The Comparison Problem for Approximating Epistemic Ideals

Marc-Kevin Daoust, École de technologie supérieure. marc-kevin.daoust@etsmtl.ca

Abstract: Some epistemologists think that the Bayesian ideals matter because we can *approximate* them. That is, our attitudes can be more or less close to the ones of our ideal Bayesian counterpart. In this paper, I raise a worry for this justification of epistemic ideals. The worry is this: In order to correctly compare agents to their ideal counterparts, we need to imagine idealized agents who have the same relevant information, knowledge, or evidence. However, there are cases in which one's ideal counterpart cannot have one's information, knowledge, or evidence. In these situations, agents cannot compare themselves to their ideal counterpart.

1 INTRODUCTION

What is the place of ideals in epistemology? Are they necessary? Are they relevant?[1] Just think of the demands of Bayesian epistemology. For Bayesians, agents should have credences that satisfy the axioms of probability,[2] they should form and revise their credences in accordance with Conditionalization,[3] they should be omniscient for logical truths,[4] they should satisfy some 'inter–level' coherence requirements, such as Rational Reflection or Immodesty,[5] and they should also satisfy other substantive requirements such as the Principal Principle.[6] However, we are not in a position to meet some, or all, of the above ideals of Bayesian epistemology. So, why do these epistemic ideals matter to ordinary agents like us?[7] Call this the Relevance Challenge for Bayesian epistemology.

---

1   See Christensen (2007, 2013), DiPaolo (2019), Staffel (2019), and Thorstad (2021) for discussion.

2   See Joyce (1998), Leitgeb & Pettigrew (2010a; 2010b), and Pettigrew (2016).

3   See Greaves & Wallace (2006), Meacham (2015), and Schoenfield (2017).

4   See Easwaran (2011) and Dogramaci (2018).

5   See Lewis (1971), Joyce (2009), Horowitz (2014; 2019), and Elga (2010; 2013).

6   See Pettigrew (2012; 2016).

7   See de Bona & Staffel (2018), DiPaolo (2019), Halpern & Pass (2015), Icard (2018), Lorkowski & Kreinovich (2018), Morton (2012, chap. 1), Paul & Quiggin (2018), Staffel (2017; 2019; forthcoming), Skipper & Bjerring (2022), Thorstad (2021), and Wiens (2020).

Here is a possible response to this worry: Ideals of Bayesian epistemology matter because we can *approximate* them. We can be *more or less close* to the epistemic ideal. Given this, what we should do is approximate the ideal the best we can. For instance, suppose it would be ideal for me to have a credence of 1 in P. If I have a credence of 0 in P, I am maximally far from the ideal (at least, with respect to this specific credence). If I have a credence of 0.5 in P, I am getting closer to the ideal credence in P, and so forth. Maybe we cannot fully comply with the demands of Bayesian epistemology, but we can come close to it. Call this the Bayesian Approximation Thesis.[8]

Claims concerning the approximation of ideals raise various questions, such as: Why think that approximating ideals is a good thing? What do we mean, exactly, by 'approximating' an ideal? Which measures principles are appropriate? Those who accept the Approximation Thesis should answer these questions. In the past few years, various contributions to the literature have helped us make significant progress in answering these questions. De Bona and Staffel (2018) give us good reasons to approximate some requirements of formal coherence. Staffel (2017; 2019; forthcoming) explores what it means to approximate an ideal and explores various distance principles for correctly measuring the distance between an agent's credences and his or her ideal counterpart's credences. DiPaolo (2019) has offered arguments in favour of an epistemic theory of the Second Best, which clarifies the extent to which we should approximate the epistemic ideals. These contributions help us substantiate the Approximation Thesis and its limits.

