

Epistemic Sanity or why you shouldn't be opinionated or skeptical

Danilo Fraga Dantas (dfdantas@ucdavis.edu)

Penultimate draft, please cite the published version

<https://doi.org/10.1017/epi.2022.43>

Abstract: I propose the notion of ‘epistemic sanity’, a property of parsimony between the holding of true but not false beliefs and the consideration of our cognitive limitations. Where ‘alethic value’ is the epistemic value of holding true but not false beliefs, the ‘alethic potential’ of an agent is the amount of extra alethic value that she is expected to achieve, given her current environment, beliefs, and reasoning skills. Epistemic sanity would be related to the holding of (true or false) beliefs that increase the agent’s alethic potential (relevant beliefs) but not of beliefs that decrease it (this is related to cognitive parsimony). Suspension of judgment, forgetting, and clutter avoidance are the main contributors to an agent’s epistemic sanity, where this paper focuses on suspension. I argue that rational suspension favors the holding of true and relevant beliefs, which is not the case for the extremes of opinionation (no suspension) and skepticism (general suspension). In the absence of evidence, opinionated agents are often forced to rely on principles such as the principle of indifference, but suspension dominates indifference in terms of alethic value in some conditions. A rational agent would only find it beneficial to adopt skepticism if she considers herself to be an anti-expert about her entire agenda, but then ‘flipping’ beliefs maximizes expected alethic value in relation to skepticism. The study of epistemic sanity results in an ‘impure’ veritism, which can deal with some limitations of veritism (e.g., explaining the existence of false but relevant beliefs). **Keywords:** suspension of judgment; rationality; epistemic utility theory; skepticism.

Introduction

An omniscient being does not need to (maybe, she shouldn't) suspend judgment about any proposition because she knows the truth-value of any proposition. An omnipotent being with unlimited cognitive resources does not need to (maybe, she shouldn't) forget any stored information. After all, she has an unlimited space in memory and unlimited computational power available for searching over any amount of retrieved information. For the same reason, she does not need to avoid cluttering her mind with irrelevant information. We are not such a being! One fundamental fact about our cognitive situation is that “human beings are in the finitary predicament of having fixed limits on their cognitive capacities and the time available to them” (Cherniak, 1986, p. 8). Epistemic rationality (in the following, ‘rationality’) seems to require ‘finite reasoners’ (those in the finitary predicament) to convert efficiently their scarce cognitive resources into epistemic value. Rationality seems to require finite reasoners to exhibit a form of ‘cognitive parsimony’.

This fact is often recognized in the cognitive sciences, where the “tractable cognition thesis” (van Rooij, 2008) is used to constrain the space of computational-level theories of rationality (e.g., Oaksford & Chater, 2007, p. 35). Epistemologists, on the other hand, tend to dismiss considerations about the cognitive limitations of finite reasoners as trading upon *practical* values, whereas epistemology should concern only the maximization of *epistemic* value. Various features (e.g., of sets of beliefs)¹ are regarded as putative sources of epistemic value: closure, coherence, amount of evidential support, etc. In the last decades, some epistemologists have argued for ‘veritism’, the thesis that *the* fundamental source of epistemic value is the believing of truths but not falsehoods: “[T]he fundamental source of epistemic value for a doxastic state is the extent to which it represents the world correctly: that is, its fundamental epistemic value is determined entirely by its truth or falsity” (Pettigrew, 2019b, p. 761). These epistemologists often sought to justify Bayesian norms of rationality in veritistic terms, resulting in the epistemic utility theory (EUT, see Pettigrew, 2019a, for a review).

I intend to propose the notion of ‘epistemic sanity’, a property of parsimony between the holding of true but not false beliefs and the consideration of our cognitive limitations. ‘Sanity’ is sometimes used for attributing the absence of psychiatric illnesses. In this sense, the paradigm ‘epistemic illness’, caused by the lack of attentiveness to the evidence, would be the tendency to hold blatantly false and unjustified beliefs (e.g., conspiracy theorists). But ‘sanity’ also has the meaning of ‘mental health’ and this is the meaning that I

¹ I use ‘belief’ as a general term, which encompasses both full beliefs and credences.

am exploiting here to designate whether a finite reasoner is ‘in a good shape’ for achieving extra epistemic value. Where ‘alethic value’ is the epistemic value of holding true but not false beliefs, the ‘alethic potential’ of an agent is the amount of extra alethic value that she is expected to achieve, given her current environment, beliefs, and reasoning skills. Epistemic sanity would be related to the holding of (true or false) beliefs that increase the agent’s alethic potential (relevant beliefs) but not of beliefs that decrease it (this is related to cognitive parsimony). Suspension of judgment, forgetting, and clutter avoidance (Harman, 1986, p. 12) would be the major contributors to an agent’s epistemic sanity, where this paper focuses on suspension.

Epistemologists often work under a tripartite account of (categorical) doxastic attitudes: a reasoner may believe or disbelieve a proposition, but she may also hold an attitude of neutrality towards it (‘suspension of judgment’)². The mere lack of belief and disbelief should not be sufficient for suspension, as someone who has never considered a proposition does not hold a doxastic attitude (e.g., an attitude of neutrality) towards it (Friedman, 2013b, p. 167). Suspending is an attitude of “committed neutrality” (Sturgeon, 2010, p. 133) in the sense of being able to be adopted or dropped given reasons. Suspending judgment may be cognitively parsimonious because it may avoid the holding of many beliefs³. How suspension could be related to the *increase* of veritistic value is more difficult to devise. I intend to argue that rational suspension is an epistemic virtue between the two vices of opinionation (holding beliefs about every proposition in your agenda) and skepticism (holding beliefs about none) because it would favor the holding of true and relevant beliefs, which is not the case for opinionation or skepticism.

In the absence of evidence, opinionated agents are often forced to rely on principles such as the principle of indifference, but suspension dominates indifference in terms of alethic value in some conditions (see sec. 2.1). Using the expression ‘epistemic sanity’ is especially fortuitous when we are dealing with suspension because it was sometimes considered central to the mental well-being of rational reasoners. The ancient skeptics describe themselves as “the investigators”, but also as “those who suspend” (see Vogt, 2018). They often tell a story where they first find themselves in turmoil due to discrepan-

² Belief and disbelief are often treated as being fundamentally the same attitude (disbelieving ϕ would be the same as believing $\neg\phi$). There are reasons for rejecting this trend (e.g., somebody who is not competent with the negation may be able to disbelieve, see Lord, 2020, fn. 1), but, for simplicity, I will follow the trend.

³ The situation is not so simple because there are different views about suspension and, in some views, suspending demands the adoption of beliefs (e.g., middling credences or second-order full beliefs). In addition, suspending may have its own cognitive cost and fail to be cognitively parsimonious. I discuss these issues in section 2.

cies in how things appear to them. They start hoping to achieve tranquility by settling on what is true and false, but their investigation leads them to find opposing views to be of equal weight. The skeptics then free themselves from turmoil by suspending, which finally brings to them “tranquility”. In section 2.2, I argue that an agent would only find it beneficial to adopt skepticism if she considers herself to be an anti-expert about her entire agenda, but then ‘flipping’ beliefs maximizes expected alethic value in relation to skepticism. Neither opinionation nor skepticism would be cognitively parsimonious because of the high cognitive cost of holding many beliefs (opinionation) and of maintaining suspension in the face of the evidence (skepticism).

In section 1, I argue that the formal models and methods currently used in epistemology are often inadequate for the study of epistemic sanity and propose adequate measures of alethic value and potential. The alethic potential of an agent should be understood as the amount of extra alethic value that she is expected to achieve, given her current environment, beliefs, and reasoning skills. I use the measure of alethic potential in a notion of epistemic relevance and argue that rational finite reasoners should strive to hold only beliefs that are true and relevant in this sense, as to maintain their epistemic sanity. In section 2, I discuss the different views about suspension and the vices of opinionation (sec. 2.1) and skepticism (sec. 2.2) and argue that these vices secure an illusory appearance of rationality by the means of an unreasonable appeal to the minimax principle (sec. 2.3). In the conclusions, I discuss how forgetting and clutter avoidance are related to epistemic sanity and how the study of epistemic sanity results in an ‘impure’ veritism, which can deal with limitations of veritism (e.g., the existence of false but relevant beliefs).

