Be modest: you're living on the edge

KEVIN DORST
University of Pittsburgh
kevindorst@pitt.edu

January 2021; forthcoming in *Analysis*.

Abstract

Many have claimed that whenever an investigation might provide evidence for a claim, it might also provide evidence against it. Similarly, many have claimed that your credence should never be on the edge of the range of credences that you think might be rational. Surprisingly, both of these principles imply that you cannot rationally be *modest*: you cannot be uncertain what the rational opinions are.

Do I like lentils?

I'm guessing you're not sure. Yet, given your evidence (philosopher; strange writer; mentioned lentils in an Analysis paper...), there are some opinions you ought to have now. And when (in a few pages) you get new evidence—and, let's suppose, you don't lose any old evidence—there will be some new opinions you ought to have then. Let 'P' and 'P^+' be definite descriptions for these current and future rational opinions.[1]

How—if at all—should your opinions about P and P^+ constrain your other opinions (say, about lentils)? Many have endorsed principles that prevent your opinions from being on the edge of the range of potentially-rational opinions (§1):

> (STEP AWAY) If $P(q) = t$ and $P(P(q) > t) > 0$, then $P(P(q) < t) > 0$.
> If you should have a given credence and should leave open that the current rational credence might be higher, then you should also leave open that it might be lower.
>
> (STAY AWAY) If $P(q) = t$ and $P(P^+(q) > t) > 0$, then $P(P^+(q) < t) > 0$.
> If you should have a given credence and should leave open that the *future* rational credence might be higher, then you should also leave open that it might be lower.

Surprisingly (§2), each principle implies that you cannot rationally be *modest*—you cannot be uncertain what the current rational opinions are. Precisely: for all q there is a t such that $P(P(q) = t) = 1$. (So if you're unsure how confident you should be that I like lentils, you're irrational.) Arguably, this implies they're both false.

[1] I'll model them with precise probabilities (White 2009; Joyce 2010; Schoenfield 2012), and assume intrapersonal uniqueness: at each time (given your evidence, standards of reasoning, etc.) there is uniquely rational credence function (cf. White 2005; Horowitz 2014b; Schoenfield 2014; Schultheis 2018).

1 Why think they're true?

Start with Stay Away. This principle follows from (but is much weaker than) rationalized versions of diachronic Reflection principles—principles which entail that the current rational credence equals the current rational estimate of the future rational credence.

To see this, let $\mathbb{E}(P^+(q))$ be the current rational expectation of the future rational credence in q. (Precisely, $\mathbb{E}(P^+(q)) := \sum_t P(P^+(q) = t) \cdot t$.) Diachronic, rational Reflection principles are those which entail that $P(q) = \mathbb{E}(P^+(q))$. Now note that if Stay Away fails, then there is an $\epsilon > 0$ such that $\mathbb{E}(P^+(q)) \;=\; P(P^+(q) = t) \cdot t + P(P^+(q) > t) \cdot (t + \epsilon) \;>\; t \;=\; P(q)$, so diachronic Reflection fails too. Thus authors who endorse such Reflection principles are committed to Stay Away (e.g. van Fraassen 1984; Gaifman 1988; Weisberg 2007; Briggs 2009; Mahtani 2017).

Moreover—and independently—several authors have argued that there should be a ban on 'no lose' investigations: investigations which have some chance of confirming a claim without a chance of disconfirming it.[2] (If you think asking might increase your confidence that I like lentils, you should also think that doing so might decrease it.) Stay Away looks to be a formalization of that seemingly-banal claim.

Yet that seemingly-banal claim has the surprising consequence that you must be immodest. That's the point of this paper.

It's a fair point. But is it an old one? Bernhard Salow has recently argued that we should generalize the ban on 'no lose' investigations to a ban on 'intentionally biased inquiries' (Salow 2018, §1.1)—i.e. inquiries that you should expect to provide evidence favoring q. And banning the latter, he claims, requires immodesty.

It's a clever argument. But there are ways to question it.

