

Is Higher-Order Evidence Evidence?

Eyal Tal

{Penultimate draft; forthcoming in *Philosophical Studies*; Please cite the published version}

Suppose we learn that we have a poor track record in forming beliefs rationally, or that a brilliant colleague thinks that we believe P irrationally. Does such input require us to revise those beliefs whose rationality is in question? When we gain information suggesting that our beliefs are irrational, we are in one of two general cases. In the first case we made no error, and our beliefs are rational. In that case the input to the contrary is misleading. In the second case we indeed believe irrationally, and our original evidence already requires us to fix our mistake. In that case the input to that effect is normatively superfluous. Thus, we know that information suggesting that our beliefs are irrational is either misleading or superfluous. This, I submit, renders the input incapable of justifying belief revision, despite our not knowing which of the two kinds it is.

Keywords: Disagreement; epistemic akrasia; higher-order evidence; steadfast

(1) Introduction

Some information that we come across seems to bear on whether we believe things rationally. For example, we might learn that we tend to discount evidence that we are bad drivers, or that we are hopeless at forming rational beliefs about probabilities. We might also learn that others agree or disagree with our assessment of the evidence, or that they take our evidence to justify some particular degree of belief. The information that we gain in these sorts of cases corresponds to what philosophers often call 'higher-order evidence' (hereafter, HOE). HOE is presented in the literature as evidence about the rationality of our beliefs, or evidence about what our evidence supports. For

instance, David Christensen (2010a) describes HOE as “evidence that the evidential relations may not be as I’ve taken them to be,” and Thomas Kelly (2010) understands HOE to be “evidence about the normative upshot of the evidence to which [one] has been exposed.”

In what follows I argue that the information we receive in the kinds of examples above is irrelevant to what we should believe.¹ I will refer to our evidential situation prior to receiving some HOE as our *original evidence*, or our *other evidence*. I will refer to the doxastic attitude whose rationality is in question as the *lower-level belief*. The discussion proceeds mainly in terms of a ternary model of belief, but is intended to apply equally well to a graded model. Thus, ‘belief’ will be used interchangeably with ‘doxastic attitude’ and with ‘credence.’ Given this terminology, the claim I defend is that we are rationally required to form our lower-level beliefs without regard for HOE.²

Throughout the paper I assume an evidentialist framework. On this framework, a belief is rational if and only if it accords with and is properly based on the agent’s total evidence. I also assume that rationality requires that we form the beliefs that our evidence supports and in the right way, whether or not we are (or even can be) aware of what our evidence supports.³ My goal is to show that a strong case can be made against the inclusion of HOE in lower-level belief formation, from a common evidentialist understanding of rational justification.

¹ Some HOE brings with it more than just information about the rationality of our beliefs. For example, if someone who we know has more evidence than we do tells us that our evidence E misleadingly supports P, we learn both that she thinks that E supports P and also that she thinks that P is false. But by implying that she thinks P is false, such HOE carries with it some run-of-the-mill evidence against P, perhaps via an *Evidence of Evidence is Evidence* principle as Feldman (2014) suggests. I will focus on cases where the HOE does not bring with it such evidence.

² The list of philosophers who have expressed sympathies to the thought that HOE requires lower-level belief revision is long. See, for instance, Feldman (2009), Kelly (2010), Christensen (2010a, 2016), Cohen (2013), Schechter (2013), Sliwa & Horowitz (2015), Schoenfield (2016).

³ I return to this assumption in section 4, where it is called into question.

To lay my cards on the table, here is a sketch of the argument to come. I say that each instance of HOE fits into one of two categories: *superfluous* or *misleading*. Superfluous HOE concurs with our original evidence. It tells us something in line with what our original evidence does or does not justify. Consequently, this kind of HOE does not change what we should believe.⁴ Misleading HOE is in conflict with our original evidence. It tells us something wrong about what attitude our original evidence justifies. This kind of HOE lacks the main feature that made HOE seem relevant to our lower-level beliefs all along, namely an ability to correctly indicate what our evidence supports. So we are always in a position to know that our HOE either does not change what we should believe, or it is wrong about what our evidence supports. This requires us to refrain from HOE-based belief revision, despite not knowing which of the two kinds of HOE we face. If in the best scenario the HOE changes nothing about what we should believe, and in the worst scenario it misleads us about what our original evidence supports, then HOE-based belief revision is irrational.

Others have used a similar observation to make related points. Hartry Field (2000) notes that the ideal credibility of a logical truth is unaffected by whether well-respected logicians accept it. Accordingly, the ideal credibility of any proposition P given evidence E is unaffected by what others take the credibility of P given E to be. Thus, an ideally rational agent who is also certain of her own rationality would be irrational to change her belief in light of HOE. Miriam Schoenfield (2015) argues that from a calibrationist perspective, beliefs that accord with our evidence and exclude HOE do better than those which take HOE into account. From that perspective it is therefore best to exclude HOE, and correctly follow our other evidence alone. Both of these ideas are

⁴ I use talk of what we *should* believe interchangeably with talk of what *rationality permits*, and with what is *rationally justified*.

right because HOE recommends a belief different from that which our other evidence does only when it is wrong about what our other evidence supports. But there is a gap between these ideas and the conclusion that we should dismiss HOE. We are not ideally rational, and those who are might not be certain that they are. It is also not clear that we ought to be calibration-maximizers, and even less clear (as Schoenfield says) that excluding HOE is the way for fallible creatures like us to successfully pursue that aim. So I will try to bridge this gap by arguing that the odd properties of HOE tell against its relevance to what we should believe.

The argument will come with two additions. The first is an account of why excluding HOE feels patently irrational. The basic idea is that ignoring HOE is imprudent relative to the goal of having beliefs that fit our evidence, but rationality does not require that we be prudent in this respect.⁵ The second is an explanation of why ignoring HOE when forming lower-level beliefs does not imply that we should sometimes be *epistemically akratic*. The worry is that if we may rationally retain our beliefs despite strong HOE that they are irrational, we would still need to believe that those beliefs are irrational. I will argue that if akratic belief combinations are irrational, we should put aside HOE even when forming beliefs about the rationality of our beliefs.

