How to (Blind)Spot the Truth: an investigation on actual epistemic value

Danilo Fraga Dantas (dfdantas@ucdavis.edu),
Forthcoming in Erkenntnis.

Abstract: This paper is about the alethic aspect of epistemic rationality. The most common approaches to this aspect are either normative (what a reasoner ought to/may believe?) or evaluative (how rational is a reasoner?), where the evaluative approaches are usually comparative (one reasoner is assessed compared to another). These approaches often present problems with blindspots. For example, ought a reasoner to believe a currently true blindspot? Is she permitted to? Consequently, these approaches often fail in describing a situation of alethic maximality, where a reasoner fulfills all the alethic norms and could be used as a standard of rationality (as they are, in fact, used in some of these approaches). I propose a function α, which accepts a set of beliefs as input and returns a numeric alethic value. Then I use this function to define a notion of alethic maximality that is satisfiable by finite reasoners (reasoners with cognitive limitations) and does not present problems with blindspots. Function α may also be used in alethic norms and evaluation methods (comparative and non-comparative) that may be applied to finite reasoners and do not present problems with blindspots. A result of this investigation is that the project of providing purely alethic norms is defective. The use of function α also sheds light on important epistemological issues, such as the lottery and the preface paradoxes, and the principles of clutter avoidance and reflection. **Keywords:** Epistemic Utility Theory; Blindspots; Bounded Rationality; Computational Epistemology.

1 Introduction

Practical rationality is related to the maximization of practical value, which is often associated with the fulfillment of the subject's preferences (e.g. in Decision Theory, see Steele and Stefansson, 2016). Accordingly, epistemic rationality (in the following, 'rationality') is often related to the maximization of some epistemic value (e.g. in Epistemic Utility Theory, see Pettigrew, 2019a). Different features (e.g. of sets of beliefs) are regarded as sources of epistemic value: closure, coherence, amount of evidential support, computational efficiency (see Dantas, 2017), etc. In recent decades, epistemologists have argued that *the* fundamental source of epistemic value is the fulfillment of the truth-goal, the goal of believing truths rather than 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. ...[A]ny further source of value for a belief is derivative, not fundamental: that is, it is valuable because of the extent to which it promotes or otherwise serves the fundamental source (Pettigrew, 2019b, p. 761).

This paper is about this *alethic* aspect of rationality: the aspect of rationality related to the fulfillment of the truth-goal. The existing approaches to this aspect of rationality are either normative (what a reasoner ought to/may believe?) or evaluative (how rational is a reasoner?), where the evaluative approaches are usually comparative (one reasoner is evaluated compared to another). These approaches often present problems with blindspots, which are possible but inaccessible propositions (e.g. propositions that are possibly true but impossible to be true when believed by a reasoner, see Sorensen, 1988). For example, ought a reasoner to believe a currently true blindspot? Is she permitted to? Consequently, these approaches often fail in describing a situation of alethic maximality (α-maximality), where a reasoner somehow fulfills the truth-goal 'completely'. An α-maximal reasoner, a reasoner who fulfills all the alethic norms, is used as a standard of rationality in some evaluative frameworks. Nevertheless, these approaches often fail in making sense of such a reasoner. For example, does this reasoner believe a currently true blindspot? Ought she to?

Instead of using alethic norms or evaluation methods to define a notion of α-maximality, I intend to pursue the opposite direction. I will define a function α, which accepts a set of beliefs as input and returns a numeric value, and then use this function to define a notion of α-maximality that does not present problems with blindspots. Finally, I will use this notion to define alethic norms and evaluation methods (comparative and non-comparative) that do not present problems with blindspots. It is usually accepted that rationality has a *subjective* character, e.g. related to what a reasoner finds epistemically possible (Wedgwood, 2015, p. 221). For this reason, rationality is usually related to the maximization of *expected* epistemic value or, at least, to the maximization of epistemic value for all epistemic possibilities (e.g. in Epistemic Utility Theory). The evaluation methods proposed in this paper (see sec. 4.2) use *actual* epistemic value. The evaluation of (sets of beliefs of) reasoners from their actual epistemic value would be an *objective* form of evaluation because the truth-value of beliefs is often not transparent for reasoners (see fn. 7). This form of evaluation should be used with caution in order to take into account the subjective character of rationality (see sec. 4.2). Measures of actual epistemic value are currently used, for example, in Computational Epistemology (e.g. Douven, 2013), where epistemologists have full access to the truth-values of their agents' beliefs[1].

[1] Douven (2013) uses computer simulations to resist the conclusion that Bayesian conditionalization should be regarded as *the* rational rule for updating credences on the grounds that if a reasoner uses a rule

I am interested in investigating which approach to the alethic aspect of rationality can be used for describing a notion of α-maximality that is satisfiable by *finite reasoners*, i.e. reasoners with cognitive limitations (e.g. humans)[2]. My focus on finite reasoners is due to the fact that, although (human) Epistemology is (or should be) mostly concerned with human rationality, epistemologists often propose standards of rationality that are only satisfiable by reasoners without cognitive limitations[3]. I take issue with this state of affairs because standards of rationality for finite reasoners may differ substantially from standards of rationality for reasoners without cognitive limitations[4]. However, even if we leave to one side these concerns about finite reasoning, the following objections to the existing approaches to the alethic aspect of rationality are independent of these cognitive limitations. Also, my alethic norms and evaluation methods are equally applicable to both finite reasoners and reasoners without cognitive limitations. However, I do find it to be an advantage of my proposal that it generates a notion of α-maximality that is satisfiable by finite reasoners and allows for alethic norms and evaluation methods that can be used to assess finite reasoners.

In Section 2, I show that most normative and comparative evaluative approaches to the alethic aspect of rationality present problems with blindspots. Normative approaches either make α-maximality impossible or leave this notion undefined; comparative evaluative

at odds with conditionalization, then the reasoner himself would find that conditionalization minimizes *expected* inaccuracy (in relation to her own rule) given her own credences (Leitgeb and Pettigrew, 2010b). He constructs simulations where two reasoners, a Bayesian and an "explanationist", who updates credences using a version of the inference to the best explanation, watch sequences of coin tosses and must estimate the extent of the coin's bias. The Bayesian minimizes expected inaccuracy of every update, but the eplanationist minimizes actual inaccuracy at the end of the sequence (*in most trials*). Douven (2013, p. 438) remarks that "it would seem absurd to claim that it is epistemically more important to have an update rule that minimises expected inaccuracy than to have one that actually minimises inaccuracy".

[2] Informally, a finite reasoner has cognitive limitations such as finite input (perception, etc transmits only a finite amount of information), finite memory (memory can store only a finite amount of information), and finite computational power (reasoning can execute only finitely many operations, such as inferential steps, in a finite time interval). I discuss the notion of a finite reasoner in Section 3.3.

[3] 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. Finite reasoners cannot, even in principle, (explicitly?) believe all the infinitely many logical consequences of their beliefs because of their finite memory.

[4] For this reason, I partially disagree with the remark in Leitgeb (2014, fn. 3) that it is (always) a good methodological strategy to describe ideal agents (without cognitive limitations), "whom we should strive to approximate". For example, I do not believe that we should strive to be logically omniscient. In other words, I do not believe that we should strive to clutter our (limited) memory with (infinitely) many logical truths and logical consequences (see Harman, 1986, p. 12). This attempt would lead to a form of cognitive paralysis since we would spend all of our cognitive resources in deriving logical truths and consequences, which would prevent us from fulfilling our (epistemic or practical) goals. Also, attempting to believe all the logical consequences of our beliefs is not truth-conducive in general because believing the logical consequences of a false belief may result in an amplification of the initial mistake. My concerns with finite reasoning are also related to the fact that reasoners without cognitive limitations are not fully implementable in a computer simulation. It is an advantage of my framework that it enables the investigation of α-maximal (but finte) reasoners using the tools of Computational Epistemology.

approaches often make 'rationality' ambiguous between several incommensurable notions. Consequently, these evaluative approaches often fail in making sense of an α-maximal reasoner, who they often treat as a standard of rationality for all reasoners. Section 3 discusses the properties of an adequate non-comparative evaluation function for measuring the alethic value of a set of beliefs. I use this function in a notion of α-maximality that is satisfiable in the face of blindspots and by finite reasoners. In Section 4, I use function α to generate alethic norms and evaluative methods (comparative and non-comparative) to assess rationality, which do not present problems with blindspots and which can be applied to finite reasoners. It is a finding of this Section that the whole project of providing purely alethic norms would be defective. In this same section, I also discuss how the use of function α sheds light on important epistemological issues, such as the lottery and the preface paradoxes, and the principles of clutter avoidance and reflection.

2 Blindspots

Broadly speaking, a blindspot is a possible but inaccessible proposition, where the relevant accessibility relation is determined by a propositional attitude (Sorensen, 1988). In this context, a proposition ϕ is a true-belief blindspot (in the following, 'blindspot') for a reasoner $\mathcal{R}$ iff ϕ is possibly true but is not truly-believable by $\mathcal{R}$[5]. For example, the proposition that I do not exist is a blindspot for me because it is possible that I do not exist (e.g. I could have died as a kid), but it is not possible that I do not exist while I believe that I do not exist. You, on the other hand, can believe (truly) that I do not exist (e.g. after I die). Consider the following two more abstract blindspots for an arbitrary reasoner $\mathcal{R}$:

($bs1$) $\mathcal{R}$ does not believe that $bs1$.

($bs2$) $\phi \wedge \psi$, where ϕ is a mundane truth[6] and ψ states that $\mathcal{R}$ does not believe that ϕ.

Sentence $bs1$ is the core example of a blindspot because it has the property that $bs1$ is true iff $\mathcal{R}$ does not believe that $bs1$. Whether $bs2$ is a blindspot depends on some presuppositions. It seems reasonable to presuppose that a rational reasoner believes the 'easy' logical consequences of her beliefs resulting from conjunction introduction or conjunction elimination. In this case, $bs2$ would be a blindspot for a rational $\mathcal{R}$ because if $\mathcal{R}$ believes that $bs2$, then $\mathcal{R}$ believes its conjunct ϕ, but then the second conjunct of

[5] In general, a blindspot for a propositional attitude A and a reasoner $\mathcal{R}$ is a proposition that is possibly true but cannot have attitude A taken towards it by $\mathcal{R}$. I am dealing only with the case of $A =$ true-belief. A proposition ϕ is truly-believable by $\mathcal{R}$ iff it is possible that ϕ is true while believed by $\mathcal{R}$.

[6] A mundane truth is a contingent truth that is not about (the beliefs of) the reasoner (e.g. the proposition that snow is white).

$bs2$ (i.e. ψ) is false and $bs2$ itself is false. Sentence $bs2$ would not be truly-believable by a rational $\mathcal{R}$. If $bs2$ is a blindspot, then it is a *complex blindspot*, i.e. a blindspot with components that are not blindspots. Complex blindspots are interesting because they often have components that are harmless truths (e.g. ϕ in $bs2$), which only form a blindspot when they are in a conjunction with the other components. Any harmless truth can be a component of a complex blindspot.

2.1 Blindspots and norms

The goal of a normative approach to the alethic aspect of rationality is to provide the correct and complete set of alethic norms about what a rational reasoner ought to believe and may believe[7]. A set of norms does not automatically generate a notion of α-maximality, but it seems reasonable to presuppose that an α-maximal reasoner believes all and only that which she ought to believe and only that which she may believe. This is the case because alethic norms are supposed to model the fulfillment of the truth-goal, which an α-maximal reasoner does completely. The crudest norm of this kind is the following:

(n1) $\mathcal{R}$ ought to believe that ϕ iff ϕ is true.

The problem with n1 is that it makes α-maximality impossible because of blindspots (Bykvist and Hattiangadi, 2007). Let $\mathcal{R}$ be α-maximal. Then $\mathcal{R}$ believes that $bs1$ iff $\mathcal{R}$ ought to believe that $bs1$ (α-maximality), but also $\mathcal{R}$ ought to believe that $bs1$ iff $bs1$ is true (n1) and $bs1$ is true iff $\mathcal{R}$ does not believe that $bs1$ ($bs1$). Then $\mathcal{R}$ believes that $bs1$ iff $\mathcal{R}$ does not believe that $bs1$, which is a contradiction. Carr (2020) investigates a similar problem for n1 (and related norms) using propositions that are closely related to blindspots. She shows that the problem persists in different modal interpretations of n1.

Modified versions of n1 were proposed in order to avoid this *general problem of blindspots*. Boghossian (2003) avoids the problem by rejecting one direction of n1. He accepts

(n2) If $\mathcal{R}$ ought to believe that ϕ, then ϕ is true,

but rejects:

(n2') If ϕ is true, then $\mathcal{R}$ ought to believe that ϕ.

[7] Philosophers often distinguish between subjective and objective norms (e.g. Carr, 2020), where the subjective but not the objective norms are sensitive to the information that is available to the reasoner (e.g. what is epistemically possible for the reasoner). For this reason, subjective norms are often *deliberative*, in the sense of being action-guiding. Alethic norms are objective because the truth-value of beliefs is often not transparent for reasoners. Consequently, these norms are often not deliberative, except in some cases when the relevant truth-values are transparent (e.g. contradictions and blindspots). The same holds for evaluation methods.

The retraction from n1 to n2 avoids the general problem of blindspots. It does not follow from n2 that $\mathcal{R}$ ought to believe a blindspot. But n2 is too weak in the sense that, regardless of ϕ being true or false, it does not entail an obligation (Raleigh, 2013, p. 248). If ϕ is true, n2 does not entail anything. If ϕ is false, n2 entails that $\mathcal{R}$ does not ought to believe that ϕ, but not that $\mathcal{R}$ ought not to believe that ϕ. As a consequence, n2 underdetermines the description of α-maximality: n2 states what α-maximal reasoners do not (ought to) believe, but it states nothing about what they (ought to) believe, which defeats the purpose of defining a notion of α-maximality. This problem is revealed, for example, by the fact that n2 is fulfilled when reasoners do not have beliefs. In other words, norm n2 does not model the positive part of the truth-goal: 'believe the truths'.

Raleigh (2013, p. 249) and Whiting (2010) propose a different norm:

(n3) $\mathcal{R}$ may believe that ϕ iff ϕ is true.

This formulation also avoids the general problem of blindspots. It does not follow from n3 that $\mathcal{R}$ ought to believe a blindspot. But n3 has a different problem with blindspots: it is a consequence of n3 that a reasoner is permitted to believe a currently true blindspot, and yet if the reasoner were to believe that blindspot, n3 itself would forbid that belief. Raleigh argues that this is a good feature of his proposal:

> It seems to me that as much as we may have an intuition that attempting to believe a blindspot that happens to be true should be discouraged, we also have a conflicting intuition that we *would like* to believe such a true proposition... A feature of my proposal is that there is no truth-norm forbidding belief in a currently [true] non-believed blindspot, but there is a norm that kicks in the moment that one ventures to believe the blindspot. I think this 'may' be seen as a virtue rather than a vice, in that it does justice, as far as possible, to *both* of these conflicting intuitions (Raleigh, 2013, p. 252).

I do not believe that appealing to our conflicting intuitions about blindspots helps here because the instantiation of n3 for a blindspot contradicts the very definition of the epistemic 'may'[8]. Suppose that $\mathcal{R}$ does not believe that $bs1$. Then $bs1$ is true and n3 states that $\mathcal{R}$ may believe that $bs1$. If $\mathcal{R}$ may believe that $bs1$, then there must exist a permissible epistemic situation in which $\mathcal{R}$ believes that $bs1$. However, no such permissible epistemic situation exists according to n3 itself: in all situations in which $\mathcal{R}$ believes that $bs1$, $bs1$ is false, and n3 forbids $\mathcal{R}$ to believe that $bs1$. Another problem is, again, that n3 is fulfilled by not having beliefs.

Another strategy to deal with the initial problem consists of adding conditions to n1[9]:

[8] In deontic logic, 'may' is usually modeled as diamond-like (as requiring truth in a permissible situation, where a permissible situation is one in which nobody does what is not permitted by the norms). The diamond does not aggregate over conjunctions in the standard modal logics.

[9] There exist interesting versions of n4 and n5 without the clause 'ϕ is true' (let those be n4' and n5'), but similar problems hold for n4 and n5 and to n4' and n5'. About norms n4' and n5', see footnote 40.

(n4) $\mathcal{R}$ ought to believe that ϕ iff (ϕ is true and truly-believable by $\mathcal{R}$).

(n5) $\mathcal{R}$ ought to believe that ϕ iff (ϕ is true and would be true were $\mathcal{R}$ to believe that ϕ).

Norm n4 avoids the general problem of blindspots. Since $bs1$ is not truly-believable by $\mathcal{R}$, it does not follow from n4 alone that $\mathcal{R}$ ought to believe that $bs1$. However, it follows from n4 and some reasonable assumptions that $\mathcal{R}$ ought to believe a complex blindspot (Bykvist and Hattiangadi, 2013, p. 110). Let this be the *problem of complex blindspots*. Suppose that $\mathcal{R}$ does not believe that ϕ. Then ψ is true and truly-believable by $\mathcal{R}$, and it follows from n4 that $\mathcal{R}$ ought to believe that ψ. Also, ϕ is true and truly-believable by $\mathcal{R}$ and it follows from n4 that $\mathcal{R}$ ought to believe that ϕ. Given this, $\mathcal{R}$ ought to believe that ϕ and $\mathcal{R}$ ought to believe that ψ. If 'ought' aggregates over conjunctions, then $\mathcal{R}$ ought to (believe that ϕ and believe that ψ). Since $\phi \wedge \psi$ is an easy logical consequence of ϕ and ψ, then a rational $\mathcal{R}$ ought to believe that $\phi \wedge \psi$, which is $bs2$. It is usually accepted that 'ought' aggregates over conjunctions[10]. In this case, it follows from n4 that a rational $\mathcal{R}$ ought to believe that $bs2$, which contradicts n4 itself because $bs2$ is not truly-believable by $\mathcal{R}$. The same holds for n5 and related norms (n4' and n5', see fn. 9).

Both the general and the problem of complex blindspots are avoided by n6 and n7:

(n6) $\mathcal{R}$ may believe that ϕ iff (ϕ is true and truly-believable by $\mathcal{R}$).

(n7) $\mathcal{R}$ may believe that ϕ iff (ϕ is true and would be true were $\mathcal{R}$ to believe that ϕ).

According to n6, it is not the case that $\mathcal{R}$ may either believe that $bs1$ or believe that $bs2$ because both $bs1$ and $bs2$ are not truly-believable by $\mathcal{R}$. Also, we cannot reinstate the argument for the problem of complex blindspots from n4-n5 to n6-n7 because it is usually not accepted that 'may' aggregates over conjunctions (see fn. 8). The problem with n6 and n7 is the underdetermination of the description of α-maximality: these norms state what α-maximal reasoners may believe, but it does not say anything about what they (ought to) believe. This defeats the purpose of defining a notion of α-maximality. This problem is, again, revealed by the fact that n6 and n7 are fulfilled when reasoners do not have beliefs. Again, these norms do not model the positive part of the truth-goal.

2.2 Blindspots and evaluations

The goal of an evaluative approach to the alethic aspect of rationality is to assess the rationality of reasoners using some measure related to the fulfillment of the truth-goal. In a comparative evaluative approach, a reasoner is evaluated in comparison to another

[10] In deontic logic, the 'ought' is often modeled as box-like and the box aggregates over conjunctions in the standard modal logics.

reasoner. A comparative evaluation of rationality may state, for example, that a reasoner $\mathcal{R}$ is at least as rational as a reasoner $\mathcal{R}'$ (in the following, $\mathcal{R} \succeq \mathcal{R}'$). It seems reasonable to presuppose that comparisons of rationality should form a transitive and connex relation (i.e. a total preorder). This is the supposition that, for all reasoners $\mathcal{R}$, $\mathcal{R}'$ and $\mathcal{R}''$, if $\mathcal{R} \succeq \mathcal{R}'$ and $\mathcal{R}' \succeq \mathcal{R}''$, then $\mathcal{R} \succeq \mathcal{R}''$ and either $\mathcal{R} \succeq \mathcal{R}'$ or $\mathcal{R}' \succeq \mathcal{R}$. Most comparative evaluation methods fail in forming a total preorder of rationality because of blindspots.

A very crude comparative evaluation method could be based on set inclusion[11]:

(m1) $\mathcal{R} \succeq \mathcal{R}'$ iff $t(\mathcal{R}') \subseteq t(\mathcal{R})$ and $f(\mathcal{R}) \subseteq f(\mathcal{R}')$,

where $t(\mathcal{R})$ is the set of $\mathcal{R}$'s true beliefs and $f(\mathcal{R})$ is the set of $\mathcal{R}$'s false beliefs. Method m1 has some plausibility because false beliefs can render a reasoner irrational, but this method fails in forming a transitive and connex notion of rationality because of blindspots. According to m1, $\mathcal{R} \succeq \mathcal{R}'$ is false when $\mathcal{R}'$ holds some true belief that $\mathcal{R}$ does not hold or $\mathcal{R}$ holds some false belief that $\mathcal{R}'$ does not hold. Then $\mathcal{R} \succeq \mathcal{R}'$ and $\mathcal{R}' \succeq \mathcal{R}$ are both false when $\mathcal{R}'$ believes some truth that $\mathcal{R}$ does not believe and $\mathcal{R}$ believes some truth that $\mathcal{R}'$ does not believe. There are as many versions of $bs1$ as there are reasoners. For example, if $bs1$ is a blindspot for $\mathcal{R}$, then $bs1'$ is a blindspot for $\mathcal{R}'$, etc. Suppose that the $bs1$s are the only existing blindspots. Further, suppose that both $\mathcal{R}$ and $\mathcal{R}'$ believe all and only the truths except for those that are blindspots for them. Then $\mathcal{R}$ believes all and only the truths except for $bs1$ and $\mathcal{R}'$ believes all and only the truths except for $bs1'$. In this case, $\mathcal{R}$ should be at least as rational as $\mathcal{R}'$ and vice versa ($\mathcal{R} \succeq \mathcal{R}'$ and $\mathcal{R}' \succeq \mathcal{R}$). However, according to m1, it is false that $\mathcal{R} \succeq \mathcal{R}'$ or $\mathcal{R}' \succeq \mathcal{R}$ because $bs1 \notin t(\mathcal{R})$ and $bs1 \in t(\mathcal{R}')$, but $bs1' \in t(\mathcal{R})$ and $bs1' \notin t(\mathcal{R}')$. Method m1 generates incommensurable partial orderings of rationality, where 'rational' is ambiguous between incommensurable notions (one for each ordering).

Following Joyce (1998), epistemologists have proposed frameworks that use principles of Decision Theory to motivate norms of rationality (see Pettigrew, 2019a, for a survey). The resulting field is the Epistemic Utility Theory (EUT), which has developed the best evaluation methods available in the literature. Some frameworks in EUT follow three steps (see Fitelson and Easwaran, 2015, p. 73). The first step is to define the *ideal set of beliefs*, which is usually taken as unproblematic:

[11] An anonymous reviewer has proposed method m1, which is very crude indeed. I have introduced m1 as a dialectical tool for discussing problems of evaluation methods in general. Another anonymous reviewer has pointed out that m1 would be inadequate independently of blindspots. If the objects of beliefs are centered propositions, then a reasoner cannot believe the propositions centered in another reasoner, which results in incommensurable partial orderings of rationality independently of blindspots. This result is expected due to the nature of centered propositions and *de se* beliefs. However, I want to maintain my arguments independent of whether the objects of beliefs are sentences, uncentered, or centered propositions.

> Step 1 is straightforward. It is clear what it means for a set $\mathtt{B}$ of this type to be perfectly accurate/vindicated in a world w. The vindicated set $\mathtt{B}^*_w$ is given by: $\mathtt{B}^*_w$ contains $\mathtt{B}(p)$ ($\mathtt{D}(p)$) just in case p is true (false) at w (Fitelson and Easwaran, 2015, p. 75),

where $\mathtt{B}(p)$ means 'believe that p' (belief-value 1) and $\mathtt{D}(p)$ means 'disbelieve that p' (belief-value 0)[12]. The second step is to define a measure of epistemic disutility (inaccuracy), as the 'distance' between the set of beliefs of a reasoner and the ideal set of beliefs. An adequate measure of inaccuracy should fulfill some properties, which are often fulfilled by a Brier score (see Pettigrew, 2016, ch. 4 and my fn. 16). The third step is to use the measure of inaccuracy and some principle of Decision Theory (e.g. dominance) in an argument for some norm of rationality. Joyce (1998), for example, argues for Probabilism, i.e. the norm that rational reasoners ought to have credence functions that are coherent with the axioms of Probability (see Kolmogorov, 1950). His argument exploits the fact that if a reasoner's credence function does not fulfill Probabilism, then there is another function that is less inaccurate than hers in all situations that she finds possible (dominance)[13].

Fitelson and Easwaran (2015) and Pettigrew (2016), among others, follow these three steps. Consequently, they use the (supposedly, α-maximal) ideal set of beliefs as a standard of rationality[14], where the rationality of regular reasoners is evaluated by comparing their sets of beliefs with that standard. This procedure presupposes the existence of the ideal set of beliefs, but this may not be the case in situations of blindspots. If the ideal set of beliefs contains a belief of value 1 for all and only the truths and a belief of value 0 for all and only the falsehoods, then the value of the belief that $bs1^*$ (the proposition that the ideal set of beliefs does not contain a belief of value 1 for $bs1^*$) is not defined. This problem resembles that of n1. A 'solution' could be simply stipulating that the ideal set of beliefs has a belief of value 0 for $bs1^*$, but then regular reasoners would be evaluated incorrectly for having a true belief that $bs1^*$. Another solution would be restricting the framework to partial orderings, where an ideal set of beliefs would be a standard of rationality only for some reasoners. For example, $\mathcal{R}$ would be a standard of rationality only for reasoners who do not believe that $bs1$, $\mathcal{R}'$ only for reasoners who do not believe that $bs1'$, etc. In this case, 'rational' would be ambiguous between incommensurable notions.

[12] A similar definition is found in Pettigrew (2016, p. 3): "My proposal is that the accuracy of a credence function for a particular agent in a particular situation is given by its proximity to the credence function that is ideal or perfect or vindicated in that situation. 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. On the other hand, if a proposition is false, the ideal credence in it is the minimal credence, which is represented as 0".

[13] For another example, Leitgeb and Pettigrew (2010b, sec. 6.2) argues for Conditionalization, i.e. the norm that a rational reasoner ought to update her credences after learning some new evidence using Bayesian conditionalization (see fn. 1).

[14] The ideal set of beliefs would be treated as a standard of rationality because it would always be favorably compared to any other reasoner. Not all frameworks in EUT use the ideal set of beliefs as a standard of rationality. For two exceptions, see Leitgeb and Pettigrew (2010a) and Easwaran (2013).

Suppose that two regular reasoners r and r' are equally rational: they have the same beliefs with the exception that, while r believes that $bs1'$ (but not that $bs1$), r' believes that $bs1$ (but not that $bs1'$). Then it would be the case that $\mathcal{R} \succeq$ r and r $\succeq$ r', but not that $\mathcal{R} \succeq$ r'. The rationality of $\mathcal{R}$ and $\mathcal{R}'$ would also be incommensurable: it would not be the case that either $\mathcal{R} \succeq \mathcal{R}'$ or $\mathcal{R}' \succeq \mathcal{R}$. This problem resembles that of m1[15].

3 α-Model

In this Section, I propose a non-comparative evaluative function α, which measures the alethic value (α-value) of a set of beliefs. Function α accepts a set of beliefs $\mathtt{B}$ as input and returns a numerical value x ($\alpha(\mathtt{B}) = x$). The value of $\alpha(\mathtt{B})$ depends on the amount of truth and falsehood in the set of beliefs $\mathtt{B}$ (t and f, respectively). The computation of the values of t and f is straightforward for the case of binary truth- and belief-values (full beliefs). In this case, t and f are, respectively, the number of true and false beliefs in $\mathtt{B}$:

$$t = \sum_{\phi \in \mathtt{B}} 1 - |v(\phi) - b(\phi)| \tag{1}$$

$$f = \sum_{\phi \in \mathtt{B}} |v(\phi) - b(\phi)|, \tag{2}$$

where $v(\phi)$ returns the truth-value of ϕ (0 when ϕ is false and 1 when ϕ is true) and $b(\phi)$ returns the belief-value of ϕ (0 when the reasoner disbelieves that ϕ and 1 when the reasoner believes that ϕ). I focus on the case of binary truth and full beliefs, but equations 1 and 2 may also be used for cases involving gradational truth or belief[16].

To avoid the value of f (inaccuracy) collapsing to infinity, proponents of EUT often restrict their frameworks to finite agendas (the set of objects that can be believed), regarding which reasoners are opinionated (e.g. Leitgeb and Pettigrew, 2010a; Fitelson and Easwaran, 2015; Pettigrew, 2016). The reasoners of EUT have a fixed number of beliefs

[15] Caie (2013) attacks Joyce's argument for Probabilism using an "obvious truth" that is related to $bs1$. He argues that a rational reasoner is guaranteed to be probabilistically incoherent given that she is moderately sensitive to her credences and has high credence on that obvious truth. Fitelson and Easwaran (2015, p. 86) discusses some solutions to this problem, which depend on presuppositions about the objects of beliefs. Pettigrew (2016, p. 4) restricts his framework to situations without blindspots.

[16] In a gradational notion of truth, $v(\phi)$ has continuously many values between 0 and 1: 0 when ϕ is absolutely false, 1 when ϕ is absolutely true, and the values between 0 and 1 meaning different degrees of truth. In a gradational notion of belief (credences), $b(\phi)$ has continuously many values between 0 and 1: 0 for absolute certainty that ϕ is false, 1 for absolute certainty that ϕ is true, and the values between 0 and 1 for other degrees of certainty. Epistemologists (e.g. Pettigrew, 2016, ch. 4) sometimes argue that the measures of t and f should have the form of a Brier score for cases involving credences. If you accept these arguments, the following equations (Brier scores) could be used for measuring t and f in the α-model for the cases involving credences: $t = \sum_{\phi \in \mathtt{B}} 1 - (v(\phi) - b(\phi))^2$ and $f = \sum_{\phi \in \mathtt{B}} (v(\phi) - b(\phi))^2$.

about the same (all) items in the agenda and the evaluation methods developed in this field can only be used to compare reasoners with the same number of beliefs[17]. I do not restrict my framework to finite agendas. On the contrary, I assume that the agenda is infinite[18]. In order to avoid that t (comprehensiveness) or f collapse to infinity, I will simply require $\mathtt{B}$ to be a finite set of beliefs (see sec. 3.3). Consequently, my reasoners are not opinionated and may have a varying number of beliefs. Also, function α can be used to compare reasoners with different numbers of beliefs[19]. I do not presuppose that $\mathtt{B}$ has much structure. My reasoners can suspend judgment about some ϕ ($\phi \notin \mathtt{B}$ and $\neg\phi \notin \mathtt{B}$). They can (but they shouldn't, see sec. 4.2) believe contradictions ($b(\phi \wedge \neg\phi) = 1$) and they can have inconsistent beliefs ($b(\phi) = b(\neg\phi) = 1$). They can also fail to believe the logical consequences of their beliefs ($\phi \wedge \psi \in \mathtt{B}$ but $\phi \notin \mathtt{B}$). In general, the measures of t and f used here consider the belief-value of each belief independently of the other beliefs (e.g. the belief-value of $\phi \wedge \psi \in \mathtt{B}$ is independent of that of $\phi, \psi \in \mathtt{B}$).

3.1 α-Value

Intuitively, the α-value of a set of beliefs should be higher the greater the amount of truth in the set and lower the greater the amount of falsehood in the set. Also, if function α is to be used in a notion of α-maximality, then it should be possible for a reasoner to somehow 'maximize' her α-value. In this context, function α must fulfill the following requirements, where $\alpha(t, f)$ is $\alpha(\mathtt{B})$ with t and f being computed using equations 1 and 2 or some other measure. e.g. Brier scores ($t, f \geq 0$):

(r1) $\alpha(t, f)$ strictly increases with respect to t;

(r2) $\alpha(t, f)$ strictly decreases with respect to f;

(r3) α has an upper bound.

[17] An opinionated reasoner has an opinion (a belief-value) for every member of the agenda. If the agenda is fixed, then all reasoners that are opinionated for that agenda hold credences for the same objects. Consequently, they hold the same number of beliefs. For an exception, Easwaran (2013) works with infinite agendas, but only with local measures of inaccuracy (i.e. that is, measures of the inaccuracy of individual credences). Local measures of inaccuracy could ground a notion of rationality for the holding of particular credences, but not for reasoners in general.

[18] The agenda is always maximal in this sense. If the objects of beliefs are propositions, then no assumption is needed because there are infinitely many propositions. If the objects of beliefs are sentences, then I assume that the relevant languages have enough expressive power (e.g. recursion) for expressing infinitely many truths/falsehoods (see sec. 3.2).

[19] For this reason, I need two equations: t measures the comprehensiveness of the set of beliefs and f measures its inaccuracy. In comparing reasoners who are opinionated regarding a fixed agenda (and, consequently, have the same number of beliefs), only one (either one) of these two equations is needed (e.g. EUT only uses the measure of f). This feature has to do with a difference in the interpretation of the fulfillment of the truth-goal (see sec. 4.2).

See Table 1 for functions $\alpha 1$-$\alpha 5$, which are either straightforward functions or functions that have been used in the literature. These functions are rather arbitrary, but studying their properties is a good way to understand the desirable properties of function α.

Requirement r1 models the positive part of the truth-goal: the more truth the better. When I say that function α strictly increases wrt t, I mean that if $t' > t$, then $\alpha(t', f) > \alpha(t, f)$. Therefore, this requirement states that having more truth in the set of beliefs always increases its α-value. Function $\alpha 1(t, f) = t/(t+f)$, which measures the proportion of truth in the set of beliefs, does not fulfill r1. If $f = 0$, then $\alpha 1(t, f) = 1$ for all $t > 0$. Therefore, it would be false that $\alpha(t' > t > 0, 0) > \alpha(t > 0, 0)$. Function $\alpha 2(t, f) = (t-f)/(t+f)$, which measures the difference between the proportions of truth and falsehood, also does not fulfill r1. Function $\alpha 2$ is undefined for $t = 0$ and $f = 0$ and, consequently, it is undefined whether $\alpha 2(t' > 0, 0) > \alpha 2(0, 0)$. (This last criticism also holds for $\alpha 1$.) Requirement r2 models the negative part of the truth-goal: the less falsehood the better. When I say that $\alpha(t, f)$ strictly decreases wrt f, I mean that if $f' > f$, then $\alpha(t, f') < \alpha(t, f)$. Therefore, this requirement states that having more falsehood in the set of beliefs always decreases its α-value. Functions $\alpha 1$ and $\alpha 2$ also do not fulfill r2 (e.g. they are undefined for $t = 0$ and $f = 0$). Function $\alpha 3(t, f) = t - f$ (used in Trpin and Pellert, 2019), which measures the difference between truth and falsehood in a set of beliefs, fulfills requirements r1 and r2.

It follows from r1 and r2 that function α must be defined for all values of t and f. If $\alpha(t, f)$ was undefined for a given value of t, then it would be undefined whether $\alpha(t' > t, f) > \alpha(t, f)$ and the function would not strictly increase wrt t. The same holds for f, $\alpha(t, f' > f) < \alpha(t, f)$, and function α not strictly decreasing wrt f. Then, for all values of t and f, there must exist an x such that $\alpha(t, f) = x$. This fact has an important consequence: since, for all (real) numbers x and y, either $x \geq y$ or $y \geq x$, then, for all values of t, t', f, and f', either $\alpha(t, f) \geq \alpha(t', f')$ or $\alpha(t', f') \geq \alpha(t, f)$. Consequently, function α can be used to measure and compare the α-values for all (finite) sets of beliefs[20].

Requirement r3 is related to the existence of α-maximal reasoners. An α-maximal reasoner should somehow fulfill the truth-goal 'completely', which is related to her somehow reaching the upper bound of function α. If an α-maximal reasoner must somehow reach the upper bound of α, then function α must have an upper bound. Requirement r1 entails that if t is unbounded (i.e. if there are infinitely many truths in the agenda, which I am assuming), then the upper bound of α cannot be a maximum[21]. Then the

[20] Independently of whether t and f have continuous values, it would be interesting to require that α is differentiable at every point (and, consequently, that it is a continuous function). In this case, the first derivative of α could measure some sort of 'alethic potential' (more in the last paragraph of sec. 3.3).

[21] The maximum of a function is the largest member in its image. Suppose that $\alpha(t, f) = max(\alpha)$. Requirement r1 entails that $\alpha(t+1, f) > \alpha(t, f)$ and that $\alpha(t+1, f) > max(\alpha)$, which is a contradiction. The idea is that a reasoner can always hold an extra true belief and that function α must react to it.

Function	r1	r2	r3
$\alpha 1(t,f) = t/(t+f)$	×	×	×
$\alpha 2(t,f) = (t-f)/(t+f)$	×	×	×
$\alpha 3(t,f) = t-f$	✓	✓	×
$\alpha 4(t,f) = (t+d)/(t+f+c)$	✓	✓	✓
$\alpha 5(t,f) = (t-f)/(t+f+c)$	✓	✓	✓

Tab. 1: The behavior of some functions regarding requirements r1-r3, where $c, d > 0$ are constants. The checks indicate the requirements that the functions fulfill.

upper bound of α must be a supremum (but not a maximum)[22]. Functions $\alpha 1$ and $\alpha 2$ do have an upper bound, but it is a maximum (1). The function $\alpha 3$ does not have an upper bound. In sum, functions $\alpha 1$ and $\alpha 2$ do not fulfill r1-r3; function $\alpha 3$ fulfills r1 and r2, but not r3. Requirement r3 also prevents the value of $\alpha(t,f)$ collapsing to positive infinity[23].

Functions $\alpha 1$ and $\alpha 2$ can be fixed for requirements r1-r3 by adding constants $c, d > 0$ to their numerators or denominators (see functions $\alpha 4$ and $\alpha 5$ in Table 1)[24]. Both functions $\alpha 4$ and $\alpha 5$ fulfill requirements r1-r3. For example, function $\alpha 5$ fulfills requirement r1 because if $t' > t$, then $(t'-f)/(t'+f+c) > (t-f)(t+f+c)$[25]; function $\alpha 5$ fulfills requirement r2 because if $f' > f$, then $(t-f')(t+f'+c) < (t-f)(t+f+c)$ (the proof is similar to that in fn. 25); function $\alpha 5$ fulfills requirement r3: it has a supremum that is not a maximum (1)[26]. The choice of a specific function α that yields the optimal results is a difficult and, maybe, an empirical question[27]. For now, I prefer function $\alpha 5$,

[22] The supremum of a function α is the least upper bound of the image of α, defined as a quantity s such that no member in the image of α exceeds s, but if ϵ is any positive quantity, then there is a member in the image of α that exceeds $s - \epsilon$. All maxima are suprema, but some suprema are not maxima.

[23] To prevent function α collapsing to $-\infty$, we could require it to have an infimum (but not a minimum). The infimum of a function α is the greatest lower bound of the image of α and the minimum of a function is the smallest member in its image. I do not want to discuss this requirement here because it is marginal to the investigation of the notion of α-maximality. Nevertheless, functions $\alpha 4$ and $\alpha 5$ have infima that are not minima (0 and -1 respectively). I refer to the infimum of function α in Section 4.2.

[24] Constant c defines the 'sensitivity' of the function: the smaller the c, the greater the benefit for having more truth in the set of beliefs and the greater the penalty for having more falsehood in the set of beliefs. A similar notion may be applied to d. An anonymous referee has called my attention to the fact that Laplace's rule of succession has the form of $\alpha 4$, where $d = 1$ and $c = 2$ (let's call this α_L).

[25] Suppose that $t' > t$. Multiplying both sides by $(f+c+f)$, it follows that $t'(f+c+f) > t(f+c+f)$. Distributing, it follows that $t'f + t'c + t'f > tf + tc + tf$. Adding $(-t'f - tf)$ to both sides, it follows that $t'f + t'c - tf > tf + tc - t'f$. Since $tf = ft$ and $t'f = ft'$ (commutativity), this is equivalent to $t'f + t'c - ft > tf + tc - ft'$. Adding $(tt' - ff - fc)$ to both sides, it follows that $tt' + t'f + t'c - ft - ff - fc > tt' + ft + tc - ft' - ff - fc$. Since $t't = tt'$ and $ft = tf$, this is equivalent to $t't + t'f + t'c - ft - ff - fc > tt' + tf + tc - ft' - ff - fc$ (commutativity). From distribution, it follows that $(t'-f)(t+f+c) > (t-f)(t'+f+c)$. Dividing both sides by $(t+f+c)(t'+f+c)$, it follows that $(t'-f)/(t'+f+c) > (t-f)/(t+f+c)$.

[26] The value of $(t-f)/(t+f+c)$ decreases only wrt f and strictly increases wrt t. Then $\lim_{t\to\infty}(t-0)/(t+0+c) = 1$ is the function's upper bound because the value of t dominates that of c. Since the value of $(t-f)/(t+f+c)$ strictly increases wrt t and t is unbounded, this upper bound is not a maximum.

[27] There is much work to be done in order to shrink the class of admissible functions α, but I do not think that the outcome of this will be the selection of *one* admissible function (*the* function α). For example, I

partially for aesthetic reasons: I like the fact that this function naturally indicates the relative α-values for accuracy (close to 1), inaccuracy (close to -1), and suspension of belief (close to 0, although sets of beliefs can have α-value close to 0 by balancing accuracy and inaccuracy). Function $\alpha 5$ can have different values of c. For simplicity, I will choose a c that is very close to 0. Finally, function $\alpha 5$ is such that true and false beliefs have the same weights. I cannot think of principled reasons for using different weights, but there are adequate functions α with different weightings (e.g. $\alpha 5'(t, f) = (2t - f)/(2t + f + c)$). In the following, 'function α' refers to $\alpha 5$ with c very close to 0, but the arguments should hold for all functions that fulfill r1-r3 (e.g. $\alpha 4$).

3.2 α-Maximality

In this Section, I use function α in a model of α-maximality that is satisfiable by finite reasoners and does not present problems with blindspots. An α-maximal reasoner must somehow fulfill the truth-goal 'completely', which is modeled by the upper bound of α. There are two issues with this idea. The first is that since the upper bound of α is a supremum (but not a maximum), this function does not return its upper bound for any set of beliefs. The second is that function α measures the α-value of a set of *beliefs*, but there are different notions of belief. I think that these issues are often overlooked because the discussion (especially the discussion about alethic norms) is often built upon informal notions about what it is for a reasoner to hold a belief. To avoid this problem, I will introduce a formal model of a reasoner, which enables formal definitions of different notions of belief.

Definition 1 (Reasoner). A reasoner $\mathcal{R} = \langle \mathcal{L}, \mathtt{B}, \pi \rangle$ is a triple, where $\mathcal{L}$ is a formal language, $\mathtt{B}$ is a set of sentences in $\mathcal{L}$, and π is a function $\pi : 2^{\mathcal{L}} \times 2^{\mathcal{L}} \times \mathbb{Z}^{+} \to 2^{\mathcal{L}}$.

A reasoner is composed of a formal language ($\mathcal{L}$) that models the concepts available for the reasoner, a 'belief-set' ($\mathtt{B}$) that models the reasoner's memory, and an update function (π) that models the reasoner's pattern of inference. A fact about patterns of inference is that reasoners can execute different inferences from the same premises. This fact is expressed in the model using a function π that accepts as input the current set of beliefs ($\mathtt{B}$), an 'evidential' input ($\mathtt{INPUT}$) that models the evidence available for the reasoner, and a numeric input (positive integer): $\pi(\mathtt{INPUT}, \mathtt{B}, 1)$ models inference 1 from $\mathtt{INPUT}$ and $\mathtt{B}$, $\pi(\mathtt{INPUT}, \mathtt{B}, 2)$ models inference 2, etc. This model has three components that have not appeared in our discussion: $\mathcal{L}$, π, and $\mathtt{INPUT}$. The language $\mathcal{L}$ will be used to delimit

cannot think of a principled way of specifying the 'correct' relative importance between believing the truth and not believing the falsehood, so as to distinguish, for example, $\alpha 5$ and $\alpha 5'(t, f) = (2t - f)/(2t + f + c)$.

the reasoner's agenda, which I assume to be *denumerable*[28]. The other two components, which are important in the definition of a notion of α-maximality that does not present problems with blindspots, will be discussed in the following.

The model of a reasoner can be used to define several notions of belief. For example, informally, a reasoner explicitly believes the propositions represented or 'stored' in memory. In the model, a reasoner explicitly believes the sentences in her belief-set[29]. Reasoners cannot reach the upper bound of α given their momentary sets of beliefs (explicit, implicit, or accessible beliefs, see fn. 29). We could try to avoid this problem by adding artificial conditions to α, such as, for example, 'if t is infinite and f is finite, then $\alpha(\mathtt{B}) = max(\alpha) = 1$'[30]. This would partially fix the problem, but, in addition to its *ad hoc* flavor, it is unclear what to do when t and f are both infinite. We may choose an arbitrary value (e.g. $\alpha(\mathtt{B}) = 0$), but any choice would be inadequate. Suppose that two reasoners $\mathcal{R}$ and $\mathcal{R}'$ hold infinitely many true and infinitely many false beliefs, that they hold the same false beliefs but that, for each true belief that $\mathcal{R}'$ holds, $\mathcal{R}$ holds that same belief and another (independent) true belief. The result that $\alpha(\mathtt{B}) = \alpha(\mathtt{B}')$ seems to be wrong because $\mathcal{R}$ seems to be in a better epistemic position than $\mathcal{R}'$ (m1 gets it right).

The procedural component in the model of a reasoner (the pattern of inference, π) is used for dealing with this problem. Reasoners cannot reach the upper bound of α given their momentary sets of beliefs, but they can try to always enhance their α-value, continuously approaching the upper bound of α. For example, if a reasoner with an initial belief-set $\mathtt{B}_0$ receives a sequence of inputs $\mathtt{INPUT}_1, \mathtt{INPUT}_2, \ldots, \mathtt{INPUT}_i, \ldots$, she could try to update her initial belief-set from the available evidence generating a sequence of belief-sets $\mathtt{B}_0, \mathtt{B}_1, \ldots, \mathtt{B}_i, \ldots$ such that the value of $