First, Salow's ban on 'biased inquiries' is much stronger than the ban on 'no lose' investigations formalized by Stay Away. It requires your expectation of the current rational credences to match your expectation of the future rational credences. (Where $\mathbb{E}(P(q)) := \sum_t P(P(q) = t) \cdot t$, that is the claim that $\mathbb{E}(P(q)) = \mathbb{E}(P^+(q))$.) He goes on to assume, in effect, that this will hold in general only if rationalized diachronic Reflection holds: $P(q) = \mathbb{E}(P^+(q))$ (Salow 2018, §3.2).

Second, Salow makes the controversial assumption that rational credences are are always recoverable by conditioning a known prior on propositional evidence (2018, §3.1). Some (e.g. Gallow 2019b, §1) have argued that this assumption is the culprit.

And third, it's controversial whether Salow has correctly formulated what it takes for an inquiry to be 'biased'. Gallow (2019b, §3) shows that we can avoid Salow's result using a causalist conception bias. Das (2020b, §1) formulates a notion of a biased inquiry as one which *guarantees* a rise in credence. And Dorst (2020a, §7) shows that some inquiries which are 'biased' in Salow's sense are such that you should expect the the future rational credences to be more accurate than your own on every proposition (cf. Levinstein 2019), and so perhaps should not be thought of as 'biased'—or, at least, should not be banned as irrational.

[2] See White 2006; Titelbaum 2010; Salow 2018. Gallow 2019b and Das 2020b discuss subtly different principles—the result of this paper does not directly apply to them.

Given these concerns, it's natural to wonder whether we can avoid Salow's immodest conclusion by weakening his ban on 'biased inquiries' back to one on 'no lose' investigations. My claim is that this will not help: using the weaker constraint of Stay Away—and making no assumptions about whether rational credences are determined via conditioning—we still end up with immodesty.

Now turn to Step Away. This is an example of a *bridging* principle connecting first- and higher-order rational opinions (Dorst 2020b, §1). It says that you cannot rationally have a credence that's on the edge of the range of credences you think might be rational. In other words, if you think 'my credence that he likes lentils is definitely not too high, and it might be too low', then you're irrational (Christensen 2010b, §1). Something like this has been defended or explored by a variety of theorists who have taken it to be a tenable bridging principle that allows rational modesty.[3] But it doesn't.

Similarly to above, Step Away follows from rationalized, *synchronic* versions of Reflection principles—principles which imply that you should match your credence to your estimate of the *current* rational credence: $P(q) = \mathbb{E}(P(q))$. It's now known that even the weakest such Reflection principles rule out immodesty (Samet 2000, §7.2, Dorst 2019, §2.3.1). But again, it's natural to wonder whether we can avoid this conclusion by weakening the principle to one that simply bans being on the edge of the potentially-rational range. My claim is that this will not help.

2 Why think they're false?

Step 1: Stay Away implies Step Away. *Step 2:* If modesty can be rational, it can be rational for someone who knows the structural constraints on rationality—like Step Away. *Step 3:* But (theorem:) if Step Away is true, then this is not possible.

Step 1: Stay Away *implies* Step Away.

Suppose Step Away fails: there's a case in which you should have credence t that I like lentils (q), leave open that the current rational credence might be higher, and be certain that it's no lower. Now take a version of this scenario in which it's certain that you'll get no lentil-relevant evidence—for example, all you'll learn is how a fair coin landed. Then $P(P(q) = P^+(q)) = 1$, so it follows that Stay Away fails too.

What if we restricted Stay Away to apply only in cases in which you *might* get evidence relevant to q? Still no good. For then we can take a case where Step Away fails at threshold $P(q) = t$ and where you might get relevant evidence about q, but where this will (i) only happen if $P(q) > t$ and (ii) it will not push the future rational credence below t. For example, suppose Step Away fails in this way: the rational credence is 0.5 that I like lentils, but you should leave open (only) that it's either 0.5 or 0.6. Then we can imagine a version of this case in which if the current rational credence is 0.5, you'll get no relevant evidence, but if it's 0.6, then the future rational credence will shift

[3]See Christensen (2007, §5), Christensen (2010b, §1), Sliwa and Horowitz (2015, §4.4), Kappel (2017, §1), Skipper et al. (2018, §2), Gallow (2019a, §4.1), Fraser (2020, §4). Also compare Step Away to the 'rationality dominance' principles in Schultheis (2018, §2) and Hawthorne and Isaacs (2020, §3).

to either 0.5 or 0.7. Then Stay Away fails even though you might get lentil-relevant evidence.[4]

Upshot: Stay Away is tenable only if Step Away is.