The discussion proceeds as follows: In section 2 I offer a catalogue of all possible kinds of HOE, and show that each fits into one of the two mentioned categories—superfluous or misleading. In section 3 I argue that due to features of those categories, HOE-based belief revision is rationally forbidden. Section 4 is dedicated to explaining away the intuitive implausibility of such resistance to HOE, and section 5 to addressing worries that the account leads to epistemic akrasia.

⁵ Even friends of pragmatic encroachment like Fantl and McGrath (2011) would grant that the credences we should form are independent of prudential considerations.

(2) Six kinds of HOE

Instances of HOE can be characterized according to three parameters: Valence, Correctness, and Directionality.⁶

Valence: HOE can be *positive* or *negative*. HOE is positive when it suggests that our held belief is rational, and negative when it suggests that our held belief is irrational. For example, learning that a smart friend shares our view given the same evidence is positive HOE, whereas learning that we have a poor track record in forming rational beliefs is negative HOE. Simply put, HOE may suggest that we did or did not succeed in forming a rational belief given our evidence.⁷

Correctness: HOE can be *right* or *wrong*. What any HOE suggests about whether we believe rationally will either be true or false. When we believe rationally and gain positive HOE, or believe irrationally and gain negative HOE, the valence parameter of the HOE corresponds to our situation. In those cases the HOE is *right*. But the opposite can happen as well. We may gain information suggesting that we believe rationally when we in fact do not. We may also gain information suggesting that we believe irrationally when we in fact believe rationally. In these cases, our HOE's valence parameter does not correspond to our situation, and the HOE is *wrong*. Thus, *rightly positive* HOE suggests that we believe rationally when we in fact do, whereas *wrongly positive* HOE suggests that we believe rationally when we in fact believe irrationally. Similarly, *rightly negative* HOE suggests that we believe irrationally when we in fact do,

⁶ Little rides on there being absolutely no exceptions to this claim. It is true of, at least, the vast majority of HOE. If the arguments to come only apply to the vast majority of HOE, that would be enough for my purposes.

⁷ We can gain negative HOE even in cases where we have no view about some matter. In those cases, any information suggesting that our evidence requires a particular credence c would imply that we did not do a good job in forming a rational credence.

whereas *wrongly negative* HOE suggests that we believe irrationally when we in fact believe rationally.⁸

Directionality: HOE can either be *directional* or *non-directional*.⁹ If some HOE suggests that we have overestimated, underestimated, or accurately assessed our evidence, it is *directional*. This is because such HOE carries with it some indication of which belief is required by our other evidence. Alternatively, if some HOE merely suggests that we believe irrationally and nothing more, it is *non-directional*. For example, learning that everyone thinks we made a reasoning error is non-directional HOE, whereas learning that everyone takes our evidence to support a particular credence c is directional HOE.

The three parameters specified yield eight potential kinds of HOE:

HOE₁: Directional and wrongly negative.

HOE₂: Directional and rightly negative.

HOE₃: Non-directional and wrongly negative.

HOE₄: Non-directional and rightly negative.

HOE₅: Directional and wrongly positive.

HOE₆: Directional and rightly positive.

HOE₇: Non-directional and wrongly positive.

HOE₈: Non-directional and rightly positive.

However, notice that positive HOE will always be directional, since any indication that we believe rationally is an indication that our belief is rational as is. It follows that neither *HOE₇* nor *HOE₈* are possible kinds of HOE.

To show that none of the six possible kinds of HOE requires belief revision, I first

⁸ So defined, the HOE would be right even in cases where it points to a doxastic attitude that is not supported by the evidence, as long as we indeed failed to form the rational attitude. For example, the HOE is right when we believe a proposition that we should disbelieve, and the HOE suggests that we should suspend judgment.

⁹ Christensen (2010b) introduces this distinction.

argue that they fit into two distinct categories, and then argue that the features of these categories justify excluding HOE when forming lower-level beliefs.¹⁰

(2.1) HOE₁ and HOE₃ are misleading

HOE₁ and HOE₃ capture all instances in which the HOE is wrongly negative, i.e., wrongly suggesting that we have failed to believe rationally. Such HOE can be directional or non-directional, to varying degrees of specificity. For example, we may come across information suggesting that we made some mistake in reasoning, or that we tend to underestimate the strength of our evidence, or that we underestimated the strength of our evidence by precisely .1.

By definition, when we gain wrongly negative HOE, we hold the relevant belief rationally. So, HOE of this kind is guaranteed to be misleading in an important sense. It provides an incorrect indication of what our original evidence supports. This fact makes wrongly negative HOE misleading about the very matter that made HOE seem evidentially relevant to our lower-level beliefs in the first place.

(2.2) HOE₄ is superfluous

HOE₄ and HOE₂ are both rightly negative, i.e., they suggest that we believe irrationally when we in fact do. Such HOE can also be directional or non-directional to varying degrees of specificity.

When rightly negative HOE is non-directional (HOE₄), it correctly indicates that we believe irrationally, without indicating what the rational belief is. Consequently, any

¹⁰ We should not deny that with suitably arranged background knowledge, almost anything can be evidence for almost anything else. On some backgrounds, testimony that P is evidence against P, and a memory seeming as of Q is evidence against Q. But the existence of such backgrounds does not touch the interesting issues of the epistemic significance of testimony and memory.

move to a new attitude based on such HOE (as opposed to based on correct reassessment of our original evidence) would be arbitrary. This claim remains true even if we suspend judgment on the basis of such HOE—a move that may seem like a principled response to non-directional HOE. Suspension of judgment, like any other doxastic attitude, is rational only when the balance of evidence goes a particular way.¹¹ Not knowing what the evidence supports no more justifies suspension of judgment than it justifies disbelief. Suspension of judgment is not a doxastic safe-zone that we may occupy while unsure about what our evidence supports, and where we are shielded from rational criticism. If we want to resort to a doxastic safe-zone, perhaps we could stop having a view on the relevant matter while we inquire further.¹² But this too cannot be what rationality requires when non-directional HOE tells us of our failure to believe rationally. The question of whether we should have a view about a matter does not fall within the jurisdiction of rational requirements.¹³ Rational requirements instruct us that *if* we form a view on a matter, it be the appropriate one. So this kind of HOE does nothing but agree with our original evidence that we should have some different attitude than the one we have.

