(Almost) All Evidence is Higher-Order Evidence

Brian Hedden* Australian National University

Kevin Dorst* University of Pittsburgh

October 2021

Abstract

Higher-order evidence is evidence about what's rational to think in light of your evidence. Many have argued that it's special—falling into its own evidential category, or leading to deviations from standard rational norms. But it's not. Given standard assumptions, almost all evidence is (in part) higher-order evidence.

Higher-order evidence is in vogue. But what it *is*—let alone what its effects are—is hotly contested.

Two classes of positions are relevant here. First, some have claimed or presupposed that your evidence about a given proposition can be "factored" into two different bodies of evidence—your first-order evidence, and your higher-order evidence.[1] Second, some have argued that in cases involving higher-order evidence, standard epistemological norms (like probabilism and conditionalization) break down.[2]

Here we make a simple point that casts doubt on both of these positions: given natural assumptions, almost all evidence is (in part) higher-order evidence. More precisely: given that there's a correlation between evidence and truth, and given the symmetry of probabilistic relevance, almost all evidence should affect your beliefs about what opinions were supported by your (initial) evidence. This suggests both that there can be no clean divide between first- and higher-order evidence, and that those who reject standard norms in cases of higher-order evidence must say—implausibly—that those norms fail in almost every case.

We first give our argument (§1), then go on to defend each premise (§2), and finally draw out some of the consequences of our conclusion (§3).

1 The Argument

First, an example. Here's what you know: there are wood-shavings on the floor of Roger's office. How much does this initial evidence support the claim that (C:) Roger does wood-Carving in his office? Presumably you should be unsure. Maybe (S:) it strongly supports

*Authors contributed equally.

[1] E.g. Feldman (2005); Christensen (2010, 2016); Horowitz (2014); Schoenfield (2015, 2018); Sliwa and Horowitz (2015).

[2] E.g. Christensen (2007a, 2010); Henderson (2021); Brössel and Eder (2014); Schoenfield (2018); Gallow (2021), and perhaps Sliwa and Horowitz (2015) and Elkin and Wheeler (2018).

the claim that he does (what else could generate wood-shavings?); but maybe it doesn't (perhaps wood-shavings could've come in on his boots). So you should be unsure about whether or not S is true, i.e. whether your initial evidence strongly supports the claim that Roger carves in his office.

What if you were to learn that your initial evidence *does* strongly support that claim? (It's hard for wood-shavings to stick to boots.) That should boost your confidence that he carves in his office: $P(C|S) > P(C)$. For C is exactly what you'd expect if S is true. But remember: hypotheses that should lead you to (more strongly) expect that C is true are themselves *confirmed* when you learn that C is true. S is such a hypothesis. Thus learning C (that Roger carves in his office) should also boost your confidence in S (that your initial evidence strongly supported this claim): $P(S|C) > P(S)$. More precisely, this follows from the symmetry of probabilistic relevance.[3] Since we know that learning that your (initial) evidence strongly supported C should boost your confidence in C—i.e. that $P(C|S) > P(C)$—it follows that learning C should boost your confidence that your initial evidence strongly supported C: $P(S|C) > P(S)$. And that's just to say that C provides evidence about S, i.e. provides (higher-order) evidence about how strongly your initial evidence supported C.

The subtle step in this reasoning is the appeal to the symmetry of probabilistic relevance. To make this more intuitive, consider an analogy in a familiar setting: coin flips. You know that a coin is about to be flipped, but you're unsure both about whether (H:) *it'll land Heads*, and about whether (B:) *it's Biased toward heads*. (Uncertainty about the bias is analogous to uncertainty about the strength of your initial evidence.)

What if you were to learn that it *was* biased toward heads? That should boost your confidence that it'll land heads: $P(H|B) > P(H)$. For H is exactly what you'd expect if B is true. But remember: hypotheses that should lead you to (more strongly) expect that H is true are themselves *confirmed* when you learn that H is true. B is such a hypothesis. Thus learning H (that the coin landed heads) should also boost your confidence in B (that it was biased toward heads): $P(B|H) > P(B)$. This should be familiar: obviously the way to get evidence about whether a coin is biased toward heads is to flip it and see if it lands heads. The *reason* this provides evidence is, again, due to the symmetry of probabilistic relevance: since learning it's biased should boost your confidence in H, learning that H should boost your confidence that it's biased.