Step 2: If modesty can be rational, it can be rational for someone who knows the structural constraints on rationality.

We can distinguish *structural* from *substantive* constraints on rationality. The former are constraints—like probabilism and the Principal Principle (Lewis 1980)—that hold regardless of what evidence you have or what propositions we're applying them to. The latter are constraints that depend on the specifics: 'Given evidence E, you should be 60% confident that Kevin likes lentils', etc. On this classification, Step Away is clearly a structural constraint.

Plausibly, if modesty can be rational, then it can be rational for someone who knows the structural constraints on rationality—and hence knows Step Away. Two reasons.

First, it's widely accepted that structural principles do not pin down a uniquely rational credence function; thus given our assumption of (intrapersonal) uniqueness (footnote 1), there plausibly are substantive constraints that outrun the true structural constraints. As such, someone who knows the latter can still be unsure about the former, and hence still be unsure what opinions to have, given their evidence. (Study all the epistemology you like, you can still be unsure how confident you ought to be that I like lentils.) At the very least, this has been the working assumption in most of the literature (e.g. Christensen 2010b; Cresto 2012; Lasonen-Aarnio 2013, 2015; Horowitz 2014a; Sliwa and Horowitz 2015; Schoenfield 2016a; Roush 2016, 2017; Salow 2018; Gallow 2019b; Das 2020a,b).

Second, even if structural constraints *do* pin down a uniquely rational credence function for each body of evidence, this doesn't make modesty irrational. For, as many have argued, you can still be unsure what your evidence *is* (e.g. Williamson 2000, Ch. 9, Williamson 2008, Horowitz 2014a, §6, Lasonen-Aarnio 2015, Das 2020a, §4.1, Das 2020b, §1.2). And even if you know what credences are warranted by each body of evidence, uncertainty about *which* body of evidence you have can still induce modesty. (For example, in the model specified in footnote 4, you know all conditionals of the form 'If I'm at world w, the rational credence function is ___', and yet you still should be modest

[4] For aficionados, here's a toy model of such a case, using stochastic-matrix notation (Dorst 2020b, §2): row i column j gives the probability that the rational credence at world i assigns to being at world j. At time 1, the possible rational credences are these: $\begin{pmatrix} 0.2 & 0.3 & 0.075 & 0.175 & 0.125 & 0.125 \\ 0.2 & 0.3 & 0.075 & 0.175 & 0.125 & 0.125 \\ 0 & 0 & 0.15 & 0.35 & 0.25 & 0.25 \\ 0 & 0 & 0.15 & 0.35 & 0.25 & 0.25 \\ 0 & 0 & 0.15 & 0.35 & 0.25 & 0.25 \\ 0 & 0 & 0.15 & 0.35 & 0.25 & 0.25 \end{pmatrix}$. There are six worlds, and the ones where I like lentils are $q = \{w_1, w_4, w_6\}$, as indicated by the blue columns. At worlds 1 and 2, $P(q) = 0.5$ (sum across blue columns in the first and second rows); and at worlds 3–6, $P(q) = 0.6$. Now let the time-2 rational credences be these: $\begin{pmatrix} 0.2 & 0.3 & 0.075 & 0.175 & 0.125 & 0.125 \\ 0.2 & 0.3 & 0.075 & 0.175 & 0.125 & 0.125 \\ 0 & 0 & 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{pmatrix}$. Thus at worlds 1 and 2 the rational credences stay the same; at worlds 3 and 4, the rational credence in q rises from 0.6 to 0.7; and at worlds 5 and 6 it falls from 0.6 to 0.5. Thus at world 1, $P(P^+(q) \geq 0.5) = 1$ and $P(P^+(q) > 0.5) > 0$, but $P(q) = 0.5$: Stay Away fails.

since you should be unsure what world you're at.)

Step 3: But if Step Away *is true, anyone who knows the structural constraints on rationality cannot be modest.*

Letting *Immodest* be the claim that the rational credence function is certain that it's rational (for all q, t, if $P(q) = t$, then $P(P(q) = t) = 1$), we have:

Theorem. If Step Away holds and $P(\text{Step Away}) = 1$, then $P(\mathit{Immodest}) = 1$.

