5-2024

Probability Assignments Under Deep Uncertainty

Hexuan Huang
Bates College, hhuang@bates.edu

Probability Assignments Under Deep Uncertainty

An Honors Thesis

Presented to The Department of Philosophy

Bates College

In partial fulfillment of the requirements for

the Degree of Bachelor of Arts

By Ricky Huang 黄贺宣

Lewiston, Maine

April 1, 2024

Acknowledgement

When I was in middle school, I learned a story.

> During his first imperial examination in the capital, Jia Dao spontaneously composed a poem on a donkey's back. One of the lines goes: "Birds lodge in trees by the pond, a monk pushes on the door under the moon." Jia Dao was indecisive between the characters "推” (to push) and "敲“(to knock) and spent the next couple of days reciting on the donkey, making pushing and knocking gestures with his hands. Han Yu, a famous Chinese poet, ran into Jia Dao. He advised him to use the word 敲 (to knock), as it is a crisp sound contrasting the tranquil theme of the poem. They rode home together, discussing poetry for days, and forged a lasting friendship. Now in mandarin, 推敲 (to push or to knock) represents the mental state of pondering.

The people I would like to thank helped me truly appreciate this story by pushing for intellectual rigor and careful reflections in Philosophy:

Thank you to Professor Haley Schilling for supervising this project. Your feedback/guidance have been immensely useful. Thank you to all the Philosophy professors I’ve had the pleasure of taking courses with in my undergraduate years: Lizzie Fricker, Lauren Ashwell, Adriana Clavel-Vázquez, Roger Teichmann, Michael Murray, Jonathan Cohen, Joshua Pike, Mike Dacey, Susan Stark. All of these courses shaped who I am today and how I conceive the world. I really enjoy doing philosophy, and I wish to continue down this journey. Thank you to the many members of the Effective Altruist community who have provided extensive feedback to this thesis.

Thank you to my family and friends, through your eyes I see the world. 谢谢。

Table of Contents

Probability Assignments Under Deep Uncertainty

Abstract:

We assign probabilities to represent our epistemic states about future events. We do this for tomorrow's weather prediction and timelines of Superintelligent A.I. systems. We may hold the same probability assignment for different events, but yet something feels different about these probabilities. The second case is oftentimes labeled as a case of deep uncertainty. This paper will offer an epistemic account of deep uncertainty with an appeal to higher-order uncertainty. It aims to be theoretical rather than prescriptive. I first defend the **Higher-Order Uncertainty Thesis (HOU)**: S has deep uncertainty towards P(A)=q, S's probability assignment *q* for some proposition A, iff S has not responsibly integrated higher-order uncertainty q*, which is below some threshold some threshold q**, towards our probability assignment. I then defend two thesis about the upshots of my argument: **Belief Thesis:** If we have deep uncertainty towards P(A)=q, then we should not believe that q represents the rational probability of A given higher-order uncertainty. **Assertion Thesis:** If we have deep uncertainty towards P(A)=q, then we should not assert P(A)=q.

Introduction

Anyone familiar with the Effective Altruism community should have seen some alarming predictions made over the prospect of humanity:

"*Existential catastrophe within next 100 years: ~ 1 in 6.* " (2020, p. 163)

Toby Ord, senior researcher at the Future of Humanities Institute

(On power-seeking A.I. leading to human extinction) *"estimate of ~5% that an existential catastrophe of this kind will occur by 2070. May 2022 update: since making this report public in April 2021, my estimate here has gone up, and is now at >10%*" (2022)

Joseph Carlsmith, senior research analyst at Open Philanthropy

"*$100 billion spent on reducing existential threats could reduce it by over 1% over the next century*" (2022)

Benjamin Todd, Founder of 80,000 Hours

Readers are oftentimes shocked over how gloomy these predictions are, and then ask the question: "where did these numbers come from?" The epistemic story is that we are assigning probabilities to represent our subjective likelihood of an event occurring or for a proposition to be true. These numbers represent our best judgment given available evidence. Subjective probabilities are distinct from chance, where the latter is a physical concept independent of what anyone thinks or holds evidence towards (Hájek, 2002). We then express these probability assignments through two distinct epistemic states: beliefs or credences. Belief is a binary epistemic state, where to believe is to regard a proposition as true or take it to be the case (Schwitgebel, 2019). Credence represents a more graded attitude we hold towards a proposition, where we represent our degree of belief on the [0,1] interval (Jackson 2020). For example, for the same probability assignment, I could express my epistemic state as either "I believe the probability of snow tomorrow is 30%" or "I hold a credence of 0.3 that it will snow tomorrow." [1] We do this for a wide range of scenarios, from estimating the likelihood of various outcomes in our personal lives and world events.

Here's an interesting problem: Suppose I assign a 30% probability to P1: Humanity will develop superintelligent A.I. systems in the next 100 years. I also assign a 30% probability to P2: There will be an empty parking space tomorrow morning. Let's also assume that I have evidence supporting my confidence level in both propositions. Yet, something *feels* different about our probability assignments in both cases as we feel that we are more uncertain about our assignment for P1 than P2. This "additional" uncertainty isn't represented in our degrees of belief, as we hold the same probability assignment towards each proposition for both propositions. Conventional wisdom kicks in here. It seems really difficult to predict what will happen in 100 years. Maybe it is more epistemically difficult to assign probabilities huh. Oc to some propositions than others, then we should hold an "additional" uncertainty attitude towards these propositions. If I'm correct here that there are meaningful differences between our probability assignments across propositions, then this has clear normative upshots over our ability to compare and act on our probability assignments.

Here's an existing angle offered outside of philosophy: Some economists think that there are some propositions we cannot assign probabilities towards. They call this Knightian uncertainty, or uncertainty where we have"no scientific basis on which to form any calculable probability" (Sunstein 2023). They then distinguish between risks and uncertainties. If we have

[1] Usually, the difference between belief of probability assignment and credence is that there is a belief-like attitude for belief (and absent in credence) and they hold different relations to assertino/knowledge. Moss contends this in *Probabilistic Knowledge,* where she claims that we hold a simple attitude (belief) towards complex content (.8 probability of x) when we formulate a credence.