1 Alethic potential

The formal models and methods currently used in epistemology are often inadequate for the study of epistemic sanity. Formal epistemologists often propose normative models as ideal reasoners (i.e., reasoners without cognitive limitations)⁴. For example, although Leitgeb (2014, fn. 3) recognizes that “ultimately, we should be concerned with real-world agents”, his “perfectly rational reasoner” (p. 137) is logically omniscient. The general strategy

⁴ I am agreeing with Garber (1984, p. 101) that normative claims about rationality and descriptive claims about an ideal reasoner are inter-translatable: “The Bayesian thought policeman [who enforce normative claims] might be thought of as clubbing us into behaving like ideal learning machines [ideal reasoners], if we like. Or we can think of the ideal learning machine as an imaginary person who behaves in such a way that he never needs correction by the Bayesian thought police. The two models thus seem inter-translatable”.

seems to be to propose normative models as ideal reasoners “whom we should strive to approximate” (fn. 3). I take issue with this strategy because what is rational for an ideal reasoner may not be rational for a finite reasoner. For example, why should we strive to approximate logical omniscience? Any attempt to do so would result in a form of cognitive paralysis (where all of our scarce cognitive resources would be wasted in deriving logical truths and logical consequences), which would prevent us from fulfilling our (epistemic and practical) goals⁵. But, leaving these concerns aside, the investigation of epistemic sanity would still be hindered because the ideal reasoners that are used as normative models often cannot suspend (forget or avoid clutter).

The ideal reasoners that are used as normative models in epistemology often cannot suspend. For example, EUT has developed some of the best evaluation methods in formal epistemology (e.g., Pettigrew, 2016a). Investigations in EUT usually follow three steps. The first step is to define the ideal set of beliefs: “if a proposition is true in a situation, the ideal credence for an agent in that situation is the maximal credence, which is represented as 1. ...[I]f a proposition is false, the ideal credence in it is the minimal credence, which is represented as 0” (Pettigrew, 2016a, p. 3). The ideal reasoner (the reasoner with the ideal set of beliefs) cannot suspend because she holds a belief-value about every proposition in the agenda (i.e., she is opinionated). This fact by itself suggests that suspension is never the “correct” attitude: “According to the fundamental norm of correct belief, suspending judgment about p is neither correct nor incorrect. If one suspends judgment about p then one has neither got things right nor got things wrong about p ” (Wedgwood, 2002, p. 272). The ideal reasoners that are often used as normative models in epistemology also cannot forget⁶ or avoid clutter⁷.

⁵ For example, believing every logical truth (e.g., $h \vee \neg h$, $(h \vee \neg h) \vee \neg(h \vee \neg h)$, etc) and every logical consequence of some evidence e (e.g., $e \wedge (h \vee \neg h)$, $e \wedge ((h \vee \neg h) \vee \neg(h \vee \neg h))$, etc) is often irrelevant to our goals, but any attempt to do so would result in the spending of our scarce cognitive resources. In addition, attempting to hold those beliefs would not be truth-conducive in general because it would often amplify minor mistakes. For example, if e is false, then all those logical consequences are false as well. These issues are related to the principle of clutter avoidance, which will be discussed in the conclusions.

⁶ The Bayesian ideal reasoner updates her credences by using Bayesian conditionalization (Leitgeb & Pettigrew, 2010). Given the standard Bayesian assumptions of normality and finite additivity, this form of conditionalization is such that if a reasoner reaches maximum credence (certainty) on a proposition at a time, her certainty will be maintained after any subsequent update. For example, the Bayesian ideal reasoner cannot be certain that she is having spaghetti for dinner today (because she is doing so) and forget this irrelevant fact a year later (i.e., lose certainty about it) (see Talbott, 1991, p. 139).

⁷ The Bayesian ideal reasoner has beliefs that are consistent with the axioms of probability, including normality and finite additivity. Normality entails that she must be certain of (i.e., hold maximum credence on) every logical truth. Also, if she comes to learn some

The second step in an investigation within EUT is to define the measure of epistemic value. The most common measure is one of inaccuracy (f), interpreted as the ‘distance’ between an agent’s beliefs and those of the ideal reasoner. The measure of f is often a Brier score: $f(\mathbf{B}) = \sum_{\phi \in \mathbf{B}} (v(\phi) - b(\phi))^2$, where \mathbf{B} is the agent’s belief-set, $v(\phi)$ is ϕ ’s truth-value (i.e., its ideal belief-value), and $b(\phi)$ is ϕ ’s belief-value for the agent. The most natural way in which agents earn epistemic value is by forming *new* beliefs (and not only by adjusting the values of prior beliefs). But if f was the correct measure of epistemic (dis)value, then we would have no reason to form new beliefs because, in doing so, we risk losing (but not earning) epistemic value⁸. If f was the correct measure, then suspension would be a ‘cheap’ way of minimizing epistemic disvalue (e.g., a full skeptic has minimum inaccuracy). The second most common measure of epistemic value is one of accuracy (t , see fn. 8 for an example), which may also be a Brier score: $t(\mathbf{B}) = \sum_{\phi \in \mathbf{B}} (1 - |v(\phi) - b(\phi)|)^2$. Carr (2015, p. 232) argues that the limitations of accuracy as a measure of epistemic value parallel those of inaccuracy: “the situation is reversed... Each new proposition added to the domain of a credence function increases the epistemic function’s epistemic value, as long as the credence it assigns isn’t maximally inaccurate”. The measure of inaccuracy forces skepticism; the measure of accuracy forces opinionation.

Inaccuracy or accuracy are inadequate measures of epistemic value for the investigation of epistemic sanity. These measures cannot assess the choice of holding (or not) some belief (e.g. adopting a new belief) and, consequently, may only be used to compare agents with belief-sets of the same size, which is often achieved by assuming that they are opinionated over a fixed and finite agenda⁹. This is a limitation for the study of epistemic sanity because this investigation demands the comparison between counterfactual situations where an agent holds and does not hold a belief. For example, suspension should be able to be adopted or dropped given reasons, which, in a veritistic framework, amounts to the belief-sets where an agent suspends and adopts a belief being comparable in terms of veritistic value (see also Carr, 2015,

evidence, then normality and finite additivity require her to be certain of every logical consequence of that evidence. This is a form of logical omniscience (see Garber, 1984, p. 104), which is incompatible with clutter avoidance (see fn. 5).

⁸ This parallels the argument against attentiveness to the currently available evidence being *the* fundamental epistemic value (Goldman, 2002, §3). Goldman’s measure of epistemic value is his “V-value”: $t(\mathbf{B}) = \sum_{\phi \in \mathbf{B}} 1 - |v(\phi) - b(\phi)|$ (Goldman, 1999, §3.4).

⁹ Carr (2015, p. 223) also stresses this point: “Epistemic decision theory [EUT] usually presupposes that the credence functions it compares are defined over the same algebra of propositions. Once we abandon this presupposition, new difficulties arise”. Dantas (2021) discusses the difficulties related to the assessment of infinite agendas.

sec. 2.1). In addition, the measures of inaccuracy and accuracy attribute the maximum and the minimum value (respectively) to the absence of beliefs about some proposition in the agent's agenda (Carr, 2015, p. 333), but an adequate measure of epistemic value should attribute to suspension a value that is in between those of holding a false full belief (the minimum value) and a true full belief (the maximum value)¹⁰.

The investigation of epistemic sanity demands both a measure of inaccuracy (f), a measure of accuracy (t), and their integration using an adequate function α (from *alethic* value). The function α should have a set of beliefs B as input and return a numerical value x , where $\alpha(B)$ would depend on the amount of truth (t) and falsehood (f) in B ($\alpha(t, f) = x$). An adequate function $\alpha(t, f)$ must: (r1) strictly increase with respect to (wrt) t (i.e., if $t' > t$, then $\alpha(t', f) > \alpha(t, f)$), which models that 'the more truth the better'; and (r2) strictly decrease wrt f (if $f' > f$, then $\alpha(t, f') < \alpha(t, f)$), which models that 'the less falsehood the better'. Requirements r1 and r2 are accepted by Douven (2013, p. 436)¹¹ and put to work by Trpin and Pellert (2019), who use the function $t - f$. This is the 'minimal' function that fulfills r1 and r2, but consider the 'problem of contradictory pairs'. The function $t - f$ evaluates equally an agent who believes (as to the same degree) both propositions in a contradictory pair and one who believes neither because $t - f = 0$ when $t = f$ (independently of whether $t = f = 0$), but the second agent should be evaluated higher than the first. This problem may be avoided by using weights $Rt - Wf$ where $R < W$ (see Fitelson & Easwaran, 2015, p. 83).