Likewise with Roger: since learning that your initial evidence strongly supported C should boost your confidence in C, learning that C should boost your confidence that your initial evidence strongly supported C.

Now onto the general argument. A clarification: The example just given was framed in diachronic terms; we talked about how you should change your confidence in a claim upon learning a hypothesis. But this can get confusing—we want to talk about what your evidence supports, and as you learn things your evidence changes. Thus in what follows we'll frame things in synchronic terms, in terms of your unconditional and conditional credences at a given (initial) time. Nevertheless, we'll still talk about some new claim E (which you don't yet know) being evidence about H in the sense that learning E would

[3]By the ratio formula, $P(C|S) > P(C)$ iff $\frac{P(C \wedge S)}{P(S)} > P(C)$ iff $\frac{P(C \wedge S)}{P(C)} > P(S)$ iff $P(S|C) > P(S)$.

help you figure out whether H.

Clarification made, the crucial parts of our argument are (a) that it can be rational to be uncertain what your evidence supports, and (b) that what your evidence supports is correlated with truth. We'll assume that we can model your rational degrees of belief given your evidence (standards of reasoning, etc.) with a unique, precise probability function P.[4] 'P' is a definite description: 'the rational credence function I should have, given my current evidence (standards of reasoning, etc.)'.

Our first premise is a sufficient condition for one thing being evidence about[5] another:

P1) If your credence in X given q should differ from your credence in X, then q is evidence about X.
I.e. if $P(X|q) \neq P(X)$, then q is evidence about X.

Our second premise is a sufficient condition for a claim being higher-order evidence:

P2) If q is evidence about the opinion you should have in p given your current evidence, then q is (in part) higher-order evidence about p.

Our third premise is that, typically, it's rational to have at least some *higher-order uncertainty* concerning a given proposition q. That is, it's rational to be unsure what is the exact rational credence for you to have in q:

P3) Typically, it's rational to be uncertain of the rational credence in q.
I.e. typically, for all t, $P(P(q) = t) < 1$.

Our final premise is a way of making precise the idea that evidence is correlated with truth. In particular, the claim that your evidence supports confidence in q provides evidence in favor of q:

P4) Typically, q is more likely to be true if your evidence supports it than if it doesn't.
I.e. typically, for all t, $P(q|P(q) \geq t) > P(q|P(q) < t)$ (if both are well-defined).[6]

From (P1)–(P4), it follows that:

C) Evidence is typically higher-order evidence.

Here's why.

Consider some possible piece of evidence q, and suppose we are in a typical case. The opinion you should have about q, given your current evidence, is $P(q)$. It follows from (P1) and (P2) that *if* your credences about the value of $P(q)$ should be different on the supposition that q than without this supposition, then q is higher-order evidence about

[4]See e.g. White (2005); Schoenfield (2014, 2019); Schultheis (2018) on uniqueness and Seidenfeld and Wasserman (1993); White (2009a); Joyce (2010); Carr (2020) on precision.

[5]As always, claims about what's evidence for what must be understood as relative to some (implicit) background information—we'll take this background to be your current rational credence function P.

[6]Technically, our argument only requires the weaker premise that there is *some* t such that $P(q|P(q) \geq t) > P(q|P(q) < t)$, since we only need to show that there is some proposition of the form $[P(q) \geq t]$ such that $P(P(q) \geq t|q) \neq P(P(q) \geq t)$. Nevertheless, we think the stronger premise is well-motivated and plausible, since denying it would imply that there's a threshold at which evidential support is *anti*-correlated (or uncorrelated) with truth, i.e. that $P(q|P(q) \geq t) \leq P(q|P(q) < t)$. See footnote 10.

q—for example, if for some t, $P(P(q) \geq t|q) \neq P(P(q) \geq t)$, then q is higher-order evidence about q. We'll show that this is so.