But there is still work to do. In particular, it is not always clear how we should apply the approximation method in a concrete situation. This article focuses on this issue. I argue that there are situations in which we cannot compare ourselves to our ideal counterpart. More precisely, there are common cases in which we cannot measure the distance between our attitudes and those of our ideal counterpart. This is a worry for the Approximation Thesis. Of course, this does not mean that the justification of ideals based on approximation is entirely irrelevant. But this

---

8   See Christensen (2007, 7) for an early articulation of this thought. See De Bona & Staffel (2018) and Staffel (2019) for details on how to measure degrees of closeness to an epistemic ideal. Note that this is one way among others that we can justify the relevance of ideals. Bayesian epistemologists could turn to models of ideal advisors (see Smith 1994, Sampson 2022, and Karlan 2021 for discussion). We could say that epistemic ideals matter because they provide robust epistemic norms (Carr 2021). We could say that epistemic ideals operate like the values of a function that we want to maximize, even though we might not be able to fully maximize all the values of the function (see Christensen, 2007, p. 24; Christensen 2010). The Approximation Thesis *is one way among others* to argue for the relevance of idealized epistemic norms.

suggests that the justification is not universal, or applicable to all the situations we may encounter.

In section 2, I describe the worry more clearly. In section 3, I explore some possible responses to the worry. In section 4, I analyse the implications of the argument for the Approximation Thesis.

2 THE PROBLEM

According to the Bayesian Approximation Thesis, agents can have attitudes that are more or less to their ideal counterpart's attitudes.[9] An 'ideal counterpart' is the idealized version of an agent, or the version of an agent who complies with certain Bayesian ideals of rationality. This expression comes from Staffel (2017). So, when we compare an agent's attitudes to the attitudes of her ideal counterpart, we compare the agent's attitudes to the ones entertained by the idealized version of the agent.

The degree to which agents approximate the ideal is a function of the distance between their actual attitudes and the ideal ones. So, if an agent has a credence of 0.5 in H and the ideal credence of H is 1, the degree of approximation of the ideal is a function of the distance between 0.5 and 1. Take the Euclidean distance. If I have a credence of 0.5 in H and my ideal counterpart has a credence of 1 in H, then the degree to which I am ideally rational is given by: $(0.5–1)^2$. In other words, relative to the Euclidean distance, my degree of ideal rationality is the square of the difference between my attitudes and the ones of my ideal counterpart.

In simple cases, there is *one* ideal credence in a proposition. But what about more complicated cases of *conflicting* ideals or requirements?[10] Suppose that requirement X calls for the thinker to have a credence of 1 in H, and requirement Y calls for the thinker to have a credence of 0.9 in H. Here, we cannot imagine one idealized counterpart that has a unique ideal credence (that is, a credence in H that satisfies *all* the ideal requirements). Instead, we need to imagine one idealized counterpart who satisfies requirement X and one idealized counterpart who

---

9   See Staffel (2019, chap. 3).

10   As discussed in, e.g., Christensen (2007, 2010, 2013). Very roughly, Christensen argues that requirements of formal coherence, integration between first– and higher–order beliefs, and evidence–responsiveness can conflict with each other. See also Hughes (2019, 2021) on epistemic dilemmas.

satisfies requirement Y. Depending on the *weight* or *priority* that each of these requirements has, we can then figure out the credence that the agent should approximate (see Staffel (2019, chap. 6) on the piecemeal strategy for details). What is important to note here is that, even in cases of conflicting ideals, approximation is still a function of the distance between an agent's attitudes and the ones entertained by ideal counterparts. That is, even in cases where ideals conflict with each other, we still need to figure out the credences of some idealized counterparts (there are simply more steps to the method).

At first sight, such comparisons between actual and ideal agents appear to be unproblematic. But how do we make these comparisons in an actual case? Are we always in a position to compare agents with their ideal counterparts?

In this section, I present the problem more clearly. My argument goes as follows:

(P1) In order to correctly compare the attitudes of A with the attitudes of A's ideal counterpart, we need to imagine an idealized agent who has the same relevant information, knowledge, or evidence than A.

(P2) However, there are (common) cases in which A's ideal counterpart cannot have A's relevant information, knowledge, or evidence.

(P2A) Non–ideal agents like A can know, learn or be correctly certain that they have violated an idealized requirement of rationality, but A's ideal counterpart cannot be in the exact same epistemic state. We could say the same thing about other facts concerning ideal rationality.