"in mind well-defined probabilities on possible outcomes", then it is the former; if not, the latter. (LeRoy & Singell, 1987, pg. 395). Some decision theorists shares this school of thought. Proponents of Decision Making Under Deep Uncertainty (DMDU) argue that there is a difference between acting under deep uncertainty and ordinary uncertainty, where predicting P1 and P2 fall under different categories of uncertainty. The distinction here is driven by practical concerns: scientists identify that there are some future events where it is very difficult to assign probabilities to the likelihood of events occurring (Marchau et al., 2013). Classic examples of deep uncertainty identified in the field include predicting far-future technological progress, assessing future impacts of climate change, and understanding implications of deeply complex financial model changes (Marchau et al., 2013). Yet, agents still need to make decisions. DMDU comes in to help identify robust and adaptable decisions amidst such uncertainty (Stanton & Roelich, 2021).

I'm not concerned about whether or not these uncertainties should be labeled as Knightian or deep uncertainty. I'm also not interested in the prescriptively classifying specific events as deep or ordinary uncertainty. This paper concerns the normative question: Are economists and decision theorists hinting at interesting epistemic distinctions with categorizing some of our uncertainty as Knightian or deep? Very little, if anything, has been written by epistemologists on the topic of deep uncertainty. I'm interested in building the epistemic story on how we got these probability assignments and whether or not they are meaningfully distinct from our probability assignments in ordinary circumstances.

Two interesting unanswered questions:

(A) *Definition Question***:** What is deep uncertainty? For the sake of clarity, I will label the uncertainty identified above as possible cases of deep uncertainty. I'm less concerned with whether or not my view precisely represents the cases offered in the DMDU literature, and more concerned with what warrants the distinction (if at all) between deep and ordinary uncertainty.

(B) *Epistemic Privilege Question*: If it is true that deep uncertainty differs from ordinary uncertainty, does that warrant different epistemic states towards our probability assignments (i.e. we should only believe our probability assignments under ordinary uncertainty). Subsequently then, do we hold the same epistemic privileges e.g,. the ability to assert, testify, and conduct expected value calculations for both types of uncertainty? In absence of these epistemic privileges, how should we express our epistemic attitudes? If we cannot assert our probability assignment, what should we do?

In this essay, I hope to offer an account for the first question, and hint at interesting arguments for the second. I attempt to answer these two questions by relating deep uncertainty with theories in epistemology about higher-order uncertainty. [2]

Our evidence which raises or lowers our subjective probability towards a proposition is is divided between our first-order evidence and our higher-order evidence. Roughly, first-order evidence is "evidence directly related to our beliefs– it's evidence that makes the content of your belief more (or less) likely to be true"(Ye, 2022, pg 1). For example, meteorological evidence about Troposphere status when we are making rain predictions. Subsequently, first-order uncertainties are uncertainties about our propositions given our first-order evidence. Higher-order uncertainties are uncertainties over how confident we should be towards our

[2] For clarity sake, this essay will adopt Dorst's reframing of Higher-order Evidence debate into higher-order uncertainty. I broadly agree with the same motivations listed in Dorst (2018) over why this reframing offers conceptual clarity, also see Hedden & Dorst (2022) on *Why (almost) all evidence is higher-order evidence.* My arguments aren't particularly reliant upon which way you frame the higher-order uncertainty/evidence debate, as one can easily rewrite my thesis into a higher-order evidence thesis without losing its core motivations.

probability assignments. These are caused by evidence not directly related to the content of your belief, but about your belief-forming process or evidence about your first-order evidence (Ye, 2022). Examples of higher-order uncertainty include uncertainty over how well you can reason in massively complex cases given cognitive limitations or evidence about evidence that we are missing evidence.

I think differences in higher-order uncertainty is why it is far harder to assign probabilities for P1 than P2, and subsequently is why we feel more certain about our probability assignment for the latter despite holding the same probability assignment. I argue that thinking through the deep uncertainty problem through the higher-order uncertainty framework provides both a clear definition on deep uncertainty and elucidates its epistemic upshots.

In this essay, I'll attempt to answer the *Definition Question, Epistemic Privilege Question* by defending two claims. I'll first offer a brief overview of my claims and argument structure.

I begin with answering the definition question in Part 1:

In 1.1, I offer an exegesis over existing definitions of deep uncertainty/knightian uncertainty. A lot of the exegesis requires deliberate reconstruction of the view, as it uses very different language than Epistemologists. I'll conclude in 1.1 what is missing from existing literature.

1.2 - 1.4 will be defending HOU:

Higher-Order Uncertainty Thesis (HOU): S has deep uncertainty towards P(A)=q, S's probability assignment *q* for some proposition A, iff S has not responsibly integrated higher-order uncertainty q*, which is below some threshold q**, towards our probability assignment. *Formally:* S has deep uncertainty towards P(A)=q iff P(P(A)=q)=q*, q* <q**.

HOU draws equivalence between higher-order uncertainty and deep uncertainty with two exceptions: cases where higher-order uncertainty is responsibly integrated into our probability assignment, and cases where our higher-order uncertainty is very little and below some threshold. Then, a defense of HOU will be composed of three parts:

1.2 will outline the broad motivations for relating deep uncertainty to higher-order uncertainty by comparing *A.I. Timeline* (deep uncertainty) and *Weather Predictions* (ordinary uncertainty). I'll then list three differences between these two cases (cognitive limitations, evidence about evidence, peer disagreements) and argue that these differences are due to higher-order uncertainty. This will motivate the idea that deep uncertainty is a type of higher-order uncertainty.

1.3 - 1.4 will spell out the two exceptions; higher-order uncertainty below a certain threshold shouldn't count as deep uncertainty, and responsibly integrated higher-order certainty into our probability assignment shouldn't count as deep uncertainty.

1.3 will defend the *threshold* condition, $q^* > q^{**}$, to the higher-order uncertainty thesis. The motivation for adding a threshold condition is that we have higher-order uncertainty towards most of our ordinary propositions, but we wouldn't want these ordinary propositions to be counted as cases of deep uncertainty.

1.4 will defend the *responsible integration* condition to the higher-order uncertainty thesis. The motivation for adding an integrated condition is that sometimes our higher-order uncertainties are responsibly integrated into our probability assignments. In these cases, we are quite clear what our probability assignment should be despite previously having higher-order uncertainty. This doesn't seem like cases of deep uncertainty.

I then move on to answering the epistemic privilege question in part 2:

In 2.1, I argue that if HOU is true, then Belief Thesis is true.

Belief Thesis: If we have deep uncertainty towards P(A)=q, then we should not believe that q represents the rational probability of A given higher-order uncertainty.

I defend the **Belief Thesis** with three claims: (a) Integrating our higher-order uncertainty into our probability assignment requires epistemic self-trust over some higher-order reflection process (b) Epistemic self-trust over higher-order reflection process is absent in cases of deep uncertainty (c) Believing that q represents rational probability of A requires our higher-order uncertainty being integrated into our probability assignment.