Since we are in a typical case, (P3) tells us that you should be uncertain of the rational credence in q—thus there is a t such that both $P(P(q) \geq t) > 0$ and $P(P(q) < t) > 0$. It follows that both $P(q|P(q) \geq t)$ and $P(q|P(q) < t)$ are well-defined, and so by (P4), the former is greater than the latter:

$$P(q|P(q) \geq t) > P(q|P(q) < t)$$

Since by total probability $P(q)$ is a weighted average of (and thus in between) $P(q|P(q) \geq t)$ and $P(q|P(q) < t)$, it follows that the claim that *your current evidence supports confidence in* q (i.e. $P(q) \geq t$) is evidence for q:

$$P(q|P(q) \geq t) > P(q)$$

By the symmetry of probabilistic relevance, it thereby follows that q is evidence for *your current evidence supports confidence in* q:

$$P(P(q) \geq t|q) > P(P(q) \geq t)$$

But this means that q should change your credence about the value of $P(q)$, and so—by our above consequence of (P1) and (P2)—that q is higher-order evidence about q.

2 The Defense

We'll now defend each premise. The first two should be uncontroversial. (P1) follows from standard Bayesian conceptions of evidence. It is widely held that q is evidence for X if and only if $P(X \mid q) > P(X)$. Some theorists reject the left-to-right ('only if') part this claim because of the problem of old evidence, but they still agree with the right-to-left direction.[7] And all these theorists would, we think, agree that if $P(X \mid q) \neq P(X)$, then q is evidence *about* X.

(P2) is a plausible sufficient condition for evidence to be higher-order. After all, higher-order evidence is supposed to be evidence *about* your evidence—hence the name (Christensen, 2010). So even if other cases can count as well[8], surely if you get evidence about how confident your evidence should make you, that counts as (at least in part) higher-order evidence. We think all epistemologists can accept this, even if they disagree about how you ought to *respond* to higher-order evidence.

The last two premises are slightly more controversial. (P3) says that you typically ought to be uncertain of what credence you ought to have in a proposition q.[9] Two considerations support this claim. First, you typically ought to be uncertain about what your evidence *is*. For instance, suppose your evidence consists of all and only the propositions that

[7] In response to this problem, Joyce (2005) holds that q is evidence for X if and only if $P(X \mid q) > P(X \mid \neg q)$. But this still entails that if $P(X \mid q) > P(X)$, then q is evidence for (and hence about) X.

[8] E.g. some focus on evidence about your *reliability*, in various senses (Elga, 2007; Christensen, 2007b; Roush, 2009, 2016; White, 2009b; Schoenfield, 2015, 2018; Isaacs, 2019).

[9] For defenses, see Williamson (2000, 2008); Lasonen-Aarnio (2013); Roush (2016); Dorst (2019).

you know. Arguably, you can know a proposition but fail to know that you know it (Williamson, 2000). And unarguably, you can fail to know a proposition (because it's false) without knowing that you fail to know it (because you have no reason to suspect that it's false). Assuming that such failures are typical, you will typically be rationally uncertain about what your evidence is, and therefore uncertain about what credences you ought to have. One can resist this argument only by defending a conception of evidence for which both positive introspection and negative introspection typically hold—that if a proposition is (not) part of your evidence, then your evidence includes the fact that it's (not) part of your evidence. We think the prospects for such a conception of evidence are dim (but see Salow, 2019). Second, you typically ought to be uncertain about what your evidence *supports*. For instance, you can and should be uncertain about what theoretical virtues there are and how exactly they should be traded off against each other to determine overall evidential support. That is, you can and should be uncertain just how much weight should be given to simplicity versus elegance versus fit with the data, and so on.

One might dispute (P3) by holding (i) that facts about evidential support relations are necessary and (ii) that (ideally) rational agents are certain of all necessary truths. While (i) is plausible, (ii) is not. Ideally rational agents might have unlimited memories and computational speeds, but there is little reason to think that these cognitive capacities would enable you to be rationally certain of all necessary truths, especially of some such necessary truths, such as the fact that water is H_2O. One might respond that fundamental facts about evidential support are *a priori*, just like fundamental facts about ethics. But while these domains may be *a priori* in the sense that many facts therein are knowable *a priori*, this does not mean that *all* facts in these domains are knowable *a priori*. Some may not be knowable at all, for instance because they are too precise for us to latch onto safely. More generally, it is implausible that even ideally rational agents would be certain of the exact degree to which any given body of evidence supports any given hypothesis. Our uncertainty about just how much weight to give to simplicity versus elegance versus fit is not due to our finite computational speed or finite memory, but rather due to our ignorance about the exact strengths of the reasons corresponding to each of these theoretical virtues.