(P2B) Learning facts like 'I have violated idealized requirements of rationality' is *relevant* for the rational evaluation of how we compensate for our imperfections. The way we deal with our imperfections can be more or less rational.

(C) So, there are (common) cases in which we cannot correctly compare the attitudes of A with the attitudes of A's ideal counterpart.

Begin with P1. Why think that, in order to compare agents with their ideal counterparts, we need to imagine idealized agents who have the same information, knowledge, or evidence? It is because information, knowledge or evidence often matters for evaluating the ideal rationality of our attitudes. And so, if we want to compare apples to apples, we need to compare our attitudes with the attitudes of idealized agents in the same epistemic position.

Suppose we flip a fair coin. You know that it is fair, but you do not know which side it landed on. Relative to your knowledge, only a credence of 0.5 in the proposition 'the coin landed on heads' (H) satisfies the Principal Principle. So, given what you know, it is ideal for you to have a credence of 0.5 in H. Or, your ideal counterpart would have a credence of 0.5 in H. We come to this conclusion because, if an idealized Bayesian agent had the *same knowledge*, then that agent would have a credence of 0.5 in H.[11]

However, if you had more information, it could be ideal for you to have a different credence in H. Suppose that you look at the coin. You come to know that it landed on tails. Then, it is not ideal for you to have a credence of 0.5 in the proposition 'the coin landed on heads.' Given the additional information you now have, your ideal counterpart has a lower credence in H. The lesson here is this: If we want to compare ourselves correctly to our ideal counterpart, we must imagine an ideal agent who is in the same 'epistemic position' as us (i.e., who has the same relevant information, evidence, or knowledge). Otherwise, we will not correctly measure the distance between our actual attitudes and the ones that would be ideal for us to have.

For similar reasons, the types of attitude we have can make a difference. There is a difference between (i) having some evidence supporting P, (ii) knowing that P, and (iii) being certain that P. Suppose I merely have some evidence for believing P (it is uncertain, on my evidence, that P). Then, we should not compare my attitudes to the ones of an ideally rational agent who is certain that P. My evidence does not make it ideally rational, for me, to be certain that P. Again, if we want to compare apples to apples, we need to measure the distance between our attitudes and the ones of an ideally rational agent in the same epistemic position.

Here is another way to see this. Suppose that I believe that P's objective probability is 1 and I have a credence of 1 in P. However, my second–order belief is outlandish. It is not supported by the evidence and does not amount to knowledge. Then, suppose I say this: 'I am ideally rational. You see, my ideal counterpart who *knows* that P's objective probability is 1 has a

---

11 Staffel's examples also support this conclusion. For instance, she says: 'We will assume that, in order to be rational, the thinker's credences should obey the following principles: the credences have to be probabilistically coherent, and they have to obey the Principal Principle. We will assume that the thinker's knowledge about the chances mandates that that both A and ∼A should get credence 0.5. The optimal credence assignment to have in this case is thus: c(A)=0.5, c(∼A)=0.5' (Staffel, 2019, p. 107). This example makes sense under the assumption that the ideal attitudes are a function of the agent's knowledge, evidence or information. An ideally rational agent with more relevant information could have a different credence in A.

credence of 1 in P. So, me and my ideal counterpart have the same attitudes.' The problem here is that we are not comparing apples to apples. My belief that P's objective probability is 1 does not amount to knowledge. And that makes a difference for evaluating my (degree of) ideal rationality. Once again, the point is that, in order to correctly compare ourselves to our ideally rational counterpart, we need to imagine an idealized agent with the same relevant epistemic states.

Here is an objection against P1. It seems that we can make 'rationality comparisons' between two agents (say, Trump and Obama) even though they do not have exactly the same evidence, knowledge, or information. For example, the proposition 'Trump is less rational than Obama' seems right, even if Trump and Obama do not have exactly the same evidence. This suggests that P1 is too strong. Response: There is a difference between comparing the epistemic character of two agents and comparing the specific attitudes of two agents. We can say that Obama is more rational than Trump because he has better dispositions—that is, dispositions to form attitudes that are more rational. But if we want to assess the degree of rationality of Trump's *specific attitudes*, we need to imagine an ideal agent with the same evidence, information or knowledge. If Trump knows that the coin landed on tails, while Obama does not know this, it would be a mistake to compare their credences in H. They can have different credences in H, and yet, their credences can be equally rational (if their credences are not sensitive to the same information).