In addition, there is no sense in which we would be *more required* to form the rational belief upon gaining information to the effect that we believe irrationally. Requirements

¹¹ This is compatible with understanding ‘suspension of judgment’ both as a middling credence (see Hájek (1998) and Christensen (2009)) and as a belief about our evidence (see Rosenkranz (2007) and Raleigh (forthcoming)).

¹² Abstaining from having a view may not be psychologically possible. But assuming it is possible, it may still seem irrational to not have a view on a matter that we have considered and have evidence about. Refusing to have any doxastic attitude whatsoever about what is right in front of our eyes certainly seems irrational. See Nelson (2010) and Alexander (2013) for views on which abstinence from having a doxastic attitude can be rationally permitted. See Friedman (2017) for a view of suspension as closely tied to inquiry.

¹³ Cases in which we know that we are guaranteed to believe irrationally may look like ones in which rationality requires that we have no view. This appearance is arguably mistaken. If such cases are possible, they could at most show that we are sometimes doomed to fail to believe as we should.

are binary, either applying to an agent or not. One cannot be more required or less required to believe P. Since our original evidence already requires us to form a rational belief, it still does (and to no greater degree) if we are told that we believe irrationally.

From this we should conclude that non-directional and rightly negative HOE is superfluous, and does not change what an agent is required to believe. We should correct our doxastic delinquencies before learning about them just as much as we should after learning about them.

(2.3) HOE₂ is superfluous or misleading

Unlike HOE₄, HOE₂ does not only rightly suggest that we believe irrationally, but also tells us something about the direction in which we must revise. In so doing, this kind of HOE can get certain things wrong. The HOE might suggest that we have overestimated our evidence by a little when in fact we overestimated it by a lot (or even underestimated it). So directional and rightly negative HOE can be *accurate* or *inaccurate*, depending on whether it correctly indicates the extent and direction of the mistake we made.

Accurate, directional, and rightly negative HOE (HOE_{2a}) accurately points in the direction of (or precisely to) the rationally required belief, when we indeed failed to form that belief. Since this HOE is accurate, it recommends the exact same belief that our original evidence does, and is thus superfluous. Inaccurate, directional, and rightly negative HOE (HOE_{2i}) points away from the attitude that is rational on our other evidence, and is thus misleading. So when inaccurate, HOE₂ is misleading, and when accurate, it is superfluous.

(2.4) HOE₅ and HOE₆ are superfluous or misleading

HOE₅ and HOE₆ capture all instances in which the HOE is positive, i.e., suggesting that we hold our belief rationally. Such information intuitively warrants no belief revision. It would be odd to think that positive feedback about our belief's rationality might require us to abandon the belief.

Notice that HOE₆ suggests that our belief is rational when it in fact is. If such HOE supports any belief it is the very same one that our other evidence does, and that we already hold. Meanwhile, HOE₅ suggests that our belief is rational when in fact it is not. HOE₅ will therefore be recommending a belief different than the one supported by our other evidence. So positive HOE is superfluous or misleading.

(3) The upshot for higher-order evidence

The possible kinds of HOE fit into two general categories. HOE_{2a}, HOE₄, and HOE₆ agree with our original evidence, and change nothing about what we should believe. HOE₁, HOE_{2i}, HOE₃, and HOE₅ conflict with what our original evidence supports, and are thus misleading. While we may not know which of these two categories our HOE falls under, we are always in a position to know that it either does not change what we should believe, or it is misleading about what we should believe. But if we know that our HOE could only affect what we should believe by misleading us about what we should believe, then revising our beliefs in light of it starts to look wrong.¹⁴

¹⁴ Some unique HOE-like inputs might resist the misleading/superfluous charge. For example, we may learn that our total evidence *including this very input* supports so and so (thanks to Antonia Peacocke for this point). If such HOE is possible, it could change what we should believe by changing our total evidence, while avoiding misleading claims about whether we assessed our original evidence rationally. However, one concern is that it is difficult to conceive of such inputs. Since this kind of HOE does not question our rationality given our original evidence, it is unclear what new evidence it brings with it so as to change what our total evidence supports. Another concern is that such HOE marks a departure from the kind of input that the literature in the area has been concerned with—like peer disagreement, evidence of bias, or reasoning-distorting drugs. So in the event

When we know we face some misleading or otherwise unreliable input about a proposition, we may not let it figure in our belief on the matter. For example, we should not let the opinion of someone who we know is guessing affect our view of what our evidence supports regarding P, and by extension our view of P. Of course, unless we know that we face some misleading or unreliable input it would typically be irrational to exclude it when forming beliefs. Misleading information that we do not know is misleading can easily affect what we should believe. For this reason, when we do not know whether our HOE is misleading, it is tempting to think we should revise our beliefs at least to some extent in response, just in case it is not misleading. But this is where HOE is unique. If we know that the only alternative to it being bad at telling what we should believe is that it changes nothing about what we should believe, we should form our views on the relevant matter without it.

To further motivate the thought, consider a case from the moral domain:

Directions

Traveler morally ought to (do her best to) reach Rome. She is given a set of directions D₁, which she knows is both accurate and easily within her ability to follow. Before using D₁, she is offered an alternative set of directions D₂, which she knows to be one of two kinds: either D₂ contains simple, accurate, and timesaving directions to Rome, or it contains misleading directions to a randomly chosen destination.