In 2.2, I use **Belief Thesis** to prove the **Assertion Thesis**

> **Assertion Thesis*:*** If we have deep uncertainty towards P(A)=q, then we should not assert P(A)=q.

I defend the **Assertion Thesis** by relating it with existing literature on norms of assertion. (Knowledge Norms of Assertion & Justified Belief Norms of Assertion) all require belief as a necessary condition for assertion. **Belief Thesis** shows that we should not believe q represents a rational probability of A, so we shouldn't assert P(A)=q.

Part 1: Definition Question

1.1 Existing definitions of deep uncertainty

Existing literature on deep uncertainty gestures around possible traits of deep uncertainty without getting too much into the distinction. Some say deep uncertainty "represents the deepest level of recognized uncertainty; in this case, what is known is only that we do not know" (Walker et al., 2013). The examples the literature offers are "black-swan" events like the concatenation of events following the 2007 subprime mortgage crisis or level 9.0 earthquakes in Japan in 2011. (Taleb, 2017). This is echoed in the Knightian Uncertainty literature, where it is labeled as the "unknowable unknown".

The claim that our only knowledge about the situations is that we do not know anything about them beforehand is a very strong one – and a claim that is nearly impossible to defend. It seems plausible that we knew something about the aftermath of the financial crisis, even if we do not know its full scale and event details. A more charitable read of this would suggest that we formulate limited beliefs about the situation and we have low credences in some of the probability assignments we have about the situation.

Other views are more comprehensive:

> "Deep uncertainty—that is, where analysts do not know, or the parties to a decision cannot agree on, (1) the appropriate conceptual models that describe the relationships among the key driving forces that will shape the long-term future, (2) the probability distributions used to represent uncertainty about key variables and parameters in the mathematical representations of these conceptual models, and/or (3) how to value the desirability of alternative outcomes."

(Lempert et al., 2003, pg xvii)

It is unclear if (1), (2), and (3) are all necessary conditions for deep uncertainty. I would imagine that that's not what the text meant. It is also unclear if each condition is independently sufficient. A charitable interpretation of the three conditions would read them as general characteristics of deep uncertainty rather than precise conditions defining it. By (1), I imagine what they mean is having some understanding over what are the causes for the long term future to change. (3) seems like a valid characteristic of deep uncertainty, but perhaps is better addressed under the moral uncertainty literature (See MacAskill et al., 2020) for more on moral uncertainty).

The most important condition is (2). This is echoed in other deep uncertainty literature. They identify deep uncertainty as uncertainty which cannot be "represented", "quantified", or "eliminated"(Leach etal., 2010; Lempert etal., 2003). According to this view, an important characteristic of deep uncertainty is that we do not know, or agree on the probability distribution used to represent uncertainty about key variables and parameters.

There are a couple of concerns I have with this view, despite sharing its broad motivation. First, "knowing" probability distribution seems too loose as a condition for deep uncertainty if we consider the relationship between credence and knowledge. Outside of probabilistic knowledge frameworks, it is commonly accepted in epistemology that knowledge requires full-beliefs (Moss 2019, pg 1). Then, if I hold a credence of 0.8 towards some proposition, these credences cannot be knowledge, even if these credences are very well justified and supported by good evidence. Then this definition of deep uncertainty naturally includes all of our credence-based probability assignment, even in situations where we are pretty sure about. This seems too inclusive as a definition and requires further justification on why these cases of well-justified credences should be counted as deep uncertainty cases.

Second, I think it is unclear what this means. There are two plausible interpretations of this. One, it is epistemically impossible to assign probabilities for these events. Two, it is epistemically irresponsible to assign probabilities for these events. The first explanation needs to justify why it is epistemically impossible for us to assign probabilities. Subjective Bayesians believe that we can hold credences for any proposition, as long as our credences coheres with basic Kolmogorov axioms where P is a non-negative real number, P(omega)=1, p follows additivity (See de Finetti 1974). Objective Bayesians believe that we can hold credences for any propositions where, in addition to coherence, we are free from bias and avoid overly strong opinions for our prior credences (Rosenkrantz 1981; J. Williamson 2010). Just because we are more uncertain about *A.I. Timeline* than *Weather Prediction* does not mean we are unable to assign probabilities for propositions about the former. The second explanation hints at valid motivations: we often question whether or not our probabilities for far-future events came from a reliable epistemic process or are factive. I think a charitable read would suggest that there are some propositions where it is far harder to assign probabilities than other propositions. However, this view needs to be flashed out on what this means and how we should draw the line on deep uncertainty. Having clear cut reasons help us justify the distinction between deep and ordinary uncertainty.

I think broadly then, both views present valid characteristics of deep uncertainty. There are some propositions where it is difficult for us to assign probabilities or rank in perceived likelihood. I think this difference is because of (1) the higher-order uncertainty we hold towards these propositions (2) these are special kinds of higher-order uncertainty where it is very difficult to reflect them in our probability assignment. In the following sections, I will present a thesis which will hopefully formalize these motivations and define deep uncertainty in a more accurate manner.

1.2 Higher-order uncertainty and deep uncertainty

This outlines the broad motivations for relating deep uncertainty to higher-order uncertainty by comparing *A.I. Timeline* (deep uncertainty) and *Weather Predictions* (ordinary uncertainty). I'll then list three differences between these two cases (cognitive limitations, evidence about evidence, peer disagreements) and argue that these differences are due to higher-order uncertainty. This will motivate the idea that deep uncertainty is a type of higher-order uncertainty.

> ***A.I. Timeline****:* Let's suppose you are a machine learning scholar trying to assign probabilities for whether or not humanity will develop a superintelligent A.I. system in 100 years. Let's suppose that you have good first-order evidence supporting your probability assignment. Available first-order evidence is modeling predictions that have been accurate in the past, experimental data that scaling law will hold true in the future, mathematical proofs that there shouldn't be any bottlenecks with A.I. development in the foreseeable future. You assign a 30% probability to the proposition: superintelligent A.I. will develop in 100 years.
>
> ***Weather Prediction:*** Let's suppose you are a climate scientist trying to assign probabilities for whether or not it will snow tomorrow. Let's suppose you have good first-order evidence supporting your probability assignments. Available first-order evidence is modeling predictions that have been accurate in the past, data about

atmosphere conditions etc. You assign a 30% probability to the proposition: It will snow tomorrow.