Next is P2. Why think that there are common cases in which my ideal counterpart cannot have my relevant information, knowledge, or evidence? It is because by assumption, my ideal counterpart fully complies with the demands of ideal rationality. However, agents like us do not fully comply with the demands of ideal rationality. So, propositions like the following can be part of my knowledge, but cannot be part of my idealized counterpart's knowledge:

1. 'I have violated an idealized requirement of rationality.'
2. 'My cognitive limitations are an obstacle to complying with the idealized requirements of rationality.'
3. 'I am not an ideally rational agent.'

Consider a simple example. Suppose you were one of the participants in Tversky and Kahneman's (1983) study on the Conjunction Rule. Or suppose, more realistically, that your friends make you undergo the experiments conducted by Tversky and Kahneman. Your friends

(or the experimenters) start by describing Linda, a fictional character. They say that Linda is young, single, smart, has majored in philosophy, is concerned with issues of social justice and discrimination, and has taken part in anti–nuclear demonstrations.[12] You are then asked if the proposition 'Linda is a bank teller' is less likely that the proposition 'Linda is a bank teller and is active in the feminist movement.' You could easily imagine that Linda is a feminist, and so you spontaneously agree. Like 80% of the participants in Tversky and Kahneman's study, you violated the Conjunction Rule. That is, you claimed that, relative to your evidence, the proposition 'Linda is a bank teller' (P) is less likely that the proposition 'Linda is a bank teller and is active in the feminist movement' (P&Q).

Then, you read Tversky and Kahneman's paper. You find their paper illuminating, as it perfectly describes how you reason. Their analysis describes your mental processes so well that you could be certain, and correctly so, that you are not ideally rational. Accordingly, you come to know that you have violated an idealized requirement of rationality, that your cognitive limitations are an obstacle to complying with the idealized requirements of rationality, and so forth.[13]

Given what you now know, you decide to *compensate for your imperfections*.[14] Since you know you are bad at logic and probability calculus, you become less confident in your own abilities to solve certain formal problems. Or you revise your epistemic standards. For instance, when you face problems that involve logic or probability calculus, you rely more on others, or you defer to the expertise of others, for solving them. These are good decisions (at least, for imperfect agents like you). If you come to know that you are less than ideally rational, compensating for your imperfections is a perfectly appropriate response. As DiPaolo says, 'normativity doesn't end with perfection. There are better and worse ways of being imperfect. Secondary norms [of compensation] express how we ought to compensate for imperfections.' (DiPaolo, 2019, p. 2048)

---

12 See Tversky and Kahneman (1983, p. 297).

13 If you do not like this example, we could easily think of other cases. Some philosophers have argued that ideally rational agents do not forget their evidence (see Carr 2015 for discussion). But agents like you and I know (and can even be certain) that they forget their evidence. Some philosophers have argued that ideally rational agents are logically omniscient (see Dogramaci 2018 for discussion). But agents like us know that we sometimes make mistakes concerning logical truths.

14 See DiPaolo (2019).

Suppose you want to know the degree to which you are ideally rational. That is, you think you made decisions that compensate well for your imperfections, and you want to know how far you are from your ideal counterpart. In accordance with the Approximation Thesis, you need to measure the distance between your attitudes and the ones of your ideal counterpart who has the same relevant evidence, information or knowledge as you. But here is the problem: Idealized counterparts cannot come to know that they are less than ideally rational, or that their cognitive limitations are an obstacle to complying with the idealized requirements of rationality. By assumption, idealized counterparts fully comply with the demands of idealized rationality. Since it is false that they violate the idealized requirements of rationality, they cannot know this. So, you have relevant knowledge that your ideal counterpart cannot have.