Assuming that Traveler has only one attempt to make it to Rome and can use only one

that such unique HOE is both possible and evades the misleading / superfluous charge, it leaves most other cases of HOE vulnerable.

set of directions, may she opt to use D₂? It seems clear that she may not. If reaching Rome (or doing her best to reach Rome) is required of Traveler, and if she knows that D₁ accurately directs her there, she must follow it alone. Following D₂ would amount to sacrificing her perfectly accurate guide D₁, which enables her to do as she ought to, in exchange for a more convenient guide at best, and a bad guide at worst. Doing so would not count as Traveler's best effort to reach Rome, and so using D₂ would be wrong.

In both Directions and in cases of HOE, the agent has perfectly accurate directions for doing what she should do. In Directions, it is D₁. In cases where we are presented with HOE, it is our original evidence. In both Directions and cases of HOE, an opportunity presents itself to forgo the perfectly accurate set of directions that we possess, in exchange for another that is at best more convenient and at worst defective. In both, making the exchange is normatively forbidden. If Traveler were to use D₂ rather than her trusty D₁, she would be morally criticizable. If we rely on HOE to revise our lower-level beliefs instead of (or in addition to) our other evidence alone, we would be rationally criticizable. This verdict holds even in the event that D₂ or our HOE is more likely to be of the accurate kind than of the misleading kind.

There is a clear disanalogy between Directions and cases of HOE. Directions is set up so that the agent can easily do as she ought to by following D₁, whereas we often cannot figure out what our evidence supports. This difference may lead some to think that making use of HOE is rationally permissible, because it helps us to do our best to obey the rational requirements imposed by our original evidence. In a sense, this is right. Prudentially speaking, in order to achieve the goal of forming the belief that fits our original evidence, we would do well to follow our HOE where it leads despite risk of it misleading us. After all, our HOE is frequently a better guide to what our original

evidence supports than we are. But the prudential benefits of using HOE can compromise our intuitions, and cause us to think that rationality requires that we use HOE in doxastic revision. Facts about what steps we can take to improve our odds of reaching the belief that fits our evidence do not figure into rational requirements. For example, getting a good night's sleep would improve these odds, but is not rationally required. It is easy to conflate the requirement to believe what the evidence supports with a need to do our best to believe what the evidence supports, especially when the relevant action (of following the HOE) is doxastic in nature. Doing one's best to believe what the evidence supports indeed tends to help with believing rationally. But this does not make the former rationally required. This point is necessary for exposing HOE-based belief revision as a kind of prudential shortcut, rather than anything rationally required.

Relatedly, HOE cannot affect what rationality requires by improving our ability to tell what our evidence supports.¹⁵ Just as rational requirements need not be sensitive to the steps that we can take to improve our odds of believing as we should, rational requirements need not be sensitive to our ability to tell what our evidence supports. Standard accounts of Evidentialism, for example, do not take such abilities to figure into what rationality requires. So there is a formidable tradition in epistemology that would reject the suggestion that HOE could operate by affecting our epistemic abilities, and thereby change what we should believe.¹⁶

Our original evidence is a perfectly accurate guide to what we ought to believe. If

¹⁵ Thanks to Alex Worsnip for an exchange leading to this line.

¹⁶ One move that is available to my opposition (and coheres with standard Evidentialism) is to take our evidence to depend on our epistemic abilities. If our ability to tell what the raw data make probable affects what evidence we have, then our evidence may change in light of input about what the data make probable. This suggestion deserves more attention than I have room to offer, but one concern is that it would no longer be taking HOE to be the kind of input that challenges our rationality given our own evidence.

we form our lower-level beliefs by incorporating HOE, we would be sacrificing that perfectly accurate guide in exchange for a potentially inaccurate one. This highlights the fact that the only thing that HOE-based belief revision has going for it is its promise to help us believe as we already should. Since epistemic rationality does not require that we do our best to believe rationally, it does not require HOE-based belief revision. HOE could still be a motivating reason to revise or retain one's beliefs. It may even be an epistemically respectable motivation, in some sense. But the normative reasons that are determinative of a belief's rationality are independent of the motivations an agent has for reassessing her beliefs. My thesis is about those normative reasons only, and not about what epistemically respectable motivating reasons there might be.

In the next section I say more about why resistance to HOE feels irrational. After that, I address the worry that the view permits maintaining a doxastic attitude while believing that the attitude is irrational.

(4) Prudential vs. rational belief

It is hard to believe that we should really take no amount of HOE to require belief revision. When many experts tell us that we have overestimated the strength of our evidence, or when we learn that we have been dosed with a powerful reasoning-distorting drug, it seems clear that the way to take that input seriously is to significantly revise our beliefs.

Earlier I claimed that this seeming stems from our confusing rationally required belief revision with prudential belief revision relative to the goal of acquiring the rationally required beliefs. This explanation is not a far cry from distinctions that Maria Lasonen-Aarnio (2010) and Joshua Schechter (2013) draw in a similar context. Lasonen-

Aarnio distinguishes between requirements of rationality and requirements of reasonableness. Schechter distinguishes between the requirements of epistemic justification and the requirements of epistemic responsibility. We are more likely to reach beliefs that fit our evidence when we form our beliefs reasonably and responsibly. Accordingly, we may evaluate HOE according to whether it changes the rational requirements that we are under, or according to whether it changes what we can do to obtain beliefs that fit our evidence. On the account I offer, HOE does not affect rational requirements, for we know that it can add nothing non-misleading to what we should already believe given our other evidence. But perhaps, in order to count as prudent believers, we must pay attention to information about our having made some reasoning error. Perhaps prudently responding to this kind of information involves following our HOE in order to maximize our odds of approaching the belief that our original evidence supports, or at least checking our reasoning and looking for errors as Schechter suggests. Failure to respond to HOE in these ways may entail a violation of a prudence requirement, even though it does not entail a violation of a rational requirement. But as with Lasonen-Aarnio's view of epistemic reasonableness, on my view we need not grant that requirements of prudent belief formation are rational requirements.¹⁷ If we deny that they are, we can explain why dismissing HOE does not interfere with believing rationally.