I think there are three differences between *A.I. Timeline* and *Weather Prediction:*

(1) Difference in complexity and cognitive limitations. In *A.I. timeline*, there are more future events for one to consider to formulate an accurate probability assignment. Not only do we have to take account of numerous parameters (such as hardware advancements, algorithmic advances, regulatory factors), we also have to attempt to imagine how different futures will play out in the interim future (say next 50 years) to make an accurate judgment over A.I. timeline in the next 100 years. We also have to imagine how changes to each parameter influences other parameters, often in ways that are very difficult to assess. One would imagine that our epistemic process becomes less reliable the more complex the event becomes, and that there are limitations to the information we are able to cognitively process. In *Weather Prediction*, despite having many parameters, we still have to make less difficult future assessments about complex scenarios for our probability assignments. One would imagine that our epistemic process becomes less reliable the more complex the event becomes.

(2) Difference in evidence about evidence. In cases of A.I. Timeline, our evidence should be seen as less reliable, because we are using evidence about the present to assess the far future, where perhaps situations are not entirely analogous. There are also often cases where we have evidence that we are missing evidence. E.g., evidence suggesting that we do not understand machine learning sufficiently. Therefore, we might be missing key evidence. In cases of Weather Predictions, there is less evidence suggesting that our evidence isn't reliable.

(3) Difference in peer disagreement. In cases of A.I. Timeline, there is significantly more disagreement from epistemic peers, each with good evidence supporting their probability assignments. Examples of this looks like A.I. expert surveys on the earliest that machines will be able to simulate learning and every other aspect of human intelligence, 18% responded $\leq$ 50 years, 41% responded >50 years, 41% responded with never (Müller & Bostrom, 2016, pg 6). In cases of Weather Predictions, there's significantly less peer disagreement, at least for short-term weather predictions.

(1), (2), (3) are all evidence that we should be uncertain about our opinion, or that we have greater higher-order uncertainty towards our probability assignment for *A.I. Timeline* than *Weather Prediction*. These are not evidence directly related to our probability assignment, e.g., whether I have cognitive limitations should have no influence on the probability of A.I. development. Instead, these are evidence about our ability to make probability assignments. I'm not making an argument that there are no cases of complexity, evidence about evidence, or peer disagreement in *Weather Prediction*, but rather there is significantly more of these evidence in *A.I. Timeline* than *Weather Prediction*. Then the difference between *A.I. Timeline* isn't a difference in first-order evidence, but rather a difference in higher-order uncertainty.

Let's formalize this idea of higher-order uncertainty:

P(A)=q represents probability assignment q I should hold for some proposition A, given my total evidence. We do this by assigning credences to proposition [A is true], credences to proposition [A is not true]. We then assign positive probabilities to each proposition P(A is true)=q, P(A is not true)=1-q. For example, by saying that there is a 60% probability that it will rain tomorrow, we are saying that P(it will rain tomorrow) =0.6, P(it will not rain tomorrow)=0.4. We may represent P(it will rain tomorrow)=0.6 as P(A)=q, which represents our probability assignment q towards A. The higher q is, the more certain I am about the proposition.

Higher-order uncertainty examines to what extent I should be certain over my probability assignment P(A)=q. We examine proposition [I-should-be-confident-that-P(A)=q] and proposition [I-should-not-be-confident-that-P(A)=q] (Dorst 2019, pg 7). We then assign positive probabilities to each proposition. [I-should-be-confident-that-P(A)=q] as P(P(A)=q)=q* and I-should-not-be-confident-that-P(A)=q] as P(P(A)=q) = 1-q*. q* represents our probability assignment for our higher-order uncertainty. The higher q* is, the lower my higher-order uncertainty, and the more certain I am that I should be confident about my probability assignment. The lower q* is, the higher my higher-order uncertainty, the higher my higher-order uncertainty towards P(A)=q.

Here's a brief example: For *Weather Prediction,* we are assigning probabilities to the proposition that: (A) it will snow tomorrow. Let's assume that P(A)=0.5, or my probability assignment for snow is 0.5 - I think it's equally likely to snow or not tomorrow. When assessing my confidence or certainty in P(A)=0.5, we frame this as another "order" of probabilities - worlds where I should be confident in that 0.5 assignment versus worlds where I shouldn't. Let's assume that we assign probability 0.95 to proposition: [I-should-be-confident-that-P(A)=q] P(confident P(A)=0.5) = 0.95; P(not confident P(A)=0.5) = 0.05. The q* = 0.95 means I have a high degree of confidence that my probability assignment of 0.5 is well supported. This contrasts with a complex case like the AI timeline where q* would be much lower: For example, if I assign 0.6 to proposition:[I-should-be-confident-that-P(A)=q], P(confident P(AI in 100 yrs)=0.5) = 0.6 P(not confident P(AI in 100 yrs)=0.5) = 0.4. The lower q* of 0.6 indicates greater uncertainty about whether 0.5 is the correct/justified probability assignment than *weather prediction.*

In cases of *A.I Timeline*, we have a higher-degree of higher-order uncertainty than *Weather Prediction.* Thus, q* is higher for *Weather Prediction* than *A.I Timeline* despite q being the same value. This is one explanation on why predictions for *A.I. Timeline* feels less certain (despite having the same probability assignment) than *Weather Prediction* and what motivates separating deep and ordinary uncertainty.

I hope at this point I've convinced you into thinking that deep uncertainty and higher-order uncertainty are closely linked. If you are convinced by 1.2, then you should be convinced by:

Thesis*: S has deep uncertainty towards P(A)=q, S's probability assignment for some proposition A, iff S has higher-order uncertainty q* towards P(A)=q.

1.3 The need of threshold q**

If deep uncertainty is motivated by similar motivations as higher-order uncertainty, why not draw simple equivalence without having a threshold?

I'll examine and reject the following thesis:

Thesis*: S has deep uncertainty towards P(A)=q, S's probability assignment for some proposition A, iff there is higher-order uncertainty q* towards P(A)=q.

I argue that accepting Thesis* makes you accept that ordinary cases just short of certainty leads to deep uncertainty. I believe labeling these as cases of deep uncertainty loses the core motivation listed in 1.1 about these propositions.

Mathematician: Suppose that I'm a talented mathematician working on a simple primary school algebra problem. I understand the concept quite extensively, and rarely make elementary mistakes in calculations. I'm now assigning probabilities over whether or not my solution for the problem is correct.

$P(A)=q$, q in this case should be a very high number, representing my high confidence level that I got the problem correctly, given my total evidence. Recall that our higher-order uncertainty is represented by $P(P(A)=q)=q^*$, where $P(P(A)=q)$ represents the probability assignment I should hold towards my probability assignment. In this case, q^* should also be a very high number, representing my high confidence level that q is the correct probability assignment for A.

