

# Christensen 2010: "Rational Reflection"

KEVIN DORST

January 12, 2015

## A Compelling Connection

Intuitive connection between (1) what it's rational to believe, and (2) what it's rational to believe one is rational to believe.

In graded contexts, there is (perhaps) more room to have rational doubts. E.g. Ava has .7 credence in (*D*) *the democrats will win* based on a whole mess of evidence.

Christensen: part of reason is that her higher-order doubts are *symmetric*. Compare with Brayden.

**Question:** Is the symmetry doing all the work? Does it matter that in the discrete case, either your *exactly* right or *far* wrong? Compare:

- Is rationally suspending judgment on *P* incompatible with rationally believing that one should *not* suspend judgment on *P*?
- What about: .5 credence in *P*, while .5 that one ought to be very confident in *P* and .5 that one ought to be very confident in  $\neg P$ ?

A natural candidate for the connection:

**RatRef:**  $Cr(A|Pr(A) = n) = n$

Even if RatRef is wrong, Brayden's **dominance-reasoning** seems immaculate.

E.g. rationally believing *P* is incompatible with rationally believing that one's belief in *P* is irrational. Likewise with justification.

She's confident that she hasn't evaluated it *exactly* right.

Symmetric, but far wrong.

"To violate it would be as if to say, for example, 'My own credence in rain, supposing the rational credence in rain in my very evidential situation is .9, is .6!'"

## A Puzzle

Chloe is looking at an unmarked clock, in fact pointing at 21. How should her credences be distributed? A tight, roughly bell-shaped curve around  $P_{21}$ .

... $P_{19} = 0.1$ ,  $P_{20} = 0.2$ ,  $P_{21} = 0.3$ ,  $P_{22} = 0.2$ ,  $P_{21} = 0.1$ ,...

Show her the chart. She will see that her confidence in  $P_{21}$  is definitely not too low, and probably too high; so, like Brayden, it seems she should lower it. Yet that seems inconsistent with the stipulation that the chart is ideal for her.

RatRef isn't necessary to get into these puzzles; but it draws it out (weighted average of *Pr*).

Options: (1) deny RatRef, (2) deny Chart, (3) say Chloe can't *know* chart.

Similarly for each other position of the hand. Write these all down in a big  $60 \times 60$ , **Chloe's Chart**.

RatRef is inconsistent with the scenario:  
 $Pr(P_{21}) = .3 \Leftrightarrow P_{21}$ ; so  
 $Cr(P_{21}|Pr(P_{21}) = 0.3) = Cr(P_{21}|P_{21}) = 1$ .

*Split-Level Strategy*

Say Chloe can learn chart, and not change her credences.

Problem: the rough connection guaranteed by RatRef seems very hard to deny:

**Hypoxia.** "It seems clear to me that you would be grossly irrational if you were to rely on your recent reasoning about having sufficient fuel to make the more distant airstrip." (126)

**Horse-racing.** I get .4 in (M) *Mr. Ed will win*. Oracle tells me I should have .7. And this holds even if she tells you a range (.6 to .8).

"Strongly suggests that the Split-Level strategy is not an attractive option in general for theorizing about rational credences: evidence that supports certain sorts of doubts about the rationality of our credences can exert rational pressure to modify those credences." (129)

RatRef is misguided just like original Reflection is.

*Practical* irrationality ensues as well.

Not to say we need RatRef in particular; but any principle that supports her **dominance-style reasoning** will do.

*Rejecting the Values on the Chart*

Reject the chart we made? What would the real chart look like?

Puzzle generates from (1) symmetry across various positions around clock, and (2) in each evidential situation the true proposition gets higher but non-maximal credence.

(1) is a harmless idealization – it certainly seems to hold for all the relevant values around 21, say.

(2): Even if implausibly require  $\frac{1}{3}$  in each of  $P_{20}, P_{21}$ , and  $P_{22}$ , the problem would arise for  $P_{20}$ . And making  $Pr(P_{21}) = 1$  seems very implausible, since we can make the jumps smaller.

**Stalnakerian Response:** We can reject (2) in a principled way.

Your evidence is set by (i) your best guess of the location, plus (ii) your margin for error. If margin is 5 ticks, and you guess 45, then so long as it's between 40-50, you know it. (Even if it's exactly at 40.) If you guess 50, and it's in fact at 40, then your MoE must be larger, so your evidence is that it is between 30-50.

**Q:** Is each point equally likely?

*Does Chloe's seeing the chart falsify it?*

You might think telling her accurate info about her epistemic state shouldn't have this effect, but it can.

But the problem, says Christensen, is that Chloe can just have knowledge of her visual system, the clock setup, and externalism before entering the room, and make the chart herself!

E.g. if I evaluate M at .4, and Oracle tells me it is *at least* .4.

**Question:** Not as clear on this. Presumably she won't be *certain* that she's made the right chart, so she won't be able to use the compelling "definitely not too low, probably too high" dominance reasoning.

### *Saved by Anti-Luminosity?*

RatRef is one precise way to draw out the puzzle, but the only argument that really seems undeniable is when one can think through the dominance reasoning: "I *know* my credence is not too low, and it's probably too high."

But *this* reasoning only needs a bit of uncertainty to debunk.

If Chloe can't know precisely what her credences are, then she can't perform it.

– Independent reasons to reject credence luminosity.

Or if Chloe can't know precisely what the correct chart is, she can't perform the dominance reasoning.

She may know the highest value of the chart is .3, and actually have .3; but if she doesn't know this, she's safe.

Williamson; *Chris Dorst*.

### *Upshot?*

Christensen's line: "in certain cases, agents such as Chloe have no choice but to violate some perfectly respectable epistemic ideal" (136). Though that's not to say that there isn't still a best option.

**Question:** If he doesn't go all the way to committing to rational dilemmas, does this even help? Can't we run all the above issues with "all-things-considered theoretical reason"?

Precedent: logical truths; anti-expertise paradox.