Requirements of prudent belief formation are undeniably closely related to rational requirements. Lasonen-Aarnio considers reasonableness "a matter of managing one's beliefs through the adoption of policies that are generally knowledge conducive, thereby manifesting dispositions to know and avoid false belief..." Good habits of belief

¹⁷ Schechter (2013) takes irresponsibility to undermine rationality.

formation promote rational belief acquisition. Checking our reasoning for mistakes helps us avoid irrationality, and since avoiding irrationality *is* rationally required of us, it might seem that checking our reasoning is too. Recall, though, that many other things facilitate avoidance from forming irrational beliefs. Getting a good night's sleep is one example I have mentioned. It is nothing more than a contingent fact that we often fall short of fulfilling rational requirements, and various courses of action (some, doxastic) can minimize these failures. Yet it is not rationally required to abide by the requirements of prudent belief formation. It is not rationally required to get plenty of sleep. It is merely prudent to do so. The rationality of a belief is determined by the belief's accordance with our evidence in the proper way. The rationality of a belief is independent of the available courses of action that could improve our odds of believing in accordance with our evidence.

Some may resist the explanation on offer by appeal to the thought that following our HOE is not clearly an action, whereas getting enough sleep is. That could explain why rationality does not require sleep, but may require HOE-based belief revision. While rationality is not in the business of telling us to be well-rested, hydrated, or focused when we form beliefs, it is in the business of telling us to change our beliefs due to new information. HOE-based belief revision is a way to improve our odds of believing what our other evidence supports, and thus it offers an indirect way to improve our odds of believing the truth (or improve our expected accuracy) regarding the relevant proposition. In this respect, HOE-based belief revision looks similar to belief revision due to any old evidence. So why should we classify HOE-based belief revision as the kind of thing that falls outside the purview of rational requirements?

The following case stresses the difference:

Decent Instincts

A group of psychologists and epistemologists examine Watson's mind. They share their findings with him: "The subject is terrible at assessing his evidence, but has decent instincts. He is a little more likely to form the doxastic attitude that his evidence justifies toward P by following his hunch about P than by carefully assessing the evidence."

By stipulation, Watson knows that his best shot at having the rationally required attitude toward P is to follow his hunch. But that does not make it rational for him to form his attitude in that way. We may never form a doxastic attitude that is not based on the evidence, even if doing so represents our best shot at forming the attitude that fits our evidence. Being bad at assessing our evidence does not justify seeking alternative ways of reaching the rationally required attitude.¹⁸

Unlike sufficient sleep and hydration, the case involves a doxastic move rather than a physical action. This fact makes the case more straightforwardly analogous to using one's HOE. The case describes a way for Watson to improve his odds of believing as he should. So does HOE. If believing based on HOE is rationally required, then so is believing based on a hunch when we know that the hunch is our best shot at having the rationally required belief. But neither is rationally required, and for the same reason that getting enough sleep and staying hydrated are not. It does not matter whether the event that could help us believe as we already should is doxastic in nature. The fact that

¹⁸ This is especially clear in the case where Watson's hunch is misleading. Suppose that Watson is a detective, and despite the evidence clearly supporting Suspect's innocence, he follows his hunch that Suspect is guilty. When we correctly assess the evidence, we may ask Watson why he is charging Suspect with the crime. If he responds by saying that he does not go by the evidence, but by his hunch, we should want him to retire.

the relevant event primarily serves the function of assisting us with believing as we already should is what excludes it from being rationally required.¹⁹

We should therefore overcome the unintuitive feel of the thought that our beliefs can be perfectly rational despite strong HOE to the contrary. It may well be that, as a matter of fact, if many experts told us that we should disbelieve rather than believe P given our evidence, then we would all follow their advice. It would indeed be a good move to change our view when the experts tell us it is irrational, if our goal is to have beliefs that fit our original evidence. We might also say that agents who do so are not to be blamed, in some sense, for following their HOE.²⁰ But all of that is consistent with the claim we are not rationally required to take the steps that maximize our odds of believing as we should.

(5) Epistemic Akrasia

I have argued that HOE does not require revision of the (lower-level) beliefs the rationality of which our HOE concerns. One key implication of this view is that we should sometimes believe P despite possessing HOE that suggests we should not. Call this implication *Steadfast*:

Steadfast: If we believe P rationally and then receive HOE suggesting that our belief is irrational, we should (still) believe P.

I have kept quiet about how HOE affects the beliefs we should have regarding

¹⁹ The idea here is similar to Berker's (2013) point against epistemic consequentialism, which is that the view's central problem is in "its focusing on the promoting or conducting relation."

²⁰ See Boulton (2016), Littlejohn (forthcoming) for discussions of epistemic excuses.

whether our beliefs are rational, i.e., our *higher-level beliefs*. Indeed, HOE often seems to require revision of higher-level beliefs. For instance, an expert who tells us that we believe irrationally seems to provide us with excellent evidence that we believe irrationally. Just as any testimony that P is typically evidence for P, testimony that we believe irrationally looks like good evidence that we believe irrationally. So, strong enough negative HOE would appear to require belief that our corresponding lower-level belief is irrational. Call this claim Higher-Level Influence:

Higher-Level Influence: Strong HOE suggesting that a belief is irrational requires believing that the belief is irrational.

Steadfast and Higher-Level Influence combined are in tension with the enkratic constraint:²¹

Enkrateia: Rationality forbids having a doxastic attitude *a* toward P while believing that having *a* toward P is irrational.²²

According to Enkrateia, agents can never rationally believe P and also rationally believe that believing P is irrational. Just as it is (morally) impermissible to act akratically, i.e., in a way that we believe to be immoral, it is (rationally) impermissible to believe akratically, i.e., to have a belief that we believe to be irrational. But if we are sometimes

²¹ On Horowitz's (2014) view, "an epistemically akratic agent believes something that she believes is unsupported by her evidence." On Titelbaum's (2015) view "no situation rationally permits any overall state containing both an attitude A and the belief that A is rationally forbidden in one's current situation."