In other words: whenever it's rational to be (correctly) sure of Step Away, it's rational to be certain that you should be immodest. (Proof in the Appendix.)

If evidence is factive (Williamson 2000, Ch. 9) and (in finite cases) you should be sure of q only if your evidence entails it, then it follows immediately from this result that you should be immodest: from $P(\mathit{Immodest}) = 1$, we can infer *Immodest*.

Moreover, if being rationally sure that you ought to ϕ implies that it's permissible to ϕ, then again it follows from this result that you should be immodest. To see this, let '$\Box p$' mean 'it ought to be that p' (and $\Diamond p := \neg\Box\neg p$), and let C be your actual credence function. Then '$P(q) = t$' can be rewritten as '$\Box(C(q) = t)$'. So the conclusion of our Theorem—that $P(\mathit{Immodest}) = 1$—implies that you're rationally sure that your credence function ought to be immodest: $P(\Box(C \text{ is immodest})) = 1$. By the above assumption, it follows that it's *permissible* for your credence function to be immodest: $\Diamond(C \text{ is immodest})$. But given uniqueness, what's permissible for your credence function is required of it, so it follows that your credence function is required to be immodest: $\Box(C \text{ is immodest})$, i.e. P is immodest.

Of course, you might reject the factivity of evidence or the claim that being rationally sure that you must ϕ makes it permissible to ϕ. Even so, it follows directly from the Theorem that a rational person who knows Step Away cannot be unsure whether the credence function they have is rational. Precisely: if 'π' is a rigid designator for their credence function, then the Theorem implies that either $\pi(P = \pi) = 1$ or $\pi(P = \pi) = 0$; they cannot have any intermediate credence in the rationality of their credence function. After all, if they are rational and know Step Away, then $\pi(\text{Step Away}) = 1$ and π satisfies Step Away. If $\pi(P = \pi) < 1$, then $\mathit{Immodest} \Rightarrow [P \neq \pi]$, thus by the Theorem, $\pi(\mathit{Immodest}) = 1 \leq \pi(P \neq \pi)$, hence $\pi(P = \pi) = 0$. Thus either $\pi(P = \pi) = 1$ or $\pi(P = \pi) = 0$. But if modesty can be rational, surely being *intermediately* modest—having a credence strictly between 0 and 1 in your own rationality—can be rational.

Thus, I claim, Stay Away and Step Away are not interestingly weaker than the strongest bridging principles, which simply rule out immodesty entirely (Dorst 2019; Skipper 2020). For by Step 1, if either principle is true, then Step Away is true; by Step 2, if Step Away is true and modesty can be rational, then it can be rational for someone who knows Step Away; but by Step 3, this *can't* be rational. So either rationality requires immodesty, or Stay Away and Step Away are false.

The surprising step in this argument is the Theorem. Why is it true? As with other results about the strength of bridging principles (e.g. Samet 2000; Williamson 2000, 2014, 2019; Elga 2013; Lasonen-Aarnio 2013, 2015; Dorst 2019, 2020a; Das 2020b), it's due to the fact that Step Away is an (implicitly) universally quantified claim, and thus that it's difficult to satisfy it for all propositions.

In particular note that Step Away straightforwardly entails:

Positive Access: If $P(q) = 1$, then $P(P(q) = 1) = 1$.
If you should be sure of q, then you should be sure that you should be.

For suppose that Positive Access fails: for some q, $P(q) = 1$, but $P(P(q) = 1) < 1$. Then you should have credence 1 in q, should leave open that maybe your credence should be lower, but should be certain that it shouldn't be higher—contradicting Step Away.

Once we see that Step Away implies Positive Access, it's not hard to see that it's unstable. For if you should have any higher-order uncertainty at all, then there will be some set of credences $t_1 < t_2 \cdots < t_n$, such that you should (1) leave open that any of them might be the rational credence in q, and (2) be certain that one of them is. By Positive Access, you should also be certain *that* you should be certain that one of them is rational—meaning that in the possibility where t_1 *is* the rational one, you should likewise be certain that one of the $t_1, ..., t_n$ is rational. But that means that in such a possibility, the rational credence (namely, t_1) is on the edge of the range of credences you should think might be rational—Step Away fails.