My favorite measure of alethic value is not $Rt - Wf$ but $(t - f)/(t + f + c)$, where $c > 0$ is a 'sensitivity' constant: the smaller the c , the greater the benefit for believing truths and the penalty for believing falsehoods (see Dantas, 2021, for a discussion). I prefer this function, among other things, because (i) it deals more naturally with the problem of contradictory pairs¹² and (ii) it considers the cognitive limitations of finite reasoners¹³. For simplicity, I

¹⁰ The third step in an investigation within EUT is to use the measure of inaccuracy and a principle of decision theory in an argument for some norm of rationality. I will discuss some uses of principles of decision theory in section 2.3.

¹¹ "The basic intuition underlying it is clear enough, to wit, that the higher one's degree of belief in a true proposition is, the more accurate one is, *ceteris paribus*, and also the lower one's degree of belief in a false proposition is, the more accurate one is, *ceteris paribus*" (Douven, 2013, p. 436).

¹² This function is such that $(t - f)/(t + f + c) > (t - f)/(t + f + c + 2x)$ when $t > f$, where x is the degree of belief that the agent holds in each proposition of the contradictory pair. The situation changes when $t \leq f$, but this is a case of anti-expertise where the holding of contradictory beliefs may be epistemically beneficial for the agent (see section 2.2).

¹³ I suppose that a subject believes that ϕ if she is disposed to retrieve some record from her memory and to accept it as a veridical representation that ϕ because (explicit) belief is

will refer the functions $t - f$ and $Rt - Wf$ as α , and postpone the defense of $(t - f)/(t + f + c)$ to another paper (see Dantas, 2021). The function α attributes to the absence of beliefs a value (0) that is in between those of holding a false full belief (negative) and a true full belief (positive). In this context, suspension is not a cheap way of achieving epistemic value, but the attitude of forfeiting a putative increase to avoid a putative decrease of value. Similar considerations hold for forgetting and clutter avoidance.

I propose that the epistemic sanity of an agent should be related to the maximization of her ‘alethic potential’, understood as the amount of extra alethic value that she is expected to achieve, given her current environment, beliefs, and reasoning skills:

Definition 1 (Alethic potential ($\Delta\alpha$)): $\Delta\alpha = \alpha_1 - \alpha_0$,

where α_0 is the α -value of the agent’s current (or initial) set of beliefs and α_1 is the α -value that she is expected to achieve from a priori or a posteriori reasoning, given her current environment, beliefs, and reasoning skills. The notion of alethic potential is not function-dependent and may be adapted to any adequate measure of alethic value. For example, Olsson (2011, p. 128) proposes a Δ -measure in the same lines of $\Delta\alpha$ using Goldman’s V-value as base (see fn. 8)¹⁴. That said, Goldman’s V-value does not fulfill r2.

Why should we care about our alethic potential, especially when a lower initial α -value (α_0) is a factor for increasing it? Should we strive to hold more initial false beliefs (or fewer initial true beliefs) to increase our alethic potential? An anonymous reviewer has proposed the following analogy. Imagine that there is a new kind of utilitarian who proposes the quantity of ‘wealth potential’, defined as the difference between a society’s current level of wealth and its expected level in a future time. As that society produces more wealth, its wealth potential tends to decrease. Suppose that this new kind of utilitarian claims that wealth potential is a neglected value that should be promoted.

usually thought to involve both a mental representation and the positive assessment of it (Bogdan, 1986). Although it is accepted that there is no interesting limit on the amount of information that we can hold in long-term memory (Dudai, 1997), it is accepted that the learning of new information can adversely impact our capacity of retrieving old information and vice versa (“interference”, see Baddeley et al., 2020, p. 291). This cognitive limitation is modeled if $\alpha(t, f) = (t - f)/(t + f + c)$ because $\alpha(t + 2, f) - \alpha(t + 1, f) < \alpha(t + 1, f) - \alpha(t, f)$, which may be interpreted as a diminishing reward for ‘believing too much’.

¹⁴ “Suppose that a question begins to interest agent S at time t_1 [i.e., its possible answers are on her agenda], and S applies a certain practice π in order to answer the question. If the result of applying π is to increase the V-value of the belief states from t_1 to t_2 , then π deserves positive credit. If it lowers the V-value, it deserves negative credit. If it does neither, it is neutral wrt to instrumental V-value” (Olsson, 2011, p. 128).

Should we promote it by ensuring that society doesn't become too prosperous? The reviewer has point, but before discussing it, I want to stress an important dis-analogy. The maximum possible amount of wealth is finite because the amount of ecological resources is finite. Consequently, increasing the amount of wealth of a society necessarily decreases its wealth potential. This is not true for alethic potential because the amount of alethic value is potentially infinite, since there are infinitely many truths 'out there'.

The notion of epistemic sanity should capture the dialectics between the infinite amount of truths 'out there' and the finite cognitive resources that are available for a finite reasoner to 'convert' those truths into alethic value (i.e., into true beliefs, see fn. 13). Then there is indeed a trade-off between rationality, understood as the straightforward maximization of (expected) epistemic value at a moment, and epistemic sanity (this is the reviewer's point). This trade-off is even more apparent if we do not assume that agendas are fixed and finite. If the number of true-believable truths is infinite, then there is no sense in which a finite reasoner can maximize epistemic value at a moment (or even get closer to maximizing it, see Dantas (2021)). In this case, her capacity to keep increasing her epistemic situation is as important as the epistemic value that she currently possesses. This claim does not entail that we should strive to hold more initial false beliefs (or fewer initial true beliefs) because the expected α -value of an agent upon reasoning depends on her initial set of beliefs. False initial beliefs may cause a lower expected α -value by being used as premises for false conclusions. The lack of relevant true beliefs may cause the same by impairing the agent's capacity to draw conclusions from the available evidence.

A rational finite reasoner should increase her current α -value by holding true beliefs that increase her alethic potential (i.e., beliefs that are epistemically relevant to her):

Definition 2 (Epistemic relevance): The belief that ϕ is positively relevant to an agent with a set of beliefs B iff $\Delta\alpha(B \cup \{\phi\}) > \Delta\alpha(B \setminus \{\phi\})$,

where $\Delta\alpha(B)$ is the alethic potential of the agent if her initial set of beliefs were B . The belief that ϕ is irrelevant iff $\Delta\alpha(B \cup \{\phi\}) = \Delta\alpha(B \setminus \{\phi\})$ and it is negatively relevant iff $\Delta\alpha(B \cup \{\phi\}) < \Delta\alpha(B \setminus \{\phi\})$. For simplicity, I will use 'relevant' for positively relevant beliefs and 'non-relevant' for both irrelevant and negatively relevant beliefs. The metaphor is that relevant beliefs 'encapsulate' the alethic value of many beliefs, in such a way that a finite reasoner may have that value 'at her reach' without needing to hold many beliefs. Epistemic sanity would be related to the holding of relevant but not non-relevant beliefs, independent of whether they are true or false.

Epistemic sanity would be an *additional* requirement of rationality for finite reasoners. Rational finite reasoners should maximize α -value by holding true but not false beliefs, but they also should not hold non-relevant beliefs, even if they are true. For example, if I believe that every A is a B , then the belief that a is B for every a that is A may be non-relevant to me (even if they are true) because they are easily derivable from the general belief when it is necessary. I should not hold those as explicit beliefs. A rational finite reasoner should not hold non-relevant beliefs because the holding of those beliefs could be a waste of cognitive resources that does not enhance her epistemic situation. False beliefs may increase the alethic potential of an agent, but only if they promote the acquisition of new true beliefs (i.e., if they are false but relevant beliefs)¹⁵. Nevertheless, the alethic potential of that agent would be further increased if these false beliefs were withdrawn upon investigation. Rational finite reasoners should strive to hold only true and relevant beliefs to maintain their epistemic sanity.