²² Following Worsnip (2018), we might consider Enkrateia to be a *coherence* requirement, and consider Steadfast and Higher-Level Influence to be *evidential* requirements. Yet as long as rationality requires both coherence and responsiveness to evidence, the tension between the three theses remains.

required to retain a belief that P despite HOE that it is irrational (per Steadfast), and at the same time required to believe that believing P is irrational (per Higher-Level Influence), then we are sometimes required to believe P and also believe that believing P is irrational (contra Enkrateia). So Steadfast, Higher-Level Influence, and Enkrateia are jointly inconsistent.²³ Since Higher-Level Influence would seem to follow from any sensible account of testimony, Enkrateia and Steadfast are in potential trouble.

Rejecting Enkrateia comes at a high cost. Beyond its intuitive appeal, forceful considerations against its rejection are readily available. Horowitz (2014) argues that views that allow akratic combinations of doxastic attitudes “license patently bad reasoning and irrational action.” For example, consider a variant of a case by Horowitz:

Akratic Detective

Detective correctly assesses her evidence E and believes that Suspect committed the crime. Detective then acquires strong but misleading HOE, which suggests that E supports Suspect’s innocence. She reasons as follows: “Suspect is guilty, as I believe based on E. But my HOE suggests that E does not support Suspect’s guilt. So the evidence E on which I’m basing my belief is misleading. Nevertheless, based on E, I believe that Suspect committed the crime and must be arrested!”

Two aspects of this case are disconcerting. The first is Detective’s inference that her evidence E is misleading. The second is Detective’s maintaining her view despite basing it only on evidence that she considers to be misleading, and despite believing that this

²³ Others have discussed versions of this inconsistency. See Huemer (2011), Christensen (2013), Horowitz (2014), and Worsnip (2018).

belief is irrational. Views on which rationality both tolerates akratic attitudes and lets agents reason and act based on their rationally held beliefs seem committed to these implications.

In light of the initial plausibility of Enkrateia and Higher-Level Influence, and given their inconsistency with Steadfast, proponents of Steadfast appear to owe a story about which of the two theses we should reject and why. A few ways of rejecting Enkrateia are found in the literature, and may be combined with the account defended so far.²⁴ In what comes next I offer a way to resist Higher-Level Influence that builds on Michael Titelbaum's (2015) approach to doing so. The idea is that Enkrateia is already in tension with Higher-Level Influence given some plausible assumptions, and so proponents of the former should reject the latter regardless of whether Steadfast is right.

(5.1) Level-connection

Proponents of Enkrateia should be sympathetic to a rather strong level-connection principle. They deny that we could rationally believe P , disbelieve P , or suspend about P , and at the same time rationally believe that our attitude toward P is irrational. This implies that whatever attitude a we should have toward P , we should either believe that a is rational, or suspend judgment about whether a is rational.²⁵ But it is arguably another form of epistemic akrasia to have an attitude and also suspend judgment about its rationality. Both suspension and disbelief that an attitude a is rational do not sit well with having a at the same time. Michael Huemer (2011) and Declan Smithies (2012) make the intuitive case for this claim. Huemer takes suspension about P to license the

²⁴ See Weatherson (ms.), Williamson (2011, 2014), Coates (2012), Lasonen-Aarnio (2014), and Worsnip (2018).

²⁵ This presupposes that we are justified in having at least one doxastic attitude toward the proposition that our attitude a is rational. On Smithies's (2012) Exhaustiveness and Turri's (2012) Optimism we are justified in having at least one doxastic attitude toward any proposition.

assertion *it may or may not be the case that P*. Similarly, Smithies takes suspension about P to license the assertion *it is an open question whether P*. Accordingly, if we suspend about whether we are justified in believing P, we could say that *we may or may not be justified in believing P*, and that *it is an open question whether we are justified in believing P*. Both assertions appear to fit poorly with asserting P at the same time.²⁶ So, much like disbelief that we believe rationally, suspension about whether we believe rationally seems to involve irrationality.

What follows is that whenever we are rationally justified in having a doxastic attitude *a* toward a proposition P, we may not disbelieve that *a* is rational nor suspend on the matter. Instead, no matter what attitude *a* is rational for us to have, we are justified in believing that *a* is rational:

Level-Connection: It is rational for one to have a doxastic attitude *a* toward a proposition P only if one has sufficient justification to believe that having *a* toward P is rational.²⁷

Using Level-Connection we can now give an argument not entirely unlike the one offered earlier, but this time against HOE-based higher-level belief formation. Level-Connection tells us that no matter what attitude *a* is rational for us to have toward P, we possess sufficient justification to believe that *a* is rational.²⁸ In other words, we are

²⁶ See Hazlett (2012) for the opposing view.

²⁷ Kelly (2010) offers some Enkrateia-independent reason to think that a level-connection principle obtains. He argues that recognition of the import of our evidence is the reason we form the paradigmatically rational beliefs that we do. It is not a major leap from this observation to thinking that a necessary requirement on rational belief is that agents recognize what their evidence supports. Moreover, as Kelly argues, recognition that such-and-such entails justification to believe that such-and-such.

²⁸ Kiesewetter (2016) briefly notes this.

necessarily in a position to rationally believe the truth about whether a is rational in our situation. Forming our higher-level beliefs with sensitivity to HOE would discount this favorable state, as the HOE could misleadingly suggest that a is irrational. It therefore stands to reason that we may not contaminate what must be justification for the truth about whether a is rational by including possibly misleading HOE.²⁹ As before, the convenience that the HOE might afford is a prudential consideration at best.

