The Dogmatism Paradox

Belief and its Limits (BaiL); Seminar 1 ERG
Instructor: Joshua Pearson; joshuaedwardpearson@gmail.com 30/9/25

I. Opening Task on Preservation

Try to think of an example that fits the following structure:[1]

- You believe claim q. (And justifiably so.)
- p is a live possibility for you—you think it may be that p. (And justifiably so.)
- Were you to learn p—and nothing more than p—you'd no longer be justified in believing q.

If you can: what's the example? If not: what's making it difficult?

In fact, an important principle says there is no such p and q:[2]

(**Preservation**) If p is consistent with everything you are justified in believing, then you are justified in retaining all those beliefs upon discovering p.

Explanation: If learning p somehow means you won't be justified in believing q, that must be because some p-possibility that's live for you is a not-q-possibility. So it seems like you don't believe q.

Slogan: your beliefs can only be defeated by learning information you currently consider false.[3]

Interlude on the Belief-Credence Connection

'S believes p' seems to ascribe S a binary attitude (either S believes p or she doesn't). Many epistemologists prefer to work with a *graded* notion of belief—'credence'—measured on a scale between 0 and 1. To assign S a high/low credence in p is to say something like 'S is highly confident that p'/'S is not very confident that p'.[4]

How do beliefs and credences relate?[5] Two influential views:

(**Belief=credence1**) You're justified in believing p iff the rational credence for you to assign p is 1.[6]

(**Lockeanism**) You're justified in believing p iff the rational credence for you to assign p is sufficiently high (but < 1).[7]

There's an important connection between Preservation and these two views. Preservation follows from **Belief=credence1**. However, many find this view too extreme, and instead endorse **Lockeansim**. **Lockeanism**, however, does not validate Preservation.[8]

[1] Just use your intuitive sense of "believe", whatever that may be!

[2] See, especially, Leitgeb (2017). The principle first comes from "AGM belief revision theory" Alchourrón et al. (1985)—but don't worry too much about that!

[3] I sometimes only say "belief" when I mean "justified belief"—my main focus here is on the normative side, rather than the metaphysical/phil-mind side.

[4] Though, even this is recently controversial. See Williamson (fc) and Goodman (ms).

[5] There's both a *metaphysical* and a *normative* question about their relationship. I'm focusing on the latter here, but the former is also important.

[6] A perhaps surprisingly large amount of notable recent defenders. See especially Clarke (2013) and Greco (2015).

[7] A key defender is Foley, see e.g. Foley (2009).

[8] See Leitgeb (2014, 2017), who shows that endorsing Preservation is at least *consistent* with at least some forms of Lockeanism.

II. The "First" Dogmatism Paradox

The Last Oreo. It's teatime. You see Al slip the last of a limited supply of Oreos in his pocket. So, you know Al took the last Oreo (*q*). Later, you decide to go to his office to confront him. However, Al then proceeds to claim—citing highly compelling evidence (*p*)—that he has an evil twin brother, Fal, and Fal took the last Oreo.[9]

[9] A version of the library book case Harman (1973) uses.

Two Marbles. You see me place a green and a blue marble in an otherwise empty, opaque bag. Hence, you know the bag contains a blue marble (*q*). You then proceed to repeatedly draw a marble from the bag, inspect its color, and then place it back in the bag. You find that after 100 draws every marble you've drawn has been green (*p*).[10]

[10] A famous case Williamson (2000) uses.

These are both cases in which, intuitively, you start out knowing, and so justifiably believing *q*, but on learning *p*, you should stop believing *q*.[11]

[11] Fill out the cases however you need to make this judgment plausible—the evidence Al cites, or the number of times you draw a green marble.

The "first" dogmatism paradox,[12]presents a seemingly compelling argument that you should *not* drop your belief in these cases:

[12] Harman (1973), who based it from a lecture Kripke gave, published later in Kripke (2011)

(1) Initially, you know *q*.

(2) If you initially know *q*, and you are presented with evidence *p* against *q*, you can know by deduction that *p* is evidence against a truth.[13]

[13] *q* entails *q is true*, which entails *Evidence against q is evidence against a truth*; so assuming you know *p is evidence against q* you can know *p is evidence against a truth*.

(3) If you know *p* is evidence against a truth, *q*, then you should ignore any evidential impact *p* has on whether *q*.[14]

[14] For then *p* is in a sense *misleading* evidence.

(4) You should ignore *p*'s evidential impact on whether *q*.

Initial thoughts on what's wrong with the argument?