There are technical issues with the measurement of the alethic potential of an agent. For example, how much reasoning should we allow between the measures of α_0 and α_1 ? I think we should let the agent reason until her beliefs become stable (i.e., they would not change if she continued reasoning, see Dantas (2021)). Another issue is that this measurement can hardly be carried out analytically. The measurement is feasible in computational epistemology (e.g., Douven, 2013; Olsson, 2011; Trpin & Pellert, 2019), where epistemologists design computer simulations of agents interacting with environments that are randomly generated from fixed parameters (a class of environments). The measure of α_0 is feasible because we have access to the belief-values of the agent and the features of the environment before a simulation runs. For the same reason, we can measure the α -value of the agent after the simulation halts. The ‘final’ α -value of an agent in a random environment is a contingent notion, but if the number of environments is large enough, then the mean of those values should approximate the agent’s expected α -value after the investigation (α_1). The computational study of epistemic sanity is a matter for another paper.

¹⁵ A Ptolemaic astronomer may use her false beliefs about the deferent and epicycle orbits of the Earth, Mars, and the Sun and what she takes to be their current positions, to predict correctly that Mars will be visible from Earth on September 2, 2003 (Elgin, 2019, p. 26). I will return to false but relevant beliefs in the conclusions.

2 Rational suspension

I intend to argue that rational suspension favors the holding of true and relevant beliefs and that it is cognitively parsimonious. This discussion is not straightforward because there are different views on suspension. The credal view (e.g., Sturgeon, 2008) states that suspending about a proposition is to hold ‘middling’ credences about it and its negation (i.e., credences that do not surpass the thresholds for belief and disbelief). The second-order view (e.g., Raleigh, 2019) states that suspending about a proposition involves (i) not holding first-order beliefs about it or its negation and (ii) holding a second-order belief such as that you cannot yet tell whether that proposition is true or false (Raleigh, 2019, p. 9). The interrogative view (Friedman, 2015) states that suspending about a proposition involves adopting an interrogative attitude about it: to inquire actively about its truth. The anti-interrogative view (Lord, 2020) states that suspending about a proposition involves adopting an anti-interrogative attitude about it: to overlook the evidence about its truth (e.g., because you consider it non-relevant). There is an interesting discussion about which of these views (if any) describes our pre-theoretical notion of suspension¹⁶. I will not attempt to settle this discussion here.

In discussing the epistemic features of suspension, I will assume the ‘normative core’ of the second-order, interrogative, and anti-interrogative views: if a rational agent suspends about a proposition, then she should not hold beliefs about it or its negation¹⁷. This combination of attitudes is impossible by definition in the second-order view (item ii in the last paragraph). Suspending about a proposition while holding beliefs about it or its negation is possible in the interrogative and anti-interrogative views, but this

¹⁶ For example, Lord (2020, p. 134) argues that suspension cannot require second-order beliefs because those beliefs demand some intellectual sophistication that is out of reach of young children and non-human animals who could suspend. Raleigh (2019, p. 10) replies that we hesitate to ascribe suspension to young children and non-human animals because suspension in fact demands some intellectual sophistication.

¹⁷ This assumption does not hold for the credal view as I have described it: if full beliefs and suspension are defined as the holding of certain credences, then it is impossible to suspend about a proposition without holding credences about it. The assumption holds only partially (but sufficiently for our purposes) if the credal view is understood as a normative account about how a subject who holds both full beliefs and credences should update her full beliefs given her credences (e.g., Dorst, 2017). In this case, a rational reasoner who suspends about a proposition should not hold full beliefs about it or its negation. Nevertheless, the credal view (in both versions) may be in tension with our pre-theoretical notion of suspension (Friedman, 2013a, p. 62) because, supposedly, a rational agent who suspends about a number of probabilistic independent propositions should be able to suspend about their conjunction, but, if the number of propositions is large, the view states that she disbelieves it (or should do so).

combination of attitudes should be seen as counternormative. For example, Friedman (2015, p. 11) argues that holding an interrogative attitude about a proposition that you already believe may lead to irrational (because incessant) double-checking (the argument does not apply to middling credences). The holding of an anti-interrogative attitude about a proposition that you believe (including middling credences) would manifest a form of dogmatism that is irrational for fallible reasoners like us. A rational agent who believes a proposition should eventually stop actively looking for evidence about it but should keep herself open to that evidence (if it happens to appear).

The claim that rational suspension avoids the adoption of many beliefs does not declare its cognitive parsimony, because suspending may have its own cognitive cost. The discussion about the cognitive cost of suspending is also not straightforward because of the different views of suspension. There is empirical literature about the cognitive cost of belief formation that is independent of those views and applies to this discussion. The ‘Cartesian model’ states that the acceptance or rejection of incoming information (belief formation) is the product of an effortful assessment process that is after the ‘automatic’ (and relatively costless) processing of the information. Suspending in the presence of evidence would be as effortful as forming a belief about it. The ‘Spinozean model’ states that the acceptance of incoming information is part of its automatic processing and that its rejection occurs after (and is more effortful than) this processing. Suspending in the presence of direct evidence would be more effortful than forming a belief about it. Nadarevic and Erdfelder (2013) presents empirical data in favor of the Cartesian model, while Gilbert (1991), Hasson et al. (2005), and Richter et al. (2009) present data in favor of the Spinozean. I will return to these models and the cognitive parsimony of suspension.

2.1 Opinionation

Opinionation is the non-parsimonious practice of an agent who holds beliefs about every proposition on her agenda, whereas an agent’s agenda is the set of those propositions whose truth-values interest her. Opinionation may be rational when the agent possesses adequate evidence about every proposition on her agenda. The ‘problem of opinionation’ regards how an agent should set her beliefs for propositions that are on her agenda but about which she does not have evidence. This problem is without a solution for opinionated agents who hold only full beliefs. In the absence of evidence for propositions on her agenda, these agents could only adopt random and unmotivated beliefs. The situation is not so obvious for opinionated agents who hold credences because there are normative models regarding credences that require opinionation

(e.g., the Bayesian model) and arguments within those models that prescribe how agents should set their credences in the lack of evidence.

The Bayesian model of rationality comprises the norms of probabilism and conditionalization. Probabilism states that a rational credence function must be consistent with the axioms of probability (see Kolmogorov, 1950). Conditionalization states that rationality requires a reasoner who learns some new piece of evidence to update her previous credences by using Bayesian conditionalization. These norms are supported by Dutch book arguments (see Vineberg, 2016), the accuracy arguments from EUT (e.g., Joyce, 1998; Leitgeb & Pettigrew, 2010), among others (e.g., the arguments from Cox's theorem). The accuracy arguments assume (officially, as a simplifying idealization) that agents are opinionated over a fixed agenda. But opinionation is also a normative consequence of the Bayesian model because probability functions are total and not partial functions and there is no update using Bayesian conditionalization from the absence of belief-values to a belief-value¹⁸. If probabilism and conditionalization are norms of rationality, so must be opinionation.

The Bayesian model of rationality has the following core requirements:

- b1. The agent's beliefs have continuously many values between 0 and 1 (credences);
- b2. The agent's credences are consistent with the axioms of probability;
- b3. The agent updates her credences in the face of additional evidence using Bayesian conditionalization.

Requirement b1 states that the Bayesian agent holds credences. The other requirements guarantee that she fulfills probabilism (b2) and conditionalization (b3). B1-b3 does not determine how a Bayesian agent should set her credences in the absence of evidence (other than that they should be probabilistic). The model is silent about the problem of opinionation. An often discussed particular case of this problem regards priors: how should an agent set her credences at the very beginning of her credal life? The problem of opinionation is more general than that of priors because the agent may still lack evidence about some proposition *after* the investigation and not only *prior* to it. Subjective Bayesians think that b1-b3 exhausts the Bayesian model. But then a Bayesian agent would be in the same position

¹⁸ As Easwaran recognizes: "Situations in which the agent comes to have credences in new propositions seem very different from the standard examples where an agent just learns that some proposition is true. ...Bayesians already know that these cases are difficult ones to account for" (Easwaran, 2013, p. 122).

as an opinionated agent who holds only full beliefs: she may only adopt random and unmotivated beliefs (although probabilistic) in the lack of evidence about some proposition on her agenda. This is unsatisfactory because the Bayesian model requires rational agents to be opinionated.