This suggestion pairs well with Titelbaum's (2015) view of why we should resist HOE. On that view, Enkrateia is best explained by our having empirically indefeasible justification for believing what attitudes are rational in what situations. However, the explanation does not tell us how this indefeasible justification comes about. One option is that the sufficient justification we have regarding what attitudes are rational is maximal and calls for higher-level certainty. This line would make it clear why we should dismiss any misleading HOE as misleading, and why non-misleading HOE could not affect the required higher-level attitude. Another option is that the sufficient justification we have about what attitudes are rational is not maximal, but somehow *insulated*. The idea is that something about our epistemic state requires us to form beliefs only on the basis of that justification and nothing else.

Admittedly, both options are *prima facie* suspicious. The first is suspicious because nothing about Enkrateia clearly warrants thinking that our higher-level justification is maximal. There is distance between the claim that we may not have false beliefs about what rationality requires and the claim that we should be certain about what rationality requires. The second is suspicious because it is unclear what could justify insulation from inputs like HOE, which speak directly to the higher-level proposition that we care

²⁹ I express a similar thought in Tal (2018).

about. Yet both options become more palatable once we observe that we necessarily have sufficient justification for the truth about what attitudes are rational.

Knowing that our possessed justification is for the truth is a highly unusual and significant bit of knowledge. It lets us know that all we have to do to reach a true belief is correctly assess our justification. As I suggested, it makes sense that in such unique circumstances rationality would require us to do away with any input that might be misleading, so as to preserve our access to the truth. This is reason to think that our justification about what we should believe is insulated, even if not maximal.³⁰

The observation also supports the possibility that our justification about what attitudes are rational is maximal. Other occasions on which we know that our justification is for the truth are ones where our justification appears to be maximal. One example is our justification for claims in mathematics. It does not matter if a highly reliable friend tells us which of two negative numbers is greater. We know that we have independent justification for the truth about which is greater, and no testimony can outweigh that justification. This is right even if negation symbols tend to confuse us about which number is greater, and we are more likely to reach the truth by relying on the testimony. We should be certain about which number is greater regardless of our mathematical competence and regardless of who testifies to what. Perhaps another

³⁰ To avoid falling prey to the Dogmatism paradox, it must be that we do not lose knowledge that our justification supports the truth upon gaining HOE. If we lost that knowledge, we would no longer be justified in using our preexisting justification alone. Fortunately, given Level-Connection, whatever justification we have at the higher level must be for the truth. No HOE could undermine our knowledge of this fact. Relatedly, it must also be that we do not lose knowledge of what our evidence is upon receiving HOE (thanks to Ram Neta for this point). However, we should bracket issues concerning doubts about what our evidence is for present purposes. The main reason for doing so is that many paradigm instance of HOE do not undermine our knowledge of what the evidence is. For example, a disagreeing peer may well agree that the evidence is indeed what we take it to be, but disagree about its upshot. The question of what to believe when we are uncertain about what our evidence is seems separate from the question of what to believe when we are uncertain about whether we believe rationally.

example is our justification for claims about how things seem to us. Our seemings constitute our justification for the way things seem to us, and so must support the truth about the way things seem to us. Here too certainty about the way things seem to us is rational, and additional inputs on the matter are dismissible. So it would be par for the course of necessarily having justification for the truth that the justification was maximal.

To sum up, the aim here was to show that Enkrateia and Higher-Level Influence are in tension, and without presupposing Steadfast (on pain of begging the question). If Enkrateia clearly implied that we have indefeasible justification for the truth about what rationality requires, then that tension would be easy to show. But the implication is not so clear. Only once we tie Enkrateia to Level-Connection can we tell a convincing story as to why Enkrateia goes hand in hand with our having such indefeasible justification. Level-Connection acts as a bridge between the enkratic constraint and the view that HOE does not affect what higher-level beliefs we should have. So, we can avoid the accusation that Steadfast forces us to reject one of two intuitively plausible theses. We should reject either Enkrateia or Higher-Level Influence regardless.³¹

Lastly, one might worry that dismissing HOE so extensively leaves us with a view that is vulnerable to reasoning-distortion cases.³² For suppose that S learns that something (a drug, lack of oxygen, etc.) has very likely distorted her reasoning, so as to

³¹ It is also possible to reject my presupposition of a thesis like Optimism, and thus block the inference from Enkrateia to Level-Connection. It would be an interesting result of the argument here if we had to give up Optimism to resist it. That said, notice two considerations against such a move. First, the kind of Optimism that the inference relies on is quite modest. Our evidence need not justify some doxastic attitude relative to every possible proposition, but only relative to propositions of the form *attitude a is rational*. Second, rejecting Optimism would not be in the spirit of Enkrateia. Just as it seems inconsistent to assert *P but believing P is irrational*, it seems inconsistent to assert *P but despite considering the matter I have no view about whether believing P is rational*.

³² See Christensen (2016), Schoenfield's (2016). Thanks to James Fritz for suggesting this objection to steadfastness at all levels.

make her perform irrational inferences that seem perfectly rational to her. On the proposed steadfast view, the agent who rationally believes P could rationally believe that she is rational, and rationally infer that she was immune to the distortion. This seems irrationally narcissistic.

There is room to wonder what precisely is involved in performing irrational inferences that seem perfectly rational to us, and whether our intuitions about the case track prudent belief rather than rational belief. But beyond that, cases like this appear to assume equivalence between those whose reasoning is affected and those who are unaffected. The idea is that the unaffected agent who believes that she is rational should worry that everything would seem exactly the same if she were affected and irrational. Yet if everything would indeed seem *exactly* the same to the affected agent, then that affected agent would have misleading justification about what rationality requires. Not only would such equivalence undermine the intuition that affected agents are irrational, it would also beg the question against Level-Connection. As long as we grant Level-Connection, and as long as we hold fixed that the affected and unaffected differ in how rational they are, there would be some accessible symmetry-breaker for the unaffected agent to appeal to.³³

³³ I am grateful to the participants of the 2019 Chapel Hill Normativity Workshop, as well as to anonymous *Philosophical Studies* referees, for offering insightful feedback. Special thanks goes to Stewart Cohen and Juan Comesaña for their help with developing the ideas in this paper.