Harman (1973) argues against (2). Once you are presented with compelling evidence *p*, your knowledge that *q* is "defeated"—you no longer know that Al took the last Oreo, or that there's a blue marble in the bag—meaning you cannot know by deduction that evidence against *q* is misleading...problem solved?[15]

[15] There's a lot of recent work on this paradox that finds Harman's solution unsatisfactory. We'll look at Bernhard Salow's paper—a recent example that ties in well with this class. But see also Lasonen-Aarnio (2013), Ye (2016), Beddor (2019) and Fraser (fc).

III. The "Second" Dogmatism Paradox

Harman's solution explains why we shouldn't ignore the evidence *once we receive it*. But it doesn't explain why we shouldn't take precautions in *avoiding* the evidence altogether. Consider:

Belief Pill. You've just learned that *q*. You then discover that scientists have invented a "belief pill"—a pill that guarantees you'll continue to believe *q* even if you encounter evidence against it.

The "second" dogmatism paradox asks, if you are only concerned with being accurate as to whether *q*, *why not take the pill?*[16] Thoughts?

[16] According to (Kripke, 2011, Appendix C), this is the version of the paradox he initially intended when he first gave his lecture on it. Hawthorne (2003) also discusses this version of the paradox.

Here's a rough reconstruction of the second argument, broadly following Salow..[17] Let *A* and *B* denote a choice between finding out whether *p*, *A*, or avoiding doing so (i.e. being dogmatic), *B*—say between asking Al whether he stole the last Oreo or not, or whether to start drawing marbles.

[17] See (Salow, fc, §2). Salow switches from talking about initial knowledge to initial justified belief. He argues later in the paper that what he says for justified belief can equally go for knowledge, but that requires some subtleties he'd rather postpone.

(1) You believe *A* might result in losing your true belief that *q*.

(2) You believe *B* won't result in losing your true belief that *q*.

(3) You care only about being accurate as to whether *q*.

(4) So, you should choose *B* over *A*.

(4) follows from (1), (2) and (3) if we accept:

(**Weak Dominance**) If you justifiably believe that *B*'s consequence would be at least as good as *A*s, and that they might be better, you should choose *B* over *A*.

And we can just stipulate that (3) is true.[18] So whether this argument works depends on if we can find a case with a decision between *A* and *B* such that (1) and (2) are true.

[18] Even though it is weird to only care about *q*, it's still extremely paradoxical that, if you care about being accurate as to whether *q*, you should avoid evidence that bears on it.

Salow observes that this seems to depend on whether Preservation is true. To see this, consider an equivalent statement of Preservation:

(**Confidence**) If you would be required to give up your justified belief in *q* upon discovering *p*, then you are justified in believing not-*p*.

If **Confidence**/**Preservation** is right, justifiably believing *q* means justifiably believing that no action *A* will lead you to evidence forcing you to give up your belief in *q*. So (1) cannot be true. So, perhaps the key to avoiding dogmatism is just to endorse **Preservation**...?[19]

> Much like the proverbial schoolyard bully, the dogmatist behaves poorly not because he is too self-assured, but because he is, at heart, not self-assured enough. If only he stood properly behind his belief that *q*—by maintaining that *p*, being strong evidence against *q*, is false and thus not something he would discover—he would not act as he does. (Salow, fc, p. 28)

[19] Importantly, the two examples above — **The Last Oreo** and **Two Marbles** —- aren't obviously counterexamples to **Preservation**. Of course, Al's convincing evidence should convince you he didn't steal the Oreo, but it's not as though you initially thought Al might present you with such convincing evidence. If you *did*, you shouldn't have believed he stole it. Likewise, drawing a green marble 100 times in a row should convince you that there's no blue marble in there—but drawing a green marble 100 times in a row would be an extremely remarkable event that you're presumably initially justified in ruling out.

IV. Counterexamples to Preservation

Salow isn't done yet: he's worried Preservation may be false. Here are three purported counterexamples.

> **Flipping for Heads.** You're about to flip a coin—which you know is fair—until it lands on heads. You flip it 8 time and it lands tails every time.[20]

[20] Originally from Dorr et al. (2014), but see Goodman and Salow (2018), Smith (2018), Goodman and Salow (ming) and Pearson (2025) for extended discussion.

Explanation: You're initially justified in believing it won't take very long to get a heads—let's say you can justifiably believe you'll get a heads before 12 flips. But you should leave it open that you'll get 8 tails in a row. However, presumably, on seeing it land 8 tails in a row, you should drop your belief that you'll get a heads before 12 flips.

> **Suitcase.** You pick up your suitcase, and it strikes you as weighing about 15kg. You then learn it weights at least 17.5kg.[21]

[21] See e.g. Williamson (2014) and Goodman and Salow (2023) for extended discussion of these kinds of examples—sometimes referred to as cases of "inexact knowledge".