However, q^* should still be <1. If $P(P(A)=q) = 1$, then you should be willing to bet anything on the truth of the proposition (e.g. you get a nickel if it is true, and are tortured for eternity if false). We often are not willing to bet in these cases. I also think specific to this case, the evidence that you still sometimes (although very rarely) makes mistakes is evidence why $q^*<1$, as your probability assignments are still fallible to mistakes. If $q^*<1$, then there is still some degree of higher-order uncertainty.

This would then satisfy Thesis*. Cases of low higher-order uncertainty due to recognizing that our beliefs are fallible are now categorized as cases of deep uncertainty. This seems wrong. We recognize that our beliefs are fallible for most ordinary propositions, the difference is more of scope or impact of recognized fallibility, rather than presence versus absence of falliability. For most of these propositions, despite having some level of higher-order uncertainty, we are quite certain that we are correct. Labeling these as cases of deep uncertainty loses the core motivation listed in 1.1 about these propositions.

Then, we should amend Thesis * with to:

Thesis: Higher-Order Uncertainty Thesis (HOU)**: S has deep uncertainty towards P(A)=q, S's probability assignment *q* for some proposition A, iff S has not responsibly integrated higher-order uncertainty q*, which is below some threshold some threshold q**, towards our probability assignment.

By adding the threshold condition (q*<q**), then given q* is a very high number in cases of low higher order uncertainty, it will easily exclude cases like *Mathematician* as cases of deep uncertainty.

1.4 Integrated higher-order uncertainty

In this section, I call for a responsible integration criteria by rejecting Thesis **. I wish to present cases where our higher-order uncertainty(above a certain threshold) is responsibly integrated into our probability assignment, so they do not fall under cases of deep uncertainty.

Drugged Mathematician: I'm a talented mathematician working on an intricate mathematical proof. I came up with a proof on Sunday, and I believe that there is a 70% probability that I got the proof (or hold a 0.7 credence). I later realized that I was drugged on Sunday. I know that when someone is drugged, there's a 60% probability that they are unable to reason properly. There is a 40% probability that the drug does nothing to them. If you believe that you are unable to reason properly, there is a very low probability (0.01%) of getting the proof correct. Let's assume that all of these probabilities are well-justified.

What should be my probability assignment that I got the proof correct? Dorst (2019) provides a formal framework of thinking through how our opinions are warranted by our total evidence. Specifically, Principle HiFi: The rational probability assignment equals the rational expectation of the probability assignment in q that's warranted by the first-order evidence (Dorst 2019, pg 12).

HiFi: $P(A) = Ep[(\tilde{P}(A)]$, $Ep[(\tilde{P}(A)] = \sum(P(\tilde{P}(A) = q) \cdot q)$

$\tilde{P}(A)=q$ represents the probability assignment q we have towards our first-order evidence ($\tilde{P}$). $Ep[(\tilde{P}(A)]$ is a weighted average of the various possible values of first-order evidence, with weights determined by how confident you should be in each. There is some degree of built in uncertainty for HiFi. Our rational expectation represents our number we expect to be near correct. Let's cross-apply HiFi to the *Drugged Mathematician* case.

There are two propositions. propositions 1: [I-should-be-confident-in-$\tilde{P}$(A)=q-given-drugged] and proposition 2: [I-should-not-be-confident-in-$\tilde{P}$(A)=q-given-drugged]. The weighted average should be = 0.6*0.0001+0.4*0.7=0.28006, or (my confidence level that I should be confident in proposition 1) * (probability assignment supported by my first-order evidence in proposition 1) + (my confidence level that I should be confident in proposition 2) * (probability assignment supported by my first-order evidence in proposition 2).

I will refer to this as a reflective process. Essentially, we are reflecting our probability assignment from first-order evidence with some rational expectation of our higher-order uncertainty. This process requires you to (1) take best estimates at the rational expectation of our higher-order uncertainty, in the case above, (my confidence level that I should be confident in proposition 1) & (my confidence level that I should be confident in proposition 2) (2) reflect these rational expectations by compounding it with what your first-order evidence justifies. This is required so that your P(A)=q reflects your higher-order uncertainty.

Back to *Drugged Mathematician*, assuming that HiFi is the correct principle, in these cases, I have higher-order uncertainty towards my probability assignment, but if I adjusted my probability assignment given my total evidence and that (1) my confidence level assignment in each proposition is well-justified (2) my probability assignment of what my first-order evidence in each proposition supports is well-justified, then I'm not deeply uncertain towards my probability assignment.

Adding "not responsibly integrated" to HOU thesis excludes cases like this. Classifying *Drugged Mathematician* as a case of deep uncertainty seems to be unintuitive – our uncertainty levels are well-justified and clear per caveats of the case. We should then adjust Thesis ** to include a condition that deep uncertainty are cases where our higher-order uncertainty isn't responsibly integrated. I discuss more about responsible integration in 2.1.

Rewriting Thesis ** will arrive at **Higher-Order Uncertainty Thesis (HOU)**: We have deep uncertainty towards P(A)=q, our probability assignment *q* for some proposition A iff we have not responsibly integrated higher-order uncertainty q*, which is above some threshold some threshold q**, towards our probability assignment .

HOU explains well why we feel differently about cases of deep uncertainty and ordinary uncertainty – as there are unintegrated higher-order uncertainty standing in addition to our probability assignment. This formalizes the intuition in 1.1 that these are propositions we cannot assign probabilities towards.

Part 2: Epistemic Privilege Question

2.1 Defense of Belief Thesis

Belief Thesis: If we have deep uncertainty towards P(A)=q, then we should not believe that q represents the rational probability of A given higher-order uncertainty.

In this section, I will prove the Belief Thesis by proving three claims:

(a) Responsible integration of our higher-order uncertainty into our probability assignment requires epistemic self-trust over some higher-order reflection process

(b) Epistemic self-trust over higher-order reflection process is absent in cases of deep uncertainty

(c) Believing that q represents rational probability of A requires our higher-order uncertainty being responsibly integrated into our probability assignment.

Higher-order uncertainty and self trust

Recall from 1.4 over the integration process of our higher-order uncertainty into probability assignment. We need this integration process because we hold different degrees of higher-order uncertainty towards different propositions. In order to adjust first-order evidence expectation $\tilde{P}(A)$ with respect to our higher-order uncertainty, we need to make explicit estimates over the degree of higher-order uncertainty we have and represent that through the reflection process where we calculate a weighted average of what my first-order evidence warrants in each proposition. If we believe that q is the rational expectation of our first-order evidence given higher-order uncertainty, then we need to have some degree of trust that the integration process outputs responsible, rational probability assignments accounting for our first-order evidence given higher-order uncertainty.