References

- Alexander, D.** (2013). "The Problem of Respecting Higher Order Doubt". *Philosophers' Imprint* 13(18): 1–12.
- Berker, S.** (2013). "The Rejection of Epistemic Consequentialism". *Philosophical Issues* 23: 363–387.
- Boult, C.** (2017). "Epistemic normativity and the justification-excuse distinction". *Synthese* 194(10): 4065–4081.
- Christensen, D.** (2009). "Disagreement as Evidence: The Epistemology of Controversy". *Philosophy Compass* 4(5): 756–767.
- (2010a). "Higher-Order Evidence". *Philosophy and Phenomenological Research* 81(1): 185–215.
- (2010b). "Rational Reflection". *Philosophical Perspectives* 24: 121–140.
- (2013). "Epistemic Modesty Defended". Jennifer Lackey and David Christensen (Eds.), *The Epistemology of Disagreement: New Essays*. Oxford: Oxford University Press, 77–97.
- (2016). "Disagreement, Drugs, etc.: From Accuracy to Akrasia". *Episteme* 13(4): 397–422.
- Coates, A.** (2012). "Rational epistemic akrasia". *American Philosophical Quarterly* 49: 113–124.
- Cohen, S.** (2013). "A Defense of the (almost) Equal Weight View". Jennifer Lackey and David Christensen (Eds.), *The Epistemology of Disagreement: New Essays*. Oxford: Oxford University Press. 98–120.
- Fantl, J. & McGrath, Matthew.** (2011). 'Pragmatic Encroachment', *Routledge Companion to Epistemology*, (eds.) S. Bernecker & D. H. Pritchard, 558–68, London: Routledge.
- Feldman, R.** (2009). "Evidentialism, Higher-order Evidence, and Disagreement". *Episteme* 6: 294–312.
- (2014). "Evidence of Evidence is Evidence". In *The Ethics of Belief*, ed. John Mattheson and Rico Vitz. 284–99.
- Field, H.** (2000). "Apriority as an Evaluation Notion". In Boghossian and Peacocke (eds.) *New Essays on the A Priori*. Oxford University Press. 117–149.
- Friedman, J.** (2017). "Why Suspend Judging?". *Noûs* 51(2): 302–326.
- Hájek, A.** (1998). "Agnosticism meets Bayesianism". *Analysis* 58: 199–206.
- Hazlett, A.** (2012). "Higher-Order Epistemic Attitudes and Intellectual Humility". *Episteme* 9(3): 205–223.
- Horowitz, S.** (2014). "Epistemic Akrasia". *Noûs* 48(4): 718–744.
- Huemer, M.** (2011). "The Puzzle of Metacoherence". *Philosophy and Phenomenological Research* 82(1): 1–21.
- Kelly, T.** (2010). "Peer Disagreement and Higher-Order Evidence". In T. Warfield and R. Feldman (Ed.), *Disagreement*. Oxford University Press. 111–175.
- Kiesewetter, B.** (2016). "You ought to φ only if you may believe that you ought to φ ". *The Philosophical Quarterly* 66(265): 760–782.
- Lasonen-Aarnio, M.** (2010). "Unreasonable Knowledge". *Philosophical Perspectives* 24(1): 1–21.
- (2014). "Higher-Order Evidence and the Limits of Defeat". *Philosophy and Phenomenological Research* 88: 314–345.
- Littlejohn, C.** (forthcoming). "A Plea For Epistemic Excuses". In Dorsch, F. and Dutant, J., editors, *The New Evil Demon*. Oxford University Press, Oxford.
- Nelson, M. T.** (2010). "We Have No Positive Epistemic Duties". *Mind* 119(473): 83–102.
- Raleigh, T.** (forthcoming). "Suspending is Believing". *Synthese*. Doi 10.1007/s11229-019-02223-8.
- Rosenkranz, S.** (2007). "Agnosticism as a Third Stance". *Mind* 116(461): 55–104.
- Smithies, D.** (2012). "Moore's Paradox and the Accessibility of Justification". *Philosophy and Phenomenological Research* 85(2): 273–300.
- Schechter, J.** (2013). "Rational Self-Doubt and the Failure of Closure". *Philosophical Studies* 163(2): 429–452.
- Schoenfield, M.** (2015). "Bridging Rationality and Accuracy". *Journal of Philosophy* 112(12): 633–657.
- (2016). "An Accuracy Based Approach to Higher-Order Evidence". *Philosophy and Phenomenological Research* 96(3): 690–715.
- Sliwa, P. & Horowitz, S.** (2015). "Respecting All the Evidence". *Philosophical Studies* 172(11): 2835–2858.

- Tal, E.** (2018). "Self-Intimation, Infallibility, and Higher-Order Evidence". *Erkenntnis* doi: 10.1007/s10670-018-0042-4.
- Titelbaum, M.** (2015). "Rationality's Fixed Point (Or: In Defense of Right Reason)". In J. Hawthorne (ed.) *Oxford Studies in Epistemology* 5. Oxford University Press. 253–294.
- Turri, J.** (2012). "A Puzzle About Withholding". *Philosophical Quarterly* 62(247): 355–364.
- Weatherson, B.** (ms). "Do judgments screen evidence?" Unpublished manuscript.
- Williamson, T.** (2011). "Improbable knowing". In T. Dougherty (Ed.), *Evidentialism and its Discontents*, Oxford: Oxford University Press: 147–164.
- (2014). "Very Improbable Knowing". *Erkenntnis* 79: 971–999.
- Worsnip, A.** (2018). "The Conflict of Evidence and Coherence". *Philosophy and Phenomenological Research* 96 (1): 3–44.