Explanation: Idealizing, your justified in believing your suitcase has a weight within some interval, $[(15-x)kg, (15+x)kg]$. Let's suppose $x = 3$. Upon learning it weights at least $17.5kg$, it's not as though you should conclude it weights between 17.5 and 18kg; rather, you should now leave it open that it weighs more than 18kg.

> **Composers v.1.** You justifiably believe (at time t_0) that Verdi is Italian, Bizet is French and Satie is French. These three beliefs are formed on equally strong, independent bases. At t_1, you learn that Verdi and Bizet are compatriots. At t_2, you learn that all three composers are compatriots.[22]

[22] Originally from Stalnaker (1994); see my Pearson (ms) for extended discussion.

Explanation: At t_2, you should be ambivalent as to whether they are all French or all Italian—maybe you were wrong about Verdi, but you might have been wrong abut both Bizet and Satie. At t_1, you should be ambivalent as to whether Verdi and Bizet are both French or Italian, but to believe Satie is French—you haven't been learned anything him. But then t_1 to t_2 involves a failure of Preservation: at t_1, you believe Satie is French, and it's compatible with your beliefs that they're all compatriots (they might all be French); yet at t_2 you must drop your belief Satie is French.

If any of these examples work, dogmatism returns:

- Consider **Suitcase**. You are deciding between whether to A — obtain evidence settling whether the suitcase is greater than 17.5kg — or B — avoid it. You care only about the accuracy of your belief that the suitcase weights between 12kg at 18kg.
- You believe B is no worse than A: you'll preserve true belief.
- But you also believe B might be better: if you choose A and learn that suitcase weighs at least 17.5kg, you'll give up your belief that it weighs between 12kg and 18kg, and thereby drop your true belief. So **Weak Dominance** says to choose the dogmatic B.

So, either something is up with these counterexamples, or we need to find a different resolution to the paradox...

VI. Salow's Anti-Preservation solution

Salow remains neutral on **Preservation** in this paper.[23] So, he wants

[23] He argues against it extensively in other work, which some of my research is inspired by; e.g. Goodman and Salow (2023) and Goodman and Salow (ming).

a solution even if **Preservation** is false. He suggests that, in this case, we reject **Weak Dominance** for:

(Conditional Dominance (CD)) If, conditional on *A* and *B* not being equally good, you believe *B* is better than *A*, choose *B*.

To understand this principle, it's easiest to split it into two cases:

- Case 1 — you believe *A* are *B* are not equally good. Then CD says, if you believe *B* is better than *A*, choose *B*.
- Case 2 — you do not believe *A* and *B* are not equally good. Then CD says, if on supposing that *A* and *B* aren't equally as good, you believe *B* is better, choose *B*.

An example illustrating why this is plausible: "You might outright believe that Hume died before 1800, and hence that a $100 bet that Hume died before 1800 will net the same result as an unconditional pay-out of $100; still, you should choose the latter."[24]

[24] **Weak Dominance** is silent here.

Let's now return to **Suitcase**, to see whether the dogmatic choice (*B*) is licensed:

- In **Suitcase—**if it's a valid counterexample to **Preservation—**you do not believe the suitcase weighs no more than 18kg conditional on weighing at least 17.5kg.
- Now, consider the choice between *A* — obtaining evidence settling whether the suitcase is greater than 17.5kg — or *B* — avoiding such evidence. If the suitcase is less than 17.5kg, *A* and *B* are equally as good. If the suitcase is at least 17.5kg, which outcome is better depends on whether the suitcase is no more than 18kg.[25]
- **Conditional Dominance** recommends the option that, conditional on them not being equally as good, you believe to be better. But this condition is equivalent to the suitcase being greater than 17.5kg, and conditional on that, you're unsure which of *A* or *B* is better. So **Conditional Dominance** does not support the dogmatic conclusion.[26]

[25] If it is no more than 18kg, *B* is better, as then you don't lose a true belief; if it is more than 18kg, *A* is better, as then you'll drop a false belief.

[26] Instead, it remains silent.

...Are we convinced?

VII. Anticipation

Here's an even weaker principle than Preservation:

(**Anticipation**) If you justifiably believe *q*, it cannot be both that (i) learning *p* would defeat your belief in *q*; and (ii) learning not-*p* would defeat your belief in *q*.

If this principle fails, there's an even more powerful argument for dogmatism. Let A and B denote the choice concerning whether to find out whether p.

(1) You believe A will result in losing your true belief in q.

(2) You believe B won't result in losing your true belief in q.

(3) You care only about your accuracy as to whether q.

(4) So, you should choose B.