What should we make of this result? Since modesty (arguably) *is* rational[5], both Stay Away and Step Away are (arguably) false. That means that (arguably) no-lose investigations *are* possible, and that it *can* be rational to think 'My 60% credence that he likes lentils is definitely not too high, and it might be too low' and yet maintain your credence. At the very least, those who have appealed to bans on these possibilities should be aware that such they are more controversial than is standardly thought.[6]

P.S. Good guess—I *do* like lentils.

Appendix

Theorem. If Step Away holds and $P(\text{Step Away}) = 1$, then $P(\textit{Immodest}) = 1$.

Assume there is a finite number of (relevant) possibilities W, and that at each world w there is a unique probability function P_w (defined over the subsets of W) that captures the rational credences for you to have. Thus for any proposition (set of worlds) q and number t, a proposition like $[P(q) = t]$ is simply the set of worlds where this is so: $[P(q) = t] := \{w \in W |\ P_w(q) = t\}$. For example, $P_w(P(q) > t) = P_w(\{w' \in W | P_{w'}(q) > t\})$.

[5]See Williamson (2000, 2008, 2014); Christensen (2010a, 2020); Cresto (2012); Elga (2013); Lasonen-Aarnio (2013, 2015, 2019, 2020); Horowitz (2014a); Pettigrew and Titelbaum (2014); Sliwa and Horowitz (2015); Roush (2016, 2017); Schoenfield (2016b, 2017); Carr (2019a,b); Dorst (2019, 2020a,b); Fraser (2020).

[6]Many thanks to Rachel Fraser, Dmitri Gallow, Bernhard Salow, and two stellar referees for helpful discussion and feedback.

(Note: whereas 'P' in the main text was a definite description whose value varied across worlds, 'P_w' is a rigid designator for a particular probability function associated with w.) This is a standard way of modeling higher-order probabilities from epistemic logic; see e.g. Gaifman (1988); Samet (2000); Williamson (2000, 2014, 2019); Lasonen-Aarnio (2013, 2015); Salow (2018, 2019); Dorst (2019, 2020a,b); Das (2020a,b).

Step Away *holds* at world w iff every instance of it is true there, i.e. iff for all q, t: if $P_w(q) = t$ and $P_w(P(q) > t) > 0$, then $P_w(P(q) < t) > 0$. Note that this is so iff for all q, t: if $P_w(q) = t$ and $P_w(P(q) < t) > 0$, then $P_w(P(q) > t) > 0$.

Now take an arbitrary world w and suppose both that Step Away holds there and that $P_w(\text{Step Away}) = 1$. To establish the Theorem, we show that $P_w(\textit{Immodest}) = 1$, i.e. that if $P_w(x) > 0$, then for all q, t: if $P_x(q) = t$ then $P_x(P(q) = t) = 1$.

First define the binary relation R between worlds such that for all $x, y \in W$: xRy iff $P_x(y) > 0$; and let $R_x := \{y|wRy\}$. We build up to the result by establishing various properties that R must have, given our suppositions.

Lemma 1 (Shift-Transitivity) For all $x \in R_w$, if xRy then $R_y \subseteq R_x$.

Proof. Suppose $x \in R_w$, so $P_w(x) > 0$. By supposition, this means Step Away holds at x. Since xRy, we have $P_x(y) > 0$. Suppose for reductio that $R_y \not\subseteq R_x$, i.e. that there is a z such that $P_y(z) > 0$ and $P_x(z) = 0$. Then since $P_x(y) > 0$ and $y \in \{w'|P_{w'}(z) > 0\} = [P(z) > 0]$, we have that $P_x(P(z) > 0) > 0$, yet $P_x(z) = 0$ and of course $P_x(P(z) < 0) = 0$, violating Step Away at x. Contradiction. □

Lemma 2 (w-Transitivity) For all $x \in R_w$: $R_x \subseteq R_w$.