Objective Bayesians (e.g., Landes & Williamson, 2013) propose principles such as the principle of indifference as a solution to the problem of opinionation¹⁹: in the absence of evidence, agents should distribute their credences equally among the propositions that express the alternative outcomes under consideration²⁰. Pettigrew (2016b) proposes an accuracy argument for the principle of indifference as a solution for the problem of priors²¹. He observes that the indifferent credence function for a set of propositions worst-case dominates any other (opinionated) credence function for that in terms of inaccuracy minimization. For example, consider an exhaustive and exclusive pair of propositions $\{\phi, \psi\}$ (e.g., the contradictory pair, where $\psi \equiv \neg\phi$). The indifferent credence function for this set is such that $cr_0(\phi) = cr_0(\psi) = .5$. For any credence function $cr \neq cr_0$, it is the case that $\min(f(cr, w)) < \min(f(cr_0, w))$, where f is restricted to that set and relativized to the (epistemically?) possible situations w . The minimax principle (also referred to as ‘maximin’) states that a rational agent should minimize the loss in the worst-case scenario (minimize the maximum loss). If minimax was a general principle of rationality, then agents would be required to adopt the indifferent credence function in every situation. This is not the case, but it may be the case that agents should minimize the maximum loss at the beginning of their credal lives (or in the absence of evidence in general). This would be necessary for them to ‘initialize’ their opinionation, although they should update by using Bayesian conditionalization afterward.

Pettigrew’s argument assumes opinionation. Without this assumption, it follows from the same premises that agents should suspend and not adopt indifference in the absence of evidence. If inaccuracy is the measure of epistemic value, then suspending about an exhaustive and exclusive non-unitary set of propositions dominates (and worst-case dominates) indifference. Suspension about such a set guarantees that $f = 0$ and indifference guarantees that $f > 0$ because it requires the agent to hold a positive credence in at

¹⁹ The following considerations also hold for principles that are close to the principle of indifference, such that of maximum entropy (Landes & Williamson, 2013). For example, Landes and Williamson rely on a form of minimax principle in justifying their principle.

²⁰ For simplicity, I have restricted the principle to the finite case. The infinite case is not straightforward because this formulation would be incompatible with countable additivity.

²¹ Pettigrew’s argument targets only the problem of priors (and not that of opinionation) because the principle of decision theory used in the argument (the minimax principle) would only hold at the very beginning of our credal lives (see sec. 2.3).

least one false proposition. The situation is not so obvious if the function α is the measure of epistemic value because, in this case, suspension is not a cheap way of achieving epistemic value. Regardless, suspension dominates indifference in terms of alethic value in some conditions²². Indifference about an exhaustive and exclusive pair of propositions (e.g., a contradictory pair) guarantees that $t = f$ and that $t - f = 0$. Suspension guarantees that $t = f = 0$ and also that $t - f = 0$. So far so good, but if the function α is to deal with the problem of contradictory pairs by adopting weights $Rt - Wf$ such that $R < W$, then suspension still guarantees that $\alpha(t, f) = 0$, but indifference guarantees that $\alpha(t, f) < 0$. Suspension dominates indifference in this case. The same holds for other small exhaustive and exclusive sets non-unitary of propositions, depending on the relative sizes of W and R ²³.

An agent may lack evidence about propositions either at α_0 (before investigating) or at α_1 (afterward). Maintaining indifference about propositions that you still lack evidence after investigation may not be epistemically sane because, in this case, suspending about these propositions may increase α_1 and, consequently, $\Delta\alpha$. Adopting indifference about propositions that you lack evidence before investigating may be epistemically sane, but only if you quit indifference after the investigation (either by suspending or by adopting non-indifferent beliefs). In this case, the agent could have a lower α_0 in comparison to if he suspended, but not necessarily a lower α_1 because suspension does not dominate valued belief functions in general²⁴. The same holds for opinionation. Opinionation is only epistemically sane if agents have evidence for all propositions on her agenda (when they should adopt a non-indifferent belief function) or when she expects to find all the relevant evidence through investigation (when they should adopt a non-indifferent belief function). Consequently, a rational finite reasoner should be opinionated only when she currently possesses, or expects to possess by exploring the

²² This line of reasoning does not apply directly to the credal view, where indifference about non-unitary but small exhaustive and exclusive sets of propositions *is* (or requires) suspending about those propositions. Indifference about large exhaustive and exclusive sets of propositions is (or requires) disbelieving those propositions. Here, the principle of indifference commits agents to the falsity of propositions about which they have no evidence. This is a counter-intuitive prescription, which supports Friedman's complaint (see fn. 17).

²³ If the measure of epistemic value is $\alpha(t, f) = (t - f)/(t + f + c)$, then suspension dominates indifference for exhaustive and exclusive non-unitary set of propositions *of any size* as long as $t > f$ because, in this case, $\alpha(t, f) > \alpha(t + x, f + y)$ for $x, y > 0$.

²⁴ Suspending does not dominate valued belief functions in general because in some situations these functions end up getting things more right than wrong. For example, the function $cr(\phi) = .9$ and $cr(\neg\phi) = .1$ worth more than suspending about this pair in the situation that ϕ is true.

environment adequate, evidence about every proposition on her agenda.

Opinionation should not be assumed (even as a simplifying idealization) in arguments for supporting normative conclusions because it artificially eliminates ‘from the competition’ non-opinionated epistemic practices that may be worth more (expected) alethic value than opinionated practices. In addition, the assumption of opinionation defeats the purpose of veritism, of supporting norms of rationality from truth-conduciveness alone (unless opinionated is also supported in such a way). This issue is highlighted by Littlejohn:

If we want to show that Agnes [an arbitrary agent] really should aspire to have partial beliefs that have certain properties, we need to think of Agnes’ available options as involving suspension and opinionation and we need a value theory that tells that Agnes could be better off for being opinionated. (Littlejohn, 2015, p. 222)

The arguments in the former paragraphs suggest Agnes is only better off being opinionated in some cases (e.g., when she has evidence about every proposition on her agenda). In addition, the assumption of opinionation leaves out of the investigation the most natural way in which agents earn epistemic value, i.e., by acquiring new beliefs. Relaxing this assumption is problematic for EUT because their measure of epistemic value (inaccuracy) cannot be used to compare agents with different numbers of beliefs. The problem is avoided by using the function α as the measure of epistemic value.

2.2 Skepticism

Without claiming historical accuracy, I will follow Comesaña and Klein (2019) in considering the “Pyrrhonian skepticism” (in the following, ‘skepticism’) to be absolute skepticism, i.e. the idea that a rational reasoner should suspend about every proposition on her agenda (‘general suspension’)²⁵. I will also presuppose that general suspension requires an agent not to hold first-order beliefs (see ‘normative core’ in sec. 2). Finally, I will presuppose that skepticism demands a general suspension that is persistent because non-skeptical agents may provisionally suspend (e.g., because they lack evidence) while intending to adopt beliefs as soon as they encounter adequate evidence.

Sextus Empiricus thus narrates the conversion of a rational agent into skepticism:

²⁵ See Frede (1997) for a historically accurate picture of the skeptical notions of belief and suspension.

Men of talent, troubled by the anomaly in things and puzzled as to which of them they should rather assent to, came to investigate what in things is true and what is false, thinking that by deciding these issues they would become tranquil. The chief constitutive principle of skepticism is the claim that to every account an equal account is opposed; for it is from this, we think, that we come to hold no beliefs (Sextus Empiricus, 2000, 1.12)

The idea is that an inquisitive agent will occasionally encounter conflicting evidence about every proposition on her agenda, in which case rationality would require her to adopt general suspension. It is unlikely that an actual agent encounters conflicting evidence about every proposition on her agenda, but the skeptics are prone to offer ‘general counterevidence’²⁶. Independently of the nature of the evidence that would prompt a rational agent to adopt skepticism, there are objective conditions in which she would find it advantageous (or, at least, not disadvantageous). The adoption of skepticism involves general suspension, which amounts to $\alpha_1 = 0$. Consequently, a rational agent would only find it advantageous to adopt skepticism if her total evidence is evidence for $\alpha_0 \leq 0$ (i.e., $t \leq f$). Independently of its nature, the evidence necessary for the rational adoption of skepticism should be evidence sufficient for the belief that $t \leq f$.