To make a long story short, if you know that you are less than ideally rational, and you take the appropriate means to compensate for your imperfections, we cannot compare you to your idealized counterpart. You have no idealized counterpart. This is a problem for the Approximation Thesis. According to this view, we should try to minimise the distance between our attitudes and the ones that are ideal for us to have. But if we have no ideal counterpart, there is no effective way to do this.

To be clear, I am not saying that ideally rational agents are certain, or infallible, concerning their own rationality. There are cases discussed in print of misleading higher–order evidence, or rational modesty, where ideally rational agents are uncertain of their own rationality.[15] However, the problem is that ideal agents cannot know, learn, or be correctly certain that they are less than ideally rational. There is an important difference between (i) having some evidence for believing that you are non–ideal, and (ii) being in a position to know, or being certain, that you are non–ideal. Ideal agents cannot know that they are non–ideal, but we can. This difference in knowledge is a problem for comparing non–ideal agents with their putative ideal counterpart.

3 SOME POSSIBLE SOLUTIONS

In this section, I explore different ways to get around the worry.

---

15 See, e.g., DiPaolo (2019), Lasonen–Aarnio (2015), and Titelbaum (2015) for discussion.

3.1 We Should Still Approximate Ideals... Just Not the Bayesian Ideals

Those who want to solve this problem can reconsider the ideal we approximate. Epistemologists are familiar with other epistemic ideals, such as truth or accuracy. Perhaps we should not approximate the ideals of Bayesian rationality. Rather, perhaps we should approximate the attitudes of an omniscient agent (i.e., an agent with all the true beliefs and no false ones, or with all the maximally accurate credences).

However, if we take this option, we no longer have a solution to the Relevance Challenge for Bayesian epistemology. We still need to figure out why some epistemic ideals, such as the ones described in Bayesian epistemology, matter. So, given what *motivates* the Approximation Thesis in the first place, this solution is not fully satisfactory.

3.2 We Should Restrict the Approximation Thesis to A Specific Class of Requirements

Those who want to solve this problem can limit the Approximation Thesis to requirements that are not sensitive to the evidence, information, or knowledge agents have. For instance, the requirements of Probabilism say that an agent's credences should conform to the probability axioms. The satisfaction of these requirements is not dependent on the specific evidence, information or knowledge agents have. And so, if we limit the Approximation Thesis to these requirements, we might get around the problem.

However, this solution does not seem general enough. Most of the Bayesian requirements have to do with the evidence, knowledge, or information agents have. For instance, the Principal Principle, Conditionalization, and inter–level coherence requirements, have to do with the evidence agents have. Accordingly, even if we accept this solution, we still need to figure out how to solve the Relevance Challenge for the vast majority of requirements described in Bayesian epistemology.

3.3 We Should Make A Distinction Between Approximating Ideals and Approximating the Attitudes of Ideal Counterparts

A third response to the problem is that we should not understand 'approximating ideals' as 'getting closer to the attitudes entertained by our ideally rational counterparts.' Perhaps it is a mistake to think that we should approximate the attitudes of our ideal counterparts. However,

this does not mean that we should give up approximation. We just need a better account of what we are trying to approximate.

The problem with this response is this: it is unclear that there is a good alternative to approximating the attitudes of our idealized counterparts, and that this alternative justify Bayesian ideals. Let me explain. Here are two alternative accounts of the approximation of epistemic ideals:

1. **Ideal Advisor Approximation**. Approximating an ideal means following as many of the epistemic recommendations of an ideal advisor who knows all the facts of our situation.[16]

2. **Requirements Approximation**. Approximating an ideal means satisfying as many idealized requirements of epistemic rationality as possible.

Ideal Advisor Approximation would avoid the worry I describe in the previous section. John's ideal counterpart cannot know the proposition 'I am irrational', but John's ideal advisor can know that John is irrational. Ideal advisors can recommend certain policies to us even if we are irrational. However, Ideal Advisor Approximation might not justify Bayesianism. Ideal advisors will not often make recommendations that perfectly align with the demands of Bayesian epistemology. For instance, take the rules of reasoning. If my ideal advisor knows all the facts of my situation (including facts about my imperfections and my limited cognitive capacities), they will probably recommend that I reason with fast and frugal heuristics.[17] They will not recommend that I reason with complicated rules, like Conditionalization.