This process is not only required for a higher-order uncertainty approach, but any approach which tries to integrate our higher-order uncertainty/evidence into our probability assignments. If we view this through a higher-order evidence lens, Calibrationism says that one's probability assignment in a proposition p should cohere with one's expected reliability with regard to whether p (Ye, 2022). This is reliant upon some expected reliability principle: If I draw the conclusion p on the basis of any evidence e, my credence in p should equal my prior expected reliability with respect to p. In this case, we are having some degree of trust that applying the expected reliability principle outputs rational probability assignments accounting for our higher-order evidence.

What is self trust then? Zagzebski posits that to trust is to hold a 3-place relation. A trust X to Phi. A trusts X to phi if and only if: (i) A believes that X will phi; (ii) A has feelings of trust[3] with respect to this (iii) A treats X as if X will phi. (Zagzebski, 2012, pg. 37). In cases of self trust, we are trusting ourselves to phi. Cross-applying self-trust to the reflection process, our belief that the integrated process will generate rational probability assignments reflective of our total evidence is predicated on self-trust that the integration process will lead to reliable outputs.

I think there is a confirmational relationship we have between the evidence we gained on whether or not q is a rational probability we should hold for P(A) and our epistemic self-trust in the integration process. There could be evidence suggesting that q is the rational probability we should hold for P(A). If we keep getting these types of evidence (or getting strong evidence supporting this), then we should believe that our epistemic self-trust in the reflective process is well-grounded.

[3] Fricker (2016) amends to this that (ii) these are feelings of confidence/absence of anxiety and doubt with respect to this. This helps explain what trust is without using the word trust.

Here's an example. Suppose for *Weather Prediction,* climate scientists realized that there is some degree of fallibility in their models, they then through the reflective process lowered $\tilde{P}$(A) with respect to our higher-order uncertainty and generated P(A)=q. P(A)=q is the probability of snow on a particular day. Although each day of snow is an independent event, one could reasonably group "days it snows" as a reference class of snow prediction. Every time q is a very high number, and it snows that day, then that confirms q a bit more as a rational probability assignment for snow predictions – it doesn't confirm it entirely, as it may be because of luck. Cross-applying this for a large series of events, if our probability assignments roughly corresponds with the actual days it snowed, then it is good evidence that our reflective process generates probability assignments reflective of rational expectations of our first-order evidence given higher-order uncertainty.

Here's another example. Suppose for *Mathematician,* Mathematician realized that despite feeling highly confident about their mathematical solution for a primary school algebra problem, they realize that their solution is still fallible to mistakes. Let's assume that the Mathematician is also very bad at adjusting their first-order opinion with higher-order uncertainty, where they overestimate the degree of fallibility in their mathematical solutions. As a result of this, q is a lower number than the actual amount of times he got the math problem solved. After seeing the solution sets and the frequency he got the problem correct, the Mathematician realizes that this is good evidence that their reflective process is probably flawed and he should have less epistemic self-trust upon the process unless they become better at adjusting first-order opinion with higher-order uncertainty. They then are able to adjust their reflective process and regain epistemic trust.

In both of these cases, we are able to confirm our epistemic self-trust in integration processes via evidence in cases of ordinary uncertainty.

Deep uncertainty and the lack of trust

Here I argue that we do not trust our integration process in cases of deep uncertainty. I think one common trait for propositions with deep uncertainty is that these are one-off events in the future where it is difficult to find reference classes for these events. For example, it is difficult to group A.I. superintelligent development with "general technological growth prediction", simply because so much is different between the two: general technological growth does not involve approximation of human intelligence, it doesn't involve the product of technological growth engaging in self-enhancing behaviors e.t.c. It is then very difficult to assess if our integration process is generating a rational probability assignment if we do not have evidence confirming that our epistemic self-trust is well-grounded. This seems common for most of the classic examples of deep uncertainty. Then, we don't have the confirmation process of building trust like the cases listed above.

It also doesn't seem like the integration process can be something we have prima facie trust in. We are assessing how to integrate our higher-order uncertainties into our probabilities q for superintelligent A.I. systems development in 100 years. This seems really hard. (1) this is a single event in the far future. Even if I have a good estimation of how my cognitive faculties operate under massive complexity in general – it seems difficult to cross-apply this estimation to this specific case of A.I. Timeline; (2) Which direction should we adjust P(A)? It is not immediately obvious if we are overestimating or underestimating Superintelligence development – but our higher-order uncertainty suggest that we have good reasons to believe that we are either overestimating or underestimating, as it is very difficult for our (3) Even if we figure out direction, by how much should we adjust based upon our higher-order uncertainty?

But yet, it seems like integration is needed, or else we are just ignoring evidence. We know that the evidence serves as reasons why we should not believe our probability assignment, and yet we have no good way of integrating these evidence into our probability assignment.

No belief

This section takes on the assumption that the Non-Akrasia Constraint is generally required for rationality. The Non-Akrasia Constraint goes: it is never rational to believe/have a high degree of confidence in some proposition "P, but my evidence does not support P". This assumption has its intuitive appeal: to believe p but also believing that we should not rationally believe p given total evidence seems rather odd. This is similar to the intuition why we find Moorean Assertions such as "It is raining, but I believe it is not raining" as apparently absurd. Recent literature objecting to the Non-Akrasia Constraint generally does not challenge it as a general requirement, but identifies moderate cases where epistemic akrasia could be rational (Horowitz 2013).

The non-akratic constraint applies here: If we believe that q represents the rational probability of A given our higher-order uncertainty, then we are committed to believing that our total evidence, including the higher-order uncertainty, supports assigning probability q to A. However, in cases of deep uncertainty, we lack the epistemic self-trust in our reflective process that would allow us to believe q is rationally supported by our total evidence. Then, this demonstrates why we should not believe in our probability assignments.

2.2 **Defense of Assertion Thesis**

> ***Assertion Thesis:*** *If we have deep uncertainty towards P(A)=q, then we should not assert P(A)=q.*

<u>The role of assertion</u>

Roughly, to assert is to claim that something is the case. We assert probability assignments all the time: when the weather forecaster says that the probability of rain tomorrow is 0.7, they are claiming that 0.7 is the probability assignment we should hold for rain tomorrow. Yet, for some of our far-term future probabilities assignments (e.g., probability of human extinction by 2050 is 1/6th), it seems less epistemically permissible to assert these assignments. In this section, I test this intuition. I offer a normative defense that in cases when we have deep uncertainty towards P(A)=q, we should not assert P(A)=q.