Sorensen (1987, p. 312) proposes that someone is anti-expert about a proposition ϕ when ϕ is true iff she does not believe that ϕ . Egan and Elga (2005) generalizes this notion to a set of propositions: “[A]n agent is an anti-expert wrt those propositions if the agent is confident in at least one of them, in the sense that his degree of belief in it is at least 90%; and at least half of the propositions that the agent is confident in are false” (Egan & Elga, 2005, p. 84). Roughly, an agent is an anti-expert wrt a set of propositions iff her beliefs about those propositions are such that $t \leq f$. In this context, the evidence necessary for the rational adoption of skepticism should be evidence sufficient for the agent to consider herself an anti-expert about everything she currently believes. Sorensen argues that it is never rational for an agent to believe herself to be an anti-expert, where his argument relies on two putative requirements of rationality: that a rational agent fulfills probabilism and that she is correct about her own beliefs (transparency)²⁷. Since a rational

²⁶ “When someone propounds to us an argument we cannot refute, we say to him: ‘Before the founder of the school to which you adhere was born, the argument of the school, which is no doubt sound, was not yet apparent, although it was really there in nature. In the same way, it is possible that the argument opposing the one you have just propounded is there in nature but is not yet apparent to us; so we should not yet assent to what is now thought to be a powerful argument’” (Sextus Empiricus, 2000, 1.34).

²⁷ Suppose that a rational agent believes herself to be an anti-expert about ϕ . If she

agent cannot believe herself to be an anti-expert, Sorensen argues, she should suspend about her anti-expertise while maintaining her other beliefs²⁸.

Egan and Elga agree that believing yourself to be an anti-expert is not rational, but they disagree about how a rational agent should react to evidence of her anti-expertise. They suggest that, in the face of evidence of her anti-expertise about a set of propositions, an agent should suspend about those propositions. In their argument, Egan and Elga rely on a somewhat stronger notion of anti-expertise: “when one becomes convinced that one’s all-things-considered judgments in a domain are produced by an anti-reliable process, then one should suspend judgment in that domain” (Egan & Elga, 2005, p. 83). This claim relates anti-expertise not only to $t \leq f$ but also to $\Delta\alpha \leq 0$. Egan and Elga’s proposal is in line with the skeptical story of conversion: if an agent believes herself to be an anti-expert about everything she believes (or may come to believe), then she should adopt general suspension. But there is a problem with the application of Egan and Elga’s reaction to anti-expertise about everything that an agent believes. If you believe you are an anti-expert about everything you believe, then you believe you are an anti-expert about your anti-expertise. Consequently, you should suspend about being an anti-expert about everything you believe. But given that you suspended about your anti-expertise, for which reason should you suspend about your other beliefs?

The arguments of Sorensen and Egan and Elga rely on idealizations (e.g., probabilism) that I cannot assume in a study about epistemic sanity (see fn. 7). I agree with Bommarito (2010) that self-ascriptions of anti-expertise are not necessarily irrational for finite reasoners²⁹. Based on veritistic considerations alone, an anti-expert who believes herself to be an anti-expert is in a better epistemic position than an anti-expert who does not believe herself to be an anti-expert: at the very bottom, the first has one more true belief than the second. In addition, the anti-expertise belief may serve as a flag for

believes that ϕ , she believes she believes that ϕ (transparency). But she also believes (and believes that she believes) that $\neg\phi$ (by modus tollens on the anti-expertise belief), which leads to incoherence.

²⁸ “Since we are warranted in making costly revisions to our background assumptions to escape acceptance of an inconsistent proposition, we are also justified in paying a high price to avoid positions which cannot be consistently accepted” (Sorensen, 1987, p. 312).

²⁹ “Whether or not Perfectly Rational Agents ever find themselves self-ascribing anti-expertise, such self-ascriptions can often be the most rational choice for epistemic mortals like us. In the fortunate cases where we find ourselves able to consciously change our beliefs or withhold regarding a topic, such doxastic changes often require a good deal of time to execute. During this time, to deny our own epistemic failures is not only to be dishonest with ourselves but also to rob ourselves of one of the strongest motives to keep attempting to bring about a change” (Bommarito, 2010, p. 418).

the reasoner to revise her other beliefs. At first, this claim could favor the conversion story of the skeptic because the maintenance of the anti-expertise belief (a second-order belief) could support the general suspension about all the agent's first-order beliefs. If the agent believes that $\alpha_0 \leq 0$ (i.e., $t \leq f$) and that $\Delta\alpha \leq 0$, then adopting skepticism would result in $\alpha_1 = 0$ and $\Delta\alpha \geq 0$. But, from the veritistic point of view, the rational revision from general anti-expertise is not one of general suspension but the 'flipping' of beliefs (i.e., the exchange of values between beliefs and disbeliefs), including the anti-expertise belief.

If an agent believes herself to be an anti-expert about everything that she believes, then she believes that $t \leq f$ and that $\alpha_0 \leq 0$. Adopting skepticism would result in $t = f = 0$ and $\alpha_1 = 0$. But the agent expects (she believes so) that flipping will result in beliefs such that $t > f$ and that $\alpha_1 > 0$ (the inequality is strict because the anti-expertise *disbelief* would also come out true). If a rational agent should maximize expected value, then a rational agent in the presence of evidence of general anti-expertise should not adopt skepticism, but flip beliefs³⁰. The same holds for alethic potential. The (expected) alethic potential of an agent who believes herself to be an anti-expert in Egan and Elga's stronger sense is $\Delta\alpha \leq 0$; the adoption of skepticism would result in $\Delta\alpha \geq 0$. But the alethic potential of flipping is even higher from the perspective of the agent because adopting skepticism results in $\alpha_1 = 0$ and the expected value of flipping is $\alpha_1 > 0$. Finally, skepticism is not cognitively parsimonious. Although it decreases the number of beliefs, both the Cartesian and the Spinozean models agree that suspending in the presence of evidence is at least as expensive as forming beliefs about the evidence. A rational finite reasoner should not adopt skepticism in any situation and she should react to evidence of anti-expertise by flipping beliefs.

2.3 Minimax

Pettigrew (2016b, p. 45) acknowledges that the "primary demerit" of his argument is that it "relies on Minimax, which many will say is not a norm of rational choice". Using the minimax principle to support the principle of indifference depends on the assumption of opinionation because suspension dominates (and worst-case dominates) indifference in some conditions. The unconstrained use of the minimax principle favors the adoption of skepticism. This is the case because suspension worst-case dominates (but does not dominate, see fn. 24) non-indifferent belief functions, except for those

³⁰ The strategy of flipping beliefs exploits the understanding that "an anti-expert is as useful as an expert since you can convert anti-expert beliefs into expert beliefs by accepting their negations" (Sorensen, 1987, p. 312).

that only ascribe maximum and minimum value to necessary and impossible propositions (respectively). Part of the appeal of the skeptical story relies on a covert use of minimax: if there is minimal evidence of your general anti-expertise, in the face of the mere possibility of you being wrong about most of your beliefs, then in the worst-case situation you are wrong about most of your beliefs. In that situation, you would be better off adopting general suspension. Consequently (by the minimax principle), you should adopt general suspension in the actual situation.

Minimax is a very conservative principle of decision-making that makes all but the most risk-averse behavior irrational. Harsanyi (1975, p. 595) comments that the minimax principle was generally accepted in decision theory from the mid-forties to the mid-fifties, but since then the general opinion is that it leads to “serious paradoxes” and “wholly unacceptable practical decisions”: “If you took the maximin principle seriously then you could not ever cross a street (after all, you might be hit by a car); you could never drive over a bridge (after all, it might collapse); you could never get married (after all, it might end in a disaster), etc”. The unrestricted use of minimax as a principle of rationality results in a very conservative model of rationality, which ultimately leads to skepticism³¹. To my mind, it is absurd to care about worst-case maximization when you can maximize *actual* value or expect to do so (e.g., by flipping beliefs). Minimax is a reasonable principle of rationality only when it does not contradict well-established principles, such as dominance and maximization of expected value.