(P4) encapsulates the idea that evidence is correlated with truth. Typically, a proposition is more likely to be true if it's supported by your evidence (to above some threshold) than if it isn't. Relatedly, it captures the idea that you should be guided by your take on the evidence. This premise may be rejected by some extreme 'level-splitters' who hold that your rational credences about the degree to which your evidence supports a proposition usually have no bearing on what credence you ought to have in that proposition (Lasonen-Aarnio, 2015; Weatherson, 2019). But among those who hold that the former do—at least *typically*—have some bearing on the latter, (P4) is common ground.[10]

[10]Dorst (2020) explicitly argues for (P4), showing that it follows from the Value of Evidence (Good, 1967). Hedden (2019) argues for a similar conclusion, and many other authors voice sympathy for the basic idea, e.g. Conee (1992, 667), Williamson (2000, 202), White (2006, 539), Sliwa and Horowitz (2015, 2846), Horowitz (2019, 107).

3 The Upshots

We'll close with two upshots. The first concerns the distinction between first-order and higher-order evidence. It's tempting to think that your total evidence can be divided into two disjoint categories, one consisting of all your first-order evidence and the other consisting of all your higher-order evidence.[11] But this doesn't work. Suppose your evidence includes the proposition that you recently ingested a reason-distorting drug (Christensen, 2010). Is that proposition part of your first-order evidence or part of your higher-order evidence? The question has a false presupposition, namely that all pieces of evidence count as first-order or higher-order (but not both) *tout court*. This proposition is first-order evidence concerning the hypothesis that your reasoning is distorted, but it is higher-order evidence concerning, say, the hypothesis that your share of the dinner bill is \$43—a conclusion you reached on the basis of reasoning that you now have grounds to distrust.

This suggests that while we might be unable to divide your evidence into two disjoint categories—first-order and higher-order—*tout court*, we might be able to divide your evidence into these two disjoint categories *relative to a given hypothesis*. We could say, for any hypothesis H, that these propositions are your first-order evidence about H, while those other ones are your higher-order evidence about H. And one proposition could be first-order relative to one hypothesis but higher-order relative to another.

If our argument is right, then this proposal is also mistaken. For any given hypothesis can—in fact, usually does—count as both first-order and higher-order evidence relative to itself. The proposition q is maximally strong first-order evidence for q, by virtue of entailing it. But q is also (typically) higher-order evidence about q, by virtue of constituting evidence about how strongly your (initial) evidence supported q.

Perhaps instead we should distinguish between first-order and higher-order *routes* or *paths* of evidential support (Christensen, 2019). We shouldn't speak, in the first instance, of propositions counting as first-order or higher-order evidence (even relative to a given hypothesis). Rather, we should speak of paths from an evidence proposition to a hypothesis as being first-order or higher-order. These paths could be represented using directed graphs. To a first approximation, an evidential support path from q to h is higher-order if it goes via a proposition concerning the degree to which your evidence supports h, and it is first-order otherwise. Sometimes, a piece of evidence bears on a hypothesis in a first-order way. Sometimes it bears on a hypothesis in a higher-order way. But sometimes, it bears on a hypothesis in both a first-order way and a higher-order way. This will be the case if there is both an evidential support path from q to h which goes via a proposition concerning the degree to which your evidence supports q, and also an evidential support path from q to h which does not go through such an intermediate proposition and is thus more direct. In a slogan: it's paths of evidential support, not propositions, that can be categorized as first-order or higher-order.

Second upshot: several authors have argued that in the context of higher-order evidence, standard rational norms break down or run into dilemmas. Christensen (2007a) and Henderson (2021) argue that such cases tell against (precise) probabilism (cf. Elkin

[11] See e.g., Schoenfield (2015, 426), who writes, 'I'll use the term "first-order evidence" to refer to any evidence you have that is not higher-order evidence.'

and Wheeler, 2018). Meanwhile, Christensen (2010), Brössel and Eder (2014), Schoenfield (2018), and Gallow (2021) argue that such cases should lead rational agents to deviate from conditionalization. If we're right, then these arguments risk over-generalizing—showing that probabilism and conditionalization are *almost never* rational norms. That seems like an implausible verdict. Thus, at the very least, authors who defend such views must be very clear on the sense of 'higher-order evidence' that is meant to warrant such deviations from standard rational norms, and explain why the above sufficient condition for higher-order evidence is not within their intended target.