Proof. The reasoning is the same is in Lemma 1. If $P_w(x) > 0$ and $P_x(y) > 0$ yet $P_w(y) = 0$, then $P_w(P(y) > 0) > 0$ yet $P_w(P(y) < 0) = 0$, violating Step Away. □

Lemma 3 (Shift-Reflexivity) If wRx, then xRx.

Proof. For reductio, suppose not: there is some x such that wRx but $x\not{R}x$. Thus $P_w(x) > 0$ but $P_x(x) = 0$. Let m be some member of arg $\max\limits_{y \in \{w\} \cup R_w} \big(P_y(x)\big)$, so m is in $\{w\} \cup R_w$ and within that set has maximal probability for x. (Since W is finite, such a maximum exists.) Suppose this probability is $P_m(x) = t_m$. We show that Step Away fails at m.

We know $R_m \subseteq R_w$, for either $m = w$ (so it holds trivially) or $m \in R_w$ (so Lemma 2 establishes it). Since $P_m(x)$ is maximal within $\{w\} \cup R_w$, it follows that $P_m(x)$ is maximal within R_m, so that **(i)** $P_m(P(x) > t_m) = 0$. Moreover note that that **(ii)** $t_m = P_m(x) > 0$ since $P_w(x) > 0$ and by construction $P_m(x)$ is at least as high as $P_w(x)$. Finally, since $P_m(x) > 0$ and $P_x(x) = 0$, we have that **(iii)** $P_m(P(x) < t_m) > 0$. But (i)–(iii) imply that Step Away fails at m.

We know $m \in \{w\} \cup R_w$. If $m = w$, this contradicts our supposition that Step Away holds there; if $m \in R_w$ this contradicts our supposition that $P_w(\text{Step Away}) = 1$. □

Lemma 4 (Shift-Symmetry) If wRx, then if xRy, also yRx.

Proof. For reductio, suppose wRx and xRy, but $y \not R x$. Take a world m in $\arg\max_{v \in \{x\} \cup R_x} \big(P_v(x)\big)$. So m is in $\{x\} \cup R_x$ and within this set has maximal probability for x; suppose $P_m(x) = t_m$. We show that STEP AWAY fails at m.

We know $R_m \subseteq R_x$, for either $m = x$ (so it holds trivially) or $m \in R_x$ (so Lemma 1 establishes it). Since $P_m(x)$ is maximal within $\{x\} \cup R_x$, it follows that $P_m(x)$ is maximal within R_m, and so that **(i)** $P_m(P(x) > t_m) = 0$. Moreover, since by Lemma 3 xRx, so $P_x(x) > 0$, it follows that since $x \in \{x\} \cup R_x$, **(ii)** $t_m = P_m(x) > 0$.

Finally, since by supposition xRy, it follows that mRy—for since $m \in \{x\} \cup R_x$, either $m = x$ (in which case it follows trivially), or else $m \in R_x$, in which case $m \in R_w$ (by Lemma 2, since $x \in R_w$ and xRm), and so since mRx, by Lemma 1 it follows that since xRy, that mRy. This means that $P_m(y) > 0$. But since $y \not R x$, $P_y(x) = 0 < t_m$, so that **(iii)** $P_m(P(q) < t_m) > 0$.

Yet (i)–(iii) show that STEP AWAY fails at m. We know that $m \in R_w$ (since $m = x$ or $m \in R_x$), so this contradicts our supposition that $P_w(\text{STEP AWAY}) = 1$. □

Lemma 5 (Shift-Equivalence) If wRx, then for all $y \in R_x$ we have $R_y = R_x$.

Proof. Suppose wRx. By Lemma 1, we know that for all $y \in R_x$: **(i)** $R_y \subseteq R_x$. By Lemma 4, we know that for all $y \in R_x$, yRx. And since Lemma 2 tells us that $y \in R_w$, it follows from Lemma 1 that since yRx, that **(ii)** $R_x \subseteq R_y$. (i) and (ii) imply that $R_y = R_x$. □

Lemma 6 (Shift-Immodesty) $wRx \Rightarrow$ for all q, t: if $P_x(q) = t$ then $P_x(P(q) = t) = 1$.