Recall the challenge we started with: We do not, and possibly cannot, live up to the standards of Bayesian epistemology. So, why should these ideals matter to agents like us? This is the challenge we are trying to answer. Ideal Advisor Approximation does not provide a direct response to the Relevance Challenge, since ideal advisors might not recommend the demanding rules of Bayesianism.

Requirements Approximation raises a different problem. When we violate at least one of the idealized requirements of rationality, we are sometimes better off violating the remaining ones.[18] For instance, suppose that I fail to respond correctly to my evidence that P's objective

16 Ideal Advisor Approximation is similar to the ideal advisor theory in ethics (see, e.g, Smith 1994 and Sampson 2022).

17 See Karlan (2021).

18 See Staffel (2019, 112). See also Daoust (2022).

10

probability is 0.5. I face certain constraints, and as a result, I have a credence of 0.6 in P. Now, suppose I have rational control over my credence in ~P. I could choose to satisfy the requirement of probabilistic consistency and form a credence of 0.4 in ~P.[19] However, relative to the constraints I face, this choice would fail to maximise expected accuracy. So, Requirements Approximation has the problematic consequence that getting closer to an ideal will sometimes make us worse off. As Staffel says, it is 'not the case that the second best credence assignment is one that still meets the principles of rationality that are not barred from being complied with by the constraint' (Staffel, 2019, p. 115).

Thus, one way to save the Approximation Thesis would be to revise what is approximated. But there is no obvious candidate that could replace idealized counterparts.

3.4 We should Reconsider the Place of Higher–Order Knowledge for Approximation
A fourth response to the problem is that higher–order knowledge concerning our own irrationality is irrelevant for the Approximation Thesis. Second–order facts, like 'I am irrational' or 'I made a reasoning mistake,' are not part of the relevant knowledge that ideally rational agents take into account in their deliberation. So, if we exclude the higher–order knowledge from the approximation model, we might get around the worry.

I have a couple of reactions to this suggestion. First, most of the discussion in section 2 focuses on what non–ideal agents like us *know*. But the point I made is not limited to knowledge. We could make the same point in terms of the information or facts agents have, or in terms of the (accurate) certainties of non–ideal agents like us. And these states are relevant in various interpretations of Bayesianism. In some versions of Bayesianism, the ideally rational attitude is determined, in part, by the *information* agents have, or the facts they *learn*. In other versions of Bayesianism, the ideally rational attitude is determined, in part, by the *maximally accurate credences* agents have. What matters for my argument is that, while it is true that agents like us can be irrational, it is false for ideal agents. This causes trouble for comparing ourselves to an ideal counterpart with the same information, knowledge, set of learned facts, maximally accurate credences, and so forth.

---

19 See, e.g., Joyce (1998; 2009) on probabilistic consistency.

Second, it is hard to see how this suggestion really solves the problem. Even if higher–order knowledge were irrelevant for the rationality of ideal agents, it is relevant for the rationality of *our* attitudes. Recall that, in section 2, the agent reacts in certain ways to her knowledge that she violated an idealized requirement of rationality. For instance, she is less confident in her ability to solve some problems involving probabilities. Arguably, the agent's reaction to this sort of knowledge can be evaluated rationally: You can react more or less rationally to knowing that you are less than ideally rational. But if we say that higher–order knowledge of propositions like 'I am irrational' is irrelevant to the ideal rationality of our attitudes, we cannot make such evaluations. At least, we cannot make such evaluations in terms of an *approximation* of the ideal counterpart. This suggests, once again, that the Approximation Thesis has shortcomings: There are situations in which we cannot evaluate the rationality of agents like us by approximation of an ideal figure.