In sum: higher-order evidence is everywhere. Because of this, we should be cautious about proposals that claim that you should respond to it in a distinctive way from how you should respond to evidence generally.[12]

References

Brössel, Peter and Eder, Anna-Maria A, 2014. 'How to Resolve Doxastic Disagreement'. *Synthese*, 191(11):2359–2381.

Carr, Jennifer Rose, 2020. 'Imprecise Evidence without Imprecise Credences'. *Philosophical Studies*, 177(9):2735–2758.

Christensen, David, 2007a. '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.

———, 2007b. 'Epistemic Self-Respect'. *Proceedings of the Aristotelian Society*, CVII(3):319–337.

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

———, 2016. 'Disagreement, Drugs, etc.: From Accuracy to Akrasia'. *Episteme*, 13(4):397–422.

———, 2019. 'Formulating Independence'. In Mattias Skipper Rasmussen and Asbjørn Steglich-Petersen, eds., *Higher-order evidence: new essays*, 13–34. Oxford University Press.

Conee, Earl, 1992. 'The Truth Connection'. *Philosophy and Phenomenological Research*, 52(3):657–669.

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.

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

Elga, Adam, 2007. 'Reflection and Disagreement'. *Noûs*, 41(3):478–502.

Elkin, Lee and Wheeler, Gregory, 2018. 'Resolving Peer Disagreements Through Imprecise Probabilities'. *Nous*, 52(2):260–278.

Feldman, Richard, 2005. 'Respecting the Evidence'. *Philosophical Perspectives*, 19(1):95–119.

Gallow, J. Dmitri, 2021. 'Updating for Externalists'. *Noûs*, 55(3):487–516.

Good, I J, 1967. 'On the Principle of Total Evidence'. *The British Journal for the Philosophy of Science*, 17(4):319–321.

Hedden, Brian, 2019. 'Hindsight Bias is Not a Bias'. *Analysis*, 79(1):43–52.

Henderson, Leah, 2021. 'Higher-Order Evidence and Losing One's Conviction'. *Noûs*, To appear.

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

———, 2019. 'The Truth Problem for Permissivism'. *The Journal of Philosophy*, cxvi(5):237—-262.

Isaacs, Yoaav, 2019. 'The Fallacy of Calibrationism'. *Philosophy and Phenomenological Research*, To appear.

[12]Thanks to Alan Hájek and two anonymous referees for helpful comments on earlier drafts.

Joyce, James M., 2005. ‘Bayesianism’. In Alfred R. Mele and Piers Rawling, eds., *The Oxford Handbook of Rationality*, April 2021, 132–153. Oxford University Press.

———, 2010. ‘A Defense of Imprecise Credences in Inference and Decision Making’. *Philosophical Perspectives*, 24:281–323.

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.

Roush, Sherrilyn, 2009. ‘Second Guessing: A Self-Help Manual’. *Episteme*, 251–268.

———, 2016. ‘Knowledge of Our Own Beliefs’. *Philosophy and Phenomenological Research*, 93(3):45—-69.

Salow, Bernhard, 2019. ‘Elusive Externalism’. *Mind*, 128(510):397–427.

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

———, 2015. ‘A Dilemma for Calibrationism’. *Philosophy and Phenomenological Research*, 91(2):425–455.

———, 2018. ‘An Accuracy Based Approach to Higher Order Evidence’. *Philosophy and Phenomenological Research*, 96(3):690–715.

———, 2019. ‘Permissivism and the Value of Rationality: A Challenge to the Uniqueness Thesis’. *Philosophy and phenomenological research*, 99(2):286–297.

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

Seidenfeld, Teddy and Wasserman, Larry, 1993. ‘Dilation for Sets of Probabilities’. *The Annals of Statistics*, 21(3):1139–1154.

Sliwa, Paulina and Horowitz, Sophie, 2015. ‘Respecting *All* the Evidence’. *Philosophical Studies*, 172(11):2835–2858.

Weatherson, Brian, 2019. *Normative Externalism*. Oxford University Press.

White, Roger, 2005. ‘Epistemic Permissiveness’. *Philosophical Perspectives*, 19(1):445–459.

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

———, 2009a. ‘Evidential Symmetry and Mushy Credence’. In Tamar Szabó Gendler and John Hawthorne, eds., *Oxford Studies in Epistemology*, volume 3, 161–186. Oxford University Press.

———, 2009b. ‘On Treating Oneself and Others as Thermometers’. *Episteme*, 6(3):233–250.

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.