Proof. For reductio, suppose not: wRx and there is a q, t such that $P_x(q) = t$ but $P_x(P(q) = t) < 1$. By Lemma 3, $P_x(x) > 0$, so that $P(P(q) = t) > 0$. It follows that there is a set of values $S = \{s \in \mathbb{R} |\ P_x(P(q) = s) > 0\}$ and that $|S| \geq 2$, since t is one such value but since $P_x(P(q) = t) < 1$, it is not the only one. Since W is finite, so is S. Now take the minimal value $s_0 \in S$, and a world y such that xRy and $P_y(q) = s_0$. By Lemma 5, $R_y = R_x$, and therefore $\{s \in \mathbb{R} | P_y(P(q) = s) > 0\} = S$. Thus by construction $P_y(q) = s_0$ while $P_y(P(q) > s_0) > 0$ and $P_y(P(q) < s_0) = 0$, so STEP AWAY fails at y. Since xRy and wRx, it follows by Lemma 2 that wRy, contradicting our supposition that $P_w(\text{STEP AWAY}) = 1$. □

It follows immediately from Lemma 6 that the THEOREM holds.

References

Briggs, R., 2009. 'Distorted Reflection'. *Philosophical Review*, 118(1):59–85.

Carr, Jennifer Rose, 2019a. 'A modesty proposal'. *Synthese*, 1–21.

———, 2019b. 'Imprecise Evidence without Imprecise Credences'. *Philosophical Studies*, To appear.

Christensen, David, 2007. 'Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals'. In Tamar Szabó Gendler and John Hawthorne, eds., *Oxford Studies in Epistemology*, volume 2, 3–31. Oxford University Press.

———, 2010a. 'Higher-Order Evidence'. *Philosophy and Phenomenological Research*, 81(1):185–215.

———, 2010b. 'Rational Reflection'. *Philosophical Perspectives*, 24:121–140.

———, 2020. 'Akratic (epistemic) modesty'. *Philosophical Studies*, To appear(September).

Cresto, Eleonora, 2012. 'A Defense of Temperate Epistemic Transparency'. *Journal of Philosophical Logic*, 41(6):923–955.

Das, Nilanjan, 2020a. 'Externalism and Exploitability'. *Philosophy and Phenomenological Research*, To Appear.

———, 2020b. 'The Value of Biased Information'. *The British Journal for the Philosophy of Science*, To Appear.

Dorst, Kevin, 2019. 'Higher-Order Uncertainty'. In Mattias Skipper Rasmussen and Asbjørn Steglich-Petersen, eds., *Higher-Order Evidence: New Essays*, 35–61. Oxford University Press.

———, 2020a. 'Evidence: A Guide for the Uncertain'. *Philosophy and Phenomenological Research*, 100(3):586–632.

———, 2020b. 'Higher-Order Evidence'. In Maria Lasonen-Aarnio and Clayton Littlejohn, eds., *The Routledge Handbook for the Philosophy of Evidence*. Routledge.

Elga, Adam, 2013. 'The puzzle of the unmarked clock and the new rational reflection principle'. *Philosophical Studies*, 164(1):127–139.

Fraser, Rachel, 2020. 'Mushy Akrasia'.

Gaifman, Haim, 1988. 'A Theory of Higher Order Probabilities'. In Brian Skyrms and William L Harper, eds., *Causation, Chance, and Credence*, volume 1, 191–219. Kluwer.

Gallow, J. Dmitri, 2019a. 'A Subjectivist's Guide to Deterministic Chance'. *Synthese*, To Appear:1–34.

———, 2019b. 'Updating for externalists'. *Noûs*, (November 2018):1–30.

Hawthorne, John and Isaacs, Yoaav, 2020. 'Permissivism, Margin-for-Error, and Dominance'. *Philosophical Studies*, 1–18.

Horowitz, Sophie, 2014a. 'Epistemic Akrasia'. *Noûs*, 48(4):718–744.

———, 2014b. 'Immoderately rational'. *Philosophical Studies*, 167:41–56.

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

Kappel, Klemens, 2017. 'Bottom Up Justification, Asymmetric Epistemic Push, and the Fragility of Higher Order Justification'. *Episteme*, To appear.

Lasonen-Aarnio, Maria, 2013. 'Disagreement and evidential attenuation'. *Nous*, 47(4):767–794.