Assertion Thesis is important. Assertion is commonly treated as necessary for testimony as testimony requires explicit statements (Fricker 1995). For S to testify P to some audience, S must be able to assert that P. It seems intuitive that we cannot just assert any proposition. For example, it seems epistemically irresponsible for me to assert my observations about organic chemistry where I have no expertise in. Then, whether or not we can epistemically assert propositions are governed by norms. Williamson frames that the assertability of propositions are governed by a single norm establishing epistemic standards (2000).

The literature on norms of assertion are focused on which norm(s) governs assertion. One plausible theory is the Knowledge Norm of Assertion (KNA): One must assert P only if one knows P. Defenders of KNA argue that KNA is well positioned at explaining the oddity of asserting on purely probabilistic grounds. For example, suppose there's a fair lottery with lots of ticket entries, A bought a lottery ticket. B asserts purely on probabilistic grounds that A will not win the lottery. Although the probability of the ticket winning is very low, it seems intuitively incorrect for B to assert to A that his tickets will not win (Williamson 2000, 247). KNA offers an intuitive explanation on why we find these assertions to be odd, as we are claiming something to be the case without knowing that something is the case. The main alternative view to KNA is the Justified Belief Norm of Assertion (JNA): One must assert *p* only if one is epistemically justified in believing *p.* Defenders of JNA argue that KNA is too strong in cases of Gettiered assertions, such as when one looks at the Cafeteria's watch and concludes that the current time is 12:35. Unbeknownst to you, Your belief is only accidentally true that the time has stuck on 12:35 for the past few days. In this case, your belief is short of knowledge, but it seems intuitively fine for you to respond "12:35" when someone asks you what time it is.

This paper does not take a stance under which norm best governs assertion. This paper will be focused on examining *Assertion Thesis* under JNA. The reason for this is that knowledge entails justification, when S knows that $p \rightarrow$ S is justified in believing that *p*. If it is epistemically impermissible to assert that *p* under JNA, then we should also not assert that *p* under KNA. This allows the *Assertion Thesis* to be non-constrained by which norm (KNA/JNA) governs assertion.

<u>What are we claiming to "assert" ?</u>

When one attempts to assert P(A)=q in a speech act, what are they claiming to assert? I believe there are three ways of viewing this:

(1) Asserting a credence state, or the degree of belief q I hold for a particular proposition. For example "I believe that my credence for A is q". Asserting this is purely descriptive, as it describes my mental state, and not asserting that the probability assignment I hold is rational. Asserting only (1) is somewhat uninteresting unless combined with some epistemic claim over

how we should evaluate this credence. Without it, the speaker doesn't assume epistemic responsibility when pressed: H:"Why did you tell me that you believe the probability of me winning that poker hand is 80%?" S: "Well, I just said I believed that, not that it is the correct probability you should hold."

(2) Asserting q as an epistemically responsible probability I should hold for a particular proposition, given available evidence. For example, "I believe that the responsible probability we should attribute to A is q, after considering the evidence I have". This may be rewritten to credence as well where "I believe that a rational Bayesian agent should hold degrees of belief q towards A, given available evidence". This assertion requires the speaker to take on some epistemic responsibility for attributing a responsible probability for A given available evidence.

Asserting (2) does not require you to assume responsibility that this is the only rational probability assignment. For (2), I may assert that the probability of climate change → human extinction is 20%, given the current evidence I have. Another scientist may assert that the probability of climate change → human extinction is 15%, given the evidence available to her. Our assertions are not contentious with each other. I have not assumed epistemic responsibility to defend my probability assignment against hers.

This paper is primarily concerned with asserting P(A)=q in (2). To assert (2) requires you to responsibly believe your probability assignment.

<u>Assertion and Belief</u>

My actual argument proving Assertion Thesis is very simple. Asserting (2) requires you to believe your probability assignment. If Belief Thesis is true, then proving Assertion Thesis under JNA with a full belief framework is just simply drawing a connection. *S* is justified in believing that *p* → S believes that *p*, as justified belief entails belief. Higher-Order Uncertainty

Thesis establishes that in cases of deep uncertainty, we should not believe our probability assignment, then we can't possibly be justified in our probability assignment, then we can't assert P(A)=q. This similarly applies to KNA, as knowledge also entails belief.

Proponents of JNA might object to this by pointing out cases of selfless assertions. For example, a Christian teacher who does not believe in evolution may still assert justified claims about how humans evolved from animals to their students. This would be a misread of my Belief Thesis, as it is not concerned with whether or not the speaker actually believes in the proposition, but that it is a normative thesis claiming that the speaker should not believe in their probability assignment.

The reason why we should not believe in our probability assignment in cases of deep uncertainty is laid out clearly in 1.2 and 1.4. We have higher-order uncertainty caused by evidence about evidence (such as we are missing evidence), cognitive limitations due to massive complexity, and peer disagreements. All three of these sources cause us to doubt our probability assignment. Cases of deep uncertainty differ from cases of ordinary uncertainty because we do not trust our integration process of higher-order uncertainties into our probability assignments. It seems like our probability assignments are not epistemically responsible ones as they are not justified properly.

Then, if we continue to assert these propositions. We will be faced by sanctions per constitutive norms of assertion. Evidence of this is strongest in two ways: (1) Hearers often press the speaker on "where did you get this probability", or "how could you know so far in ahead?". This is strong amongst cases of deep uncertainty – where some of these probabilities assignments may be labeled as too speculative (2) Speakers trying to hedge with their assertions. Recall the example at the start of the paper over 1/6th probability of human extinction, Ord also asserts a bunch of hedging propositions, that this probability is in no way the “final word”, “objective", or “precise” (pg 160-161). He is making clear that he is not assuming any epistemic responsibility that his probability assignment is the single, correct probability assignment one should hold. This is indirect evidence that these sanctioning norms exist.

Conclusion

This paper defended an account of deep uncertainty under the **Higher-Order Uncertainty Thesis (HOU)**: We have deep uncertainty towards P(A)=q, our probability assignment *q* for some proposition A iff we have not responsibly integrated higher-order uncertainty q*, which is above some threshold some threshold q**, towards our probability assignment. This hopefully will help clarify the distinctions between deep and ordinary uncertainty within the DMDU literature.

