm{Pr}_0(T_i) \cdot \Delta_i = \sum_{i \in I} \mathrm{Pr}_1(T_i \cap E_i) + \mathrm{Pr}_1(T_i \cap \overline{E_i})$ equals 1, and so the constraints of Jeffrey conditionalization are satisfied.)[4]

References

Freund, John E. 1965. "Puzzle or paradox?" *American Statistician* 19: 29-44.

Gallow, Dmitri J. 2014. "How to Learn from Theory-Dependent Evidence; or Commmutativity and Holism: A Solution for Conditionalizers." *British Journal for the Philosophy of Science*.

Halpern, Joseph Y. 2003. *Reasoning About Uncertainty*. Cambridge, Massachusetts: MIT Press.

____. 2014. *The Role of the Protocol in Anthropic Reasoning*. Unpublished manuscript. https://www.cs.cornell.edu/home/halpern/papers/doomsday.pdf (accessed July 7, 2014).

Jeffrey, Richard C. 1983. *The Logic of Decision*, 2nd ed. Chicago: University of Chicago Press.

Mosteller, Frederick. 1965. *Fifty Challenging Problems in Probability with Solutions*. Reading: Addison-Wesley.

Shafer, Glenn. 1976. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.

____. 1985. "Conditional probability." *International Statistical Review* 53: 261-277.

Spohn, Wolfgang. 2012. *The Laws of Belief. Ranking Theory and its Philosophical Applications*. Oxford: Oxford University Press.

Weisberg, Jonathan. 2009. "Commutativity or Holism? A Dilemma for Conditionalizers." *British Journal for the Philosophy of Science* 60: 793-812.

____. 2014. "Updating, Undermining, and Independence." *British Journal for the Philosophy of Science*.

[4] I am grateful to Kevin Kuhl, Wolfgang Spohn, and Jonathan Weisberg for comments on earlier versions of this note. My research was supported by the Canadian SSHRC through its Insight program and by the Connaught Foundation through its New Researcher program.