Finally, this solution raises other problems. We can think of simple cases in which higher–order knowledge is relevant to the ideal rationality of our attitudes. Some putative Bayesian requirements are sensitive to the higher–order knowledge agents have. The clearest examples are Reflection principles and the requirement of Immodesty. Consider Immodesty, which roughly says that ideally rational agents estimate that their attitudes and standards maximise expected accuracy.[20] This requirement can be sensitive to the higher–order knowledge agents have. Suppose that, at t_1, an ideal agent discovers some new standards B, and comes to know that standards B are better than standards A in terms of expected accuracy maximisation. Suppose also that, prior to t_1, the ideal agent had standards A. Then, relative to her newly acquired knowledge, we expect the ideal agent to change her standards at t_1 (after all, if she is immodest, she estimates that her standards are optimal. Accordingly, she should not keep standards A after t_1). In a case like this, higher–order knowledge is relevant for the ideal rationality of certain attitudes. In light of this, it is unclear that we can get rid of higher–order knowledge in the approximation model, or deem it irrelevant in the evaluation of ideal rationality.

---

20 See Lewis (1971).

4 CONCLUSION: SHOULD WE REJECT THE BAYESIAN APPROXIMATION THESIS?

I have raised a worry for the Bayesian Approximation Thesis. In order to compare apples to apples, we need to compare ourselves to an idealized counterpart with the same relevant epistemic states. But this is not possible in cases where, e.g., we know that we are irrational and react to this knowledge. We do not have an idealized counterpart who knows propositions like 'I am irrational.' And so, it is hard to evaluate the rationality of our reaction to this higher–order knowledge *by approximation*.

Where does that leave us with respect to the Approximation Thesis? Arguably, this means that the possibility of approximating an ideal is not an all–embracing justification of the Bayesian ideals. We cannot always approximate the ideal figures. However, this is not terrible news for those who accept the Bayesian Approximation Thesis. There are still many cases in which we can compare ourselves to an ideal counterpart and measure the degree of ideal rationality of our attitudes. It is just that there is room for other justifications of Bayesian ideals, especially in cases where it is not possible to compare ourselves to an ideal counterpart.

Thus, the Approximation Thesis might still be relevant in many contexts. But there is still work to do for identifying the cases in which it is relevant, and the cases in which it does not apply. And since there are cases in which it does not apply, there is room for identifying other justifications of the Bayesian ideals.[21]

References

Carr, J. R. (2015). Don't Stop Believing. *Canadian Journal of Philosophy*, *45*, 744–766.

Carr, J. R. (2021). Why Ideal Epistemology?. *Mind*, online first. doi:10.1093/mind/fzab023

Christensen, D. (2007). Does Murphy's Law Apply in Epistemology? Self–Doubt and Rational Ideals, 3–31. In *Oxford studies in epistemology 2*, edited by T. Szabó Gendler and J. Hawthorne. Oxford: Oxford University Press.

---

21 See footnote 8 on this point.

Christensen, D. (2010). Higher-Order Evidence. *Philosophy and Phenomenological Research*, *81*, 185–215.

Christensen, D. (2013). Epistemic Modesty Defended, 77–97. In *The Epistemology of Disagreement: New Essays*, edited by D. Christensen and J. Lackey. Oxford: Oxford University Press.

Daoust, M-K. (2022). Imperfection, Accuracy, and Structural Rationality. *Erkenntnis*, online first. doi:10.1007/s10670-022-00571-5

De Bona, G., & Staffel, J. (2018). Why Be (Approximately) Coherent? *Analysis*, *78*, 405–415.

DiPaolo, J. (2019). Second Best Epistemology: Fallibility and Normativity. *Philosophical Studies*, *176*, 2043–2066.

Dogramaci, S. (2018). Solving the Problem of Logical Omniscience. *Philosophical Issues*, *28*, 107–128.

Easwaran, K. (2011). Bayesianism II: Applications and Criticisms. *Philosophy Compass*, *6*, 321–332.

Elga, A. (2010). How to Disagree about How to Disagree, 175–186. In *Disagreement*, edited by T. Warfield and R. Feldman. Oxford: Oxford University Press.

Elga, A. (2013). The Puzzle of the Unmarked Clock and the New Rational Reflection Principle. *Philosophical Studies*, *164*, 127–39.

Greaves, H., & Wallace, D. (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, *115*, 607–632.