All this needs is an even weaker principle of decision:

(**Strong Dominance**) If you justifiably believe B's consequences would be better than As, you should choose B over A.

...Why is this a problem? Isn't Anticipation obviously true?

Well, no. The very same counterexamples to Preservation may be extendable to Anticipation:[27]

[27] See Goodman and Salow (ming), Pearson (2025) and Pearson (ms)

> **Flipping for All Heads** You're flipping 100 coins simultaneously, until all of them land on heads on the same simultaneous flip. You believe both that this won't happen on the first few simultaneous flips, but also that, for a specific n, it will happen before n simultaneous flips.

> **Composers v.2** Same as Composers v.1, except that at t_2, you learn that Satie has a *different* nationality from Veri and Bizet.

I think this puts Salow into a bit of a bind. If these counterexamples are not convincing, it places doubt on the original counterexamples to Preservation. If they are convincing, then it looks like we have to also deny the very plausible **Strong Dominance** to avoid dogmatism.

I am tempted by a different solution. I suspect that Salow is illicitly shifting between two things:

(i) In some possibility consistent with your beliefs, A produces a worse outcome than B

(ii) You believe that A might have worse consequences than B. That is, you believe it might be that, were you to A, the consequences would be worse than were you to B.

For example, it looks consistent to me to, in **Flipping for Heads**: believe the coin won't be flipped more than 12 times, leave it open that it's flipped 8 times—in which case I'd lose my belief that it won't be flipped more than 12 times—but to also fail to believe the counterfactual: *Were I to see the coin flipped 8 times, I would lose a true belief*. Rather, you're unsure whether your belief is true in that counterfactual scenario.[28]

[28] I think this means the later two arguments for dogmatism are just invalid; the dominance principles don't allow you to infer (4) from (1), (2) and (3).

References

Alchourrón, C. E., Gärdenfors, P., and Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530.

Beddor, B. (2019). The toxin and the dogmatist. *Australasian Journal of Philosophy*, 97(4):727–740.

Clarke, R. (2013). Belief is credence one (in context). *Philosophers' Imprint*, 13:1–18.

Dorr, C., Goodman, J., and Hawthorne, J. (2014). Knowing against the odds. *Philosophical Studies*, 170(2):277–287.

Foley, R. (2009). Beliefs, degrees of belief, and the lockean thesis. In Huber, F. and Schmidt-Petri, C., editors, *Degrees of belief*, pages 37–47. Springer.

Fraser, R. (fc). The will in belief. *Oxford Studies in Epistemology*.

Goodman, J. (ms). Degrees of confidence are not subjective probabilities. Manuscript.

Goodman, J. and Salow, B. (2018). Taking a chance on kk. *Philosophical Studies*, 175(1):183–196.

Goodman, J. and Salow, B. (2023). Epistemology normalized. *Philosophical Review*.

Goodman, J. and Salow, B. (forthcoming). Belief revision normalized. *Journal of Philosophical Logic*.

Greco, D. (2015). How i learned to stop worrying and love probability 1. *Philosophical Perspectives*, 29(1):179–201.

Harman, G. (1973). *Thought*. Princeton University Press.

Hawthorne, J. (2003). *Knowledge and Lotteries*. Oxford University Press.

Kripke, S. A. (2011). On two paradoxes of knowledge. In Kripke, S. A., editor, *Philosophical Troubles. Collected Papers Vol I*. Oxford University Press.

Lasonen-Aarnio, M. (2013). The dogmatism puzzle. *Australasian Journal of Philosophy*, (3):1–16.

Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, 123(2):131–171.

Leitgeb, H. (2017). *The Stability of Belief: How Rational Belief Coheres with Probability*. Oxford University Press.

Pearson, J. E. (2025). Belief revision revised. *Philosophy and Phenomenological Research*.

Pearson, J. E. (ms). Composing composers. Manuscript - `joshuaedwardpearson.com`.

Salow, B. (fc). Fallibility and dogmatism. *Australasian Journal of Philosophy*.

Smith, M. (2018). Coin trials. *Canadian Journal of Philosophy*, 48(5):726–741.

Stalnaker, R. (1994). What is a nonmonotonic consequence relation? *Fundamenta Informaticae*, 21(4):7–22.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.

Williamson, T. (2014). Very improbable knowing. *Erkenntnis*, 79(5):971–999.

Williamson, T. (fc). Knowledge, credence, and the strength of belief. In Reed, B. and Flowerree, A., editors, *Towards an Expansive Epistemology: Norms, Action, and the Social Sphere*. Routledge.

Ye, R. (2016). Misleading evidence and the dogmatism puzzle. *Australasian Journal of Philosophy*, 94(3):563–575.