———, 2015. 'New Rational Reflection and Internalism about Rationality'. In Tamar Szabó Gendler and John Hawthorne, eds., *Oxford Studies in Epistemology*, volume 5, 145–171. Oxford University Press.

———, 2019. 'Higher-Order Defeat and Evincibility'. *Higher-Order Evidence: New Essays*, 144–171.

———, 2020. 'Dispositional Evaluations and Defeat'. 1–32.

Levinstein, B. A., 2019. 'Accuracy, Deference, and Chance'.

Lewis, David, 1980. 'A subjectivist's guide to objective chance'. In Richard C Jeffrey, ed., *Studies in Inductive Logic and Probability*, volume 2. University of California Press.

Mahtani, Anna, 2017. 'Deference, respect and intensionality'. *Philosophical Studies*, 174(1):163–183.

Pettigrew, Richard and Titelbaum, Michael G, 2014. 'Deference Done Right'. *Philosopher's Imprint*, 14(35):1–19.

Roush, Sherrilyn, 2016. 'Knowledge of Our Own Beliefs'. *Philosophy and Phenomenological Research*, 93(3).

———, 2017. 'Epistemic Self-Doubt'.

Salow, Bernhard, 2018. 'The Externalist's Guide to Fishing for Compliments'. *Mind*, 127(507):691–728.

———, 2019. 'Elusive Externalism'. *Mind*, 128(510):397–427.

Samet, Dov, 2000. 'Quantified Beliefs and Believed Quantities'. *Journal of Economic Theory*, 95(2):169–185.

Schoenfield, Miriam, 2012. 'Chilling out on epistemic rationality'. *Philosophical Studies*, 158(2):197–219.

———, 2014. 'Permission to Believe: Why Permissivism is True and What it Tells Us About Irrelevant Influences on Belief'. *Nous*, 48(2):193–218.

———, 2016a. 'An Accuracy Based Approach to Higher Order Evidence'. *Philosophy and Phenomenological Research*, To Appear.

———, 2016b. 'Bridging Rationality and Accuracy'. *Journal of Philosophy*, 112(12):633–657.

———, 2017. 'Conditionalization Does Not (In General) Maximize Expected Accuracy'. *Mind*, 126(504):1155–1187.

Schultheis, Ginger, 2018. 'Living on the Edge: Against Epistemic Permissivism'. *Mind*, 127(507):863–879.

Skipper, Mattias, 2020. 'Does Rationality Demand Higher-Order Certainty?' *Synthese*, To Appear.

Skipper, Mattias, Steglich-Petersen, Asbjørn, and Bjerring, Jens Christian, 2018. 'A higher-order approach to disagreement'. *Episteme*, 15(1):80–100.

Sliwa, Paulina and Horowitz, Sophie, 2015. 'Respecting *all* the evidence'. *Philosophical Studies*, 172(11):2835–2858.

Titelbaum, Michael G., 2010. 'Tell me you love me: Bootstrapping, externalism, and no-lose epistemology'. *Philosophical Studies*, 149(1):119–134.

van Fraassen, Bas, 1984. 'Belief and the Will'. *The Journal of Philosophy*, 81(5):235–256.

Weisberg, Jonathan, 2007. 'Conditionalization, reflection, and self-knowledge'. *Philosophical Studies*, 135(2):179–197.

White, Roger, 2005. 'Epistemic Permissiveness'. *Philosophical Perspectives*, 445–459.

———, 2006. 'Problems for Dogmatism'. *Philosophical Studies*, 131:525–557.

———, 2009. 'Evidential Symmetry and mushy credence'. *Oxford Studies in Epistemology*, 161–186.

Williamson, Timothy, 2000. *Knowledge and its Limits*. Oxford University Press.

———, 2008. 'Why Epistemology Cannot be Operationalized'. In Quentin Smith, ed., *Epistemology: New Essays*, 277–300. Oxford University Press.

———, 2014. 'Very Improbable Knowing'. *Erkenntnis*, 79(5):971–999.

———, 2019. 'Evidence of Evidence in Epistemic Logic'. In Mattias Skipper and Asbjørn Steglich-Petersen, eds., *Higher-Order Evidence: New Essays*, 265–297. Oxford University Press.