I then forwarded two thesis about the upshots of my argument: **Belief Thesis:** If we have deep uncertainty towards P(A)=q, then we should not believe that q represents the rational probability of A given higher-order uncertainty. ***Assertion Thesis:*** If we have deep uncertainty towards P(A)=q, then we should not assert P(A)=q. My work here is more exploratory than certain. I think there are many interesting epistemic questions left unanswered: What does it mean to be responsible a Bayesian in cases of deep uncertainty? How should we hedge our assertions properly? Should we still do expected value calculations in cases of deep uncertainty?

For what I have answered and criticized: much of it reminds me of a famous Chinese proverb 摸着石头过河 (crossing the river by feeling the stones). The probability assignments, despite having its problems, are the rocks we grasp on when we walk across the river. I’m sympathetic to the idea that some representation of our doxastic attitude is better than none – but also, we treat these predictions as more than mere numbers. Politicians are quoting Effective Altruism extinction probability assignments in the House of Representatives as a basis for decision making (Leigh 2023). Predictive models about the future are in the background of the work of Effective Altruist organizations such as Open Philanthropy, which allocate more than 300 million USD per year to Longtermists organizations over Neartermists philanthropic organizations on the grounds that the former carries higher expected value despite uncertainty.

This essay aimed to be more normative than prescriptive – I don't have concluding remarks on any specific use of these probability assignments. But I think it is useful to remember that the lived wisdom of the proverb comes from feeling the stones when we cross the river, and not grasping on them.

References

Carlsmith, J. (2022, June 16). *Is power-seeking ai an existential risk?*. arXiv.org. https://arxiv.org/abs/2206.13353

DeFinetti, B. (1974). *Theory of probability*. Wiley & Sons.

Dorst, K. (2019). Higher-order uncertainty. *Higher-Order Evidence: New Essays*, 35–61. https://doi.org/10.1093/oso/9780198829775.003.0002

Fricker, E. (2016). Doing (better) what comes naturally: Zagzebski on rationality and epistemic self-trust. *Episteme*, *13*(2), 151–166. https://doi.org/10.1017/epi.2015.37

Fricker, E. (1995). Telling and Trusting: Reductionism and Anti-reductionism in the Epistemology of Testimony. *Mind*, *104*(414), 393–411. https://doi.org/10.1093/mind/104.414.393

Hájek, A. (2002, November 16). *Interpretations of probability*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/probability-interpret/

Hedden, B., & Dorst, K. (2022). (almost) all evidence is higher-order evidence. *Analysis*, *82*(3), 417–425. https://doi.org/10.1093/analys/anab081

Hillier, F. S., & Lieberman, G. J. (2001). *Introduction to operations research* (7th ed.). McGraw-Hill.

Horowitz, S. (2013). Epistemic akrasia. *Noûs*, *48*(4), 718–744. https://doi.org/10.1111/nous.12026

Jackson, E. G. (2020). The relationship between belief and credence. *Philosophy Compass*, *15*(6). https://doi.org/10.1111/phc3.12668

Leach, M., Scoones, I., & Stirling, A. (2010). Governing epidemics in an age of complexity: Narratives, politics and Pathways to Sustainability. *Global Environmental Change*, *20*(3), 369–377. https://doi.org/10.1016/j.gloenvcha.2009.11.008

Leigh, A. (2022, September 22). *What's The worst that could happen? existential risk and extreme politics - speech*. Andrew Leigh MP. https://www.andrewleigh.com/what_s_the_worst_that_could_happen_existential_risk_and_extreme_politics_speech

LeRoy, S. F., & Singell, L. D. (1987). Knight on Risk and Uncertainty. *Journal of Political Economy*, *95*(2), 394–406. http://www.jstor.org/stable/1832078

Lempert, R. J., Popper, S. W., & Bankes, S. C. (2003). *Shaping the next one hundred years new methods for quantitative, long-term policy analysis and bibliography*. RAND.

MacAskill, W., Bykvist, K., & Ord, T. (2020). *Moral uncertainty*. Oxford University Press.

MacFarlane, J. (2019). *On probabilistic knowledge* . https://johnmacfarlane.net/moss.pdf. https://www.johnmacfarlane.net/moss.pdf

Marchau, V. A. W. J., Walker, W. E., Bloeman, P. J. T. M., & Popper, S. W. (2019). *Decision Making Under Deep Uncertainty*. Springer Nature.

Moss, S. (2019). *Probabilistic knowledge*. Oxford University Press.

Müller, V. C., & Bostrom, N. (2016). Future progress in Artificial Intelligence: A survey of expert opinion. *Fundamental Issues of Artificial Intelligence*, 555–572. https://doi.org/10.1007/978-3-319-26485-1_33

Ord, T. (2020). *The precipice: Existential risk and the future of humanity*. Hachette Books.

Patt, A. G., & Schrag, D. P. (2003). Using Specific Language to Describe Risk and Probability. *Climatic Change*, *61*(1/2), 17–30. https://doi.org/10.1023/a:1026314523443

Patt, A., & Dessai, S. (2005). Communicating uncertainty: Lessons learned and suggestions for climate change assessment. *Comptes Rendus Geoscience*, *337*(4), 425–441. https://doi.org/10.1016/j.crte.2004.10.004

Rosenkrantz, R. D. (1981). *Foundations and applications of inductive probability*. Ridgeview.

Schwitzgebel, E. (n.d.). Belief. In *The Stanford encyclopedia of philosophy*.

Stanton, M. C., & Roelich, K. (2021). Decision making under Deep Uncertainties: A review of the applicability of methods in practice. *Technological Forecasting and Social Change*, *171*, 120939. https://doi.org/10.1016/j.techfore.2021.120939

Sunstein, C. R. (2023). Knightian uncertainty. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4662711

Taleb, N. N. (2017). *The Black Swan: The impact of the highly improbable*. Taylor and Francis.

Todd, B. (2023, December 14). *Why despite global progress, humanity is probably facing its most dangerous time ever*. 80,000 Hours. https://80000hours.org/articles/existential-risks/

Walker, W. E., Lempert, R. J., & Kwakkel, J. H. (2013). Deep uncertainty. *Encyclopedia of Operations Research and Management Science*, 395–402. https://doi.org/10.1007/978-1-4419-1153-7_1140

Williamson, J. (1999). Countable additivity and subjective probability. *The British Journal for the Philosophy of Science*, *50*(3), 401–416. https://doi.org/10.1093/bjps/50.3.401

Williamson. (2000). *Knowledge and its limits*. Oxford University Press.

Ye, R. (2022). *Higher-order evidence and calibrationism*. Cambridge University Press.

Zagzebski, L. T. (2012). *Epistemic authority*. Oxford University Press.