Halpern, J. Y., & Pass, R. (2015). Algorithmic Rationality: Game Theory with Costly Computation. *Journal of Economic Theory*, *156*, 246–268.

Horowitz, S. (2014). Immoderately Rational. *Philosophical Studies*, *167*, 41–56.

Horowitz, S. (2019). The Truth Problem for Permissivism. *Journal of Philosophy*, *116*, 237–262.

Hughes, N. (2019). Dilemmic Epistemology. *Synthese*, *196*, 4059–4090.

Hughes, N. (2021). Who's Afraid of Epistemic Dilemmas?, 253–268. In *Epistemic Dilemmas*, edited by K. McCain, S. Stapleford, and M. Steup. New York: Routledge.

Icard, T. F. (2018). Bayes, Bounds, and Rational Analysis. *Philosophy of Science*, *85*, 79–101.

Joyce, J. M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, *65*, 575–603.

Joyce, J. M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief, 263–297. In *Degrees of Belief*, edited by F. Huber and C. Schmidt–Petri. Dordrecht: Springer.

Karlan, B. (2021). Reasoning With Heuristics. *Ratio*, *34*, 100–108.

Lasonen–Aarnio, M. (2015). New Rational Reflection and Internalism about Rationality, 145–171. In *Oxford Studies in Epistemology 5*, edited by T. Szabó Gendler and J. Hawthorne. Oxford: Oxford University Press.

Leitgeb, H, & Pettigrew, R. (2010a). An Objective Justification of Bayesianism I: Measuring Inaccuracy*. *Philosophy of Science*, *77*, 201–235.

Leitgeb, H, & Pettigrew, R. (2010b). An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy*. *Philosophy of Science,* 77, 236–272.

Lewis, D. (1971). Immodest Inductive Methods. *Philosophy of Science, 38*, 54–63.

Lorkowski, J., & Kreinovich, V. (2018). *Bounded Rationality in Decision Making under Uncertainty: Towards Optimal Granularity*. Cham: Springer.

Meacham, C. J. G. (2015). Understanding Conditionalization. *Canadian Journal of Philosophy, 45*, 767–797.

Morton, A. (2012). *Bounded Thinking: Intellectual Virtues for Limited Agents*. Oxford: Oxford University Press.

Paul, L. A., & Quiggin, J. (2018). Real World Problems. *Episteme, 15,* 363–382.

Pettigrew, R. (2012). Accuracy, Chance, and the Principal Principle. *Philosophical Review, 121*, 241–275.

Pettigrew, R. (2016). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.

Sampson, E. (2022). What if Ideal Advice Conflicts? A Dilemma for Idealizing Accounts of Normative Practical Reasons. *Philosophical Studies, 179,* 1091–1111.

Schoenfield, M. (2017). Conditionalization Does Not (in General) Maximize Expected Accuracy. *Mind, 126,* 1155–1187.

Skipper, M., & Bjerring, J. C. (2022). Bayesianism for Non–Ideal Agents. *Erkenntnis, 87,* 93–115.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Staffel, J. (forthcoming). Bayesian Norms and Non–Ideal Agents. In *Routledge Handbook of the Philosophy Evidence,* edited by C. Littlejohn and M. Lasonen–Aarnio. London: Routledge.

Staffel, J. (2017). Should I Pretend I'm Perfect? *Res Philosophica, 94,* 301–324.

Staffel, J. (2019). *Unsettled Thoughts: A Theory of Degrees of Rationality*. Oxford: Oxford University Press.

Thorstad, D. (2021). The Accuracy–Coherence Tradeoff in Cognition. *British Journal for the Philosophy of science,* online first. doi:10.1086/716518

Titelbaum, M. (2015). Rationality's Fixed Point (or: In Defense of Right Reason), 253–294, In *Oxford Studies in Epistemology 5,* edited by T. Szabó Gendler and J. Hawthorne. Oxford: Oxford University Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review, 90*, 293–315.

Wiens, D. (2020). The General Theory of Second Best Is More General Than You Think. *Philosophers' Imprint, 20*, 1–26. hdl:2027/spo.3521354.0020.005.