Pettigrew’s reaction to the danger of skepticism to limit the scope of minimax to the beginning of our credal lives: “[I]t [the minimax principle] applies only at the beginning of an agent’s credal life, before she has acquired any evidence and before she has assigned credences to the propositions she entertains. ...For an agent at any other stage of her credal life, Minimax does not apply. Instead, in those situations, the agent ought to maximise her subjective expected utility” (Pettigrew, 2016b, p. 45). We have seen that Pettigrew’s restriction is not sufficient for justifying the principle of indifference without assuming opinionation. This restriction also presupposes opinionation because it assumes that the beginning of an agent’s credal life is the only moment “before she has acquired any evidence” or “before she has assigned credences to the propositions she entertains”. The restriction

³¹ Pettigrew agrees that minimax is a conservative principle of rationality, but he allies himself with such conservatism: “I have no argument for making this alliance. At this point, it seems to me, we have reached normative bedrock: one cannot argue for cognitive conservatism from more basic principles” (Pettigrew, 2016b, p. 46). Pettigrew’s conservatism is not so extreme as to lead to skepticism because he restricts the minimax principle to the beginning of an agent’s credal life.

of minimax that does not presuppose opinionation is to the situations where the agent does not have evidence for some proposition on her agenda. In those situations, the agent should often suspend, but she should do so not because she should act to maximize worst-case value, but because suspension minimizes actual or expected value. These are cases in which minimax does not contradict the well-established principles of rationality.

3 Conclusions

The three major contributors to epistemic sanity have in common that they favor the holding of true and relevant beliefs and are cognitively parsimonious. Rational finite reasoners should suspend in the absence of evidence. Suspending in the absence of evidence avoids opinionation, which is not cognitively parsimonious (it involves the adoption of many beliefs) and whose epistemic value may be dominated by that of suspension in some situations. Suspending in the absence of evidence is cognitively parsimonious independently of the Cartesian and the Spinozean models because, in both models, it does not require dedicated cognitive resources. Section 2.2 claims that a rational finite reasoner should not adopt general suspension in the face of evidence of her anti-expertise because she expects to achieve more epistemic value by flipping her beliefs in the direction that the evidence suggests³². Suspending in the presence of evidence may not be parsimonious because it may have a cognitive cost that is higher than that of belief formation (e.g., if the Spinozean model is correct). In other words, a rational finite reasoner should suspend about propositions that are not ‘suggested’ by the evidence. These propositions are non-relevant in the sense that they are unlikely to be used in inferences with other relevant beliefs as premises.

Forgetting is often considered a cognitive shortcoming. This is a consequence of the idea that the role of memory is simply to store the information acquired in the past and make it available for future use. There is a growing consensus that memory has an active role in information processing (De Brigard, 2013; Klein et al., 2009). In this context, Andonovski (2020) proposes that memory should be understood as a “faculty of triage”, whose role is to make the “right” information available given important constraints of time and cognitive resources. Forgetting would favor cognitive parsimony by, for example, reducing the demands on cognitive central processes that would otherwise be needed to suppress interference (Baddeley et al., 2020,

³² This practice is cognitively parsimonious because it avoids the cognitive cost of rejecting the incoming information, which is costly in both the Cartesian and the Spinozean models.

p. 307). Michaelian (2011) claims that virtuous forgetting may also increase the reliability of a memory system, which is related to the increase of α -value (see fn. 13). Consider a subject whose other cognitive systems are reliable so that the records stored in her memories are accurate at the moment of storage. As the world changes, some of the once-accurate records will become inaccurate. The older the record, the greater the chance that it has become inaccurate. A virtuous memory system would forget ‘older’ records (in the sense of not being retrieved lately) as a means to forget inaccurate records (Michaelian, 2011, p. 407)³³. ‘Older’ records are usually non-relevant in the sense that they are unlikely to be used in inferences with other relevant beliefs as premises.

The principle of clutter avoidance (CA) states that “one should not clutter one’s mind with trivialities” (Harman, 1986, p. 12). Harman motivates his principle by appealing to its cognitive parsimony: “There is a limit to what one can remember, a limit to the number of things one can put into long-term storage, and a limit to what one can retrieve. It is important to save room for important things and not clutter one’s mind with unimportant matters” (Harman, 1986, p. 41-2). The epistemic features of CA are difficult to discuss because of its usual subjective interpretation. Harman (1986, p. 55) warns us not to clutter our minds with “matters in which one has no interest”, but what makes a subject interesting? “Roughly, it’s to have some interest or desire served by having beliefs (or knowing) about the relevant subject matter” (Friedman, 2018, p. 3). This subjective interpretation of CA is not very interesting to epistemology³⁴: a conspiracy theorist certainly finds his favorite conspiracy interesting in this sense, but this is exactly the kind of belief that one should avoid cluttering his mind with. The notion of relevance may be used in an epistemic interpretation of CA: one may believe that ϕ only if ϕ is relevant to her. Then clutter avoidance would contribute to epistemic sanity by avoiding the holding of non-relevant beliefs³⁵. The

³³ This claim is supported by empirical research. Schooler and Hertwig (2005), for example, elaborate on the notion of “beneficial forgetting” by proposing that losing information may aid the recognition heuristic, which relies on failures of recognition to infer which of two objects scores higher on a criterion.

³⁴ Friedman (2018, p. 15) discusses the consequences of this interpretation to epistemology and concludes that “we’re left with a highly interest-driven picture of how we ought to revise our doxastic states”. For example, if CA is a meta-principle that constrains principles of belief revision, then “all sorts of purely evidentialist and reliabilist potential norms are not genuine norms” (Friedman, 2018, p. 9).

³⁵ Michaelian (2011, p. 419) proposes the analogous principle of clutter elimination (CE): “if one’s mind is cluttered with trivialities, one should remove them”. The epistemic interpretation of CE is the contrapositive of the epistemic interpretation of CA: if the belief that ϕ is non-relevant to an agent, then she ought not to believe that ϕ (e.g., she

study of forgetting and clutter avoidance will be carried out in other papers.

There is residual a problem with the notion of epistemic sanity that is related to the reviewer’s point. An agent may still ‘pump’ her alethic potential by initially holding false but irrelevant beliefs and dumping them during the investigation phase. This strategy may increase her $\Delta\alpha$ in such a way that is not rational or cognitively parsimonious. There is a slightly different notion of epistemic sanity that may be used to deal with this problem. The ‘mean relevance’ of the beliefs of an agent with a set of beliefs \mathbf{B} and alethic potential of $\Delta\alpha$ is the following quantity:

Definition 3 (Mean relevance): $\frac{\Delta\alpha}{|\mathbf{B}|}$,

where $|\mathbf{B}|$ is the number of beliefs in \mathbf{B} . The maximization of mean relevance is related to epistemic sanity not only because this quantity varies directly wrt the alethic potential ($\Delta\alpha$), but also because it varies inversely wrt to the number of beliefs of the agent ($|\mathbf{B}|$). The maximization of mean relevance demands the holding of few but relevant beliefs³⁶. Opinionation tends to decrease the mean relevance because it demands the holding of non-relevant beliefs (e.g., indifferent beliefs). Skepticism would be a degenerate case of suspension that does not increase mean relevance (this value is not even defined for full skeptics). Finally, rational suspension increases mean relevance because it involves the maintenance of fewer but relevant beliefs.

The notion of alethic potential is relative to an agent (e.g., her beliefs and reasoning skills) but also to which environment the agent is (because α_1 is calculated as the mean α -value in similar environments). In this sense, the investigation of epistemic sanity gives rise to a form of ‘impure veritism’, where the measure of epistemic value is veritistic but where which evidence is available is also relevant. The notion of mean relevance may be used in a quantitative notion of relevance for individual beliefs. The relevance of a belief that $\phi \in \mathbf{B}$ may be defined as

$$\frac{\Delta\alpha(\mathbf{B})}{|\mathbf{B}|} - \frac{\Delta\alpha(\mathbf{B} \setminus \{\phi\})}{|\mathbf{B} \setminus \{\phi\}|},$$

where \mathbf{B} is the set of beliefs of the agent. The measure of individual relevance enables impure veritism to deal with some limitations of ‘pure’ veritism. For example, the measure of individual relevance may be used to explain the existence of “epistemically useful falsehoods” (Elgin, 2019). These are

should forget that ϕ).

³⁶ The measure of mean relevance is also feasible within computational epistemology, but this investigation is also a matter for another paper.

false beliefs with high relevance, which promote the adoption of true and relevant beliefs by the agent (see fn. 15 for an example). The measure of individual relevance may also be used to answer some criticisms of veritism. For example, DePaul (2001) claims that veritism implies that all true beliefs are equally epistemically valuable, but that this implication is false because there are cases where two sets each containing an equal number of true beliefs intuitively differ in epistemic value. The impure veritism does not support that implication because (the beliefs in) the two sets of beliefs may differ in their epistemic relevance. These are matters for other papers.

References

- Andonovski, N. (2020). Memory as triage: Facing up to the hard question of memory. *Review of Philosophy and Psychology*, 12(2), 227–256.
- Baddeley, A., Eysenck, M., & Anderson, M. (2020). *Memory* (3rd ed.). Routledge.
- Bogdan, R. J. (1986). The importance of belief. In R. Bogdan (Ed.), *Belief: Form, content, and function* (pp. 1–16). Oxford University Press.
- Bommarito, N. (2010). Rationally self-ascribed anti-expertise. *Philosophical Studies*, 151(3), 413–419.
- Carr, J. (2015). Epistemic expansions. *Res Philosophica*, 92(2), 217–236.
- Cherniak, C. (1986). *Minimal Rationality*. MIT Press.
- Comesaña, J., & Klein, P. (2019). Skepticism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of Philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University.
- Dantas, D. (2021). How to (blind)spot the truth: An investigation on actual epistemic value. *Erkenntnis*.
- De Brigard, F. (2013). Is memory for remembering?: Recollection as a form of episodic hypothetical thinking. *Synthese*, 191(2), 155–185.
- DePaul, M. (2001). Value monism in epistemology. In M. Steup (Ed.), *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue*. (pp. 170–182). Oxford University Press.
- Dorst, K. (2017). Lockeans maximize expected accuracy. *Mind*, 128(509), 175–211.
- Douven, I. (2013). Inference to the best explanation, Dutch books, and inaccuracy minimisation. *The Philosophical Quarterly*, 63(252), 428–444.
- Dudai, Y. (1997). How big is human memory, or on being just useful enough. *Learning & memory*, 3(5), 341–365.

- Easwaran, K. (2013). Expected accuracy supports conditionalization - and conglomerability and reflection. *Philosophy of Science*, 80(1), 119–142.
- Egan, A., & Elga, A. (2005). I can't believe I'm stupid. *Philosophical Perspectives*, 19(1), 77–93.
- Elgin, C. (2019). Epistemically useful falsehoods. In B. Fitelson, R. Borges, & C. Braden (Eds.), *Themes from Klein* (pp. 25–38). Springer.
- Fitelson, B., & Easwaran, K. (2015). Accuracy, coherence and evidence. In T. Gendler & J. Hawthorne (Eds.), *Oxford studies in Epistemology, volume 5*. Oxford University Press.
- Frede, M. (1997). The skeptic's beliefs. In M. Burnyeat & M. Frede (Eds.), *The original sceptics: A controversy* (pp. 1–25). Hackett Publishing Co, Inc.
- Friedman, J. (2013a). Rational agnosticism and degrees of belief. In T. Gendler & J. Hawthorne (Eds.), *Oxford studies in Epistemology volume 4* (pp. 57–81). Oxford University Press.
- Friedman, J. (2013b). Suspended judgment. *Philosophical Studies*, 162(2), 165–181.
- Friedman, J. (2015). Why suspend judging? *Noûs*, 51(2), 302–326.
- Friedman, J. (2018). Junk beliefs and interest-driven Epistemology. *Philosophy and Phenomenological Research*, 97(3), 568–583.
- Garber, D. (1984). Old evidence and logical omniscience in Bayesian confirmation theory. In J. Earman (Ed.), *Testing scientific theories* (pp. 99–131). University of Minnesota Press.
- Gilbert, D. (1991). How mental systems believe. *American psychologist*, 46(2), 107.
- Goldman, A. (1999). *Knowledge in a social world*. Oxford University Press.
- Goldman, A. (2002). The unity of the epistemic virtues. In *Pathways to knowledge* (pp. 51–72). Oxford University Press.
- Harman, G. (1986). *Change in view: Principles of reasoned revision*. MIT Press.
- Harsanyi, J. C. (1975). Can the maximin principle serve as a basis for morality? a critique of John Rawls's theory. *The American Political Science Review*, 69(2), 594–606.
- Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological science*, 16(7), 566–571.
- Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science*, 65(4), 575–603.
- Klein, S., Robertson, T., & Delton, A. (2009). Facing the future: Memory as an evolved system for planning future acts. *Memory & Cognition*, 38(1), 13–22.

- Kolmogorov, A. (1950). *Foundations of Probability*. Chelsea Publishing Company.
- Landes, J., & Williamson, J. (2013). Objective Bayesianism and the maximum entropy principle. *Entropy*, *15*, 3528–3591.
- Leitgeb, H. (2014). The stability theory of belief. *The Philosophical Review*, *123*(2), 131–171.
- Leitgeb, H., & Pettigrew, R. (2010). An objective justification of Bayesianism II: The consequences of minimizing inaccuracy. *Philosophy of Science*, *77*(2), 236–272.
- Littlejohn, C. (2015). Who cares what you accurately believe? *Philosophical Perspectives*, *29*(1), 217–248.
- Lord, E. (2020). Suspension of judgment, rationality’s competition, and the reach of the epistemic. In S. Schmidt & G. Ernst (Eds.), *The Ethics of belief and beyond: Understanding mental normativity* (pp. 126–145). Routledge.
- Michaelian, K. (2011). The Epistemology of forgetting. *Erkenntnis*, *74*(3), 399–424.
- Nadarevic, L., & Erdfelder, E. (2013). Spinoza’s error: Memory for truth and falsity. *Mem. Cognit.*, *41*(2), 176–186.
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.
- Olsson, E. (2011). A simulation approach to veritistic social Epistemology. *Episteme*, *8*(2), 127–143.
- Pettigrew, R. (2016a). *Accuracy and the laws of credence*. Oxford University Press.
- Pettigrew, R. (2016b). Accuracy, risk, and the principle of indifference. *Philosophy and Phenomenological Research*, *92*(1), 35–59.
- Pettigrew, R. (2019a). Epistemic utility arguments for probabilism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University.
- Pettigrew, R. (2019b). Veritism, epistemic risk, and the swamping problem. *Australasian Journal of Philosophy*, *97*(4), 761–774.
- Raleigh, T. (2019). Suspending is believing. *Synthese*, 1–26.
- Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don’t have to believe everything you read: Background knowledge permits fast and efficient validation of information. *J. Pers. Soc. Psychol.*, *96*(3), 538–558.
- Schooler, L., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, *112*(3), 610–628.
- Sextus Empiricus. (2000). *Outlines of scepticism* (J. Annas & J. Barnes, Eds.). Cambridge University Press.

- Sorensen, R. (1987). Anti-expertise, instability, and rational choice. *Australasian Journal of Philosophy*, 65(3), 301–315.
- Sturgeon, S. (2008). Reason and the grain of belief. *Noûs*, 42(1), 139–165.
- Sturgeon, S. (2010). Confidence and coarse-grained attitudes. In T. Gendler & J. Hawthorne (Eds.), *Oxford studies in Epistemology volume 3* (pp. 126–149). Oxford University Press.
- Talbott, W. (1991). Two principles of Bayesian Epistemology. *Philosophical Studies*, 62(2), 135–150.
- Trpin, B., & Pellert, M. (2019). Inference to the best explanation in uncertain evidential situations. *British Journal for the Philosophy of Science*, 70(4), 977–1001.
- van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, 32(6), 939–984.
- Vineberg, S. (2016). Dutch book arguments. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2016). Metaphysics Research Lab, Stanford University.
- Vogt, K. (2018). Ancient skepticism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of Philosophy* (Fall 2018). Metaphysics Research Lab, Stanford University.
- Wedgwood, R. (2002). The aim of belief. *Philosophical Perspectives*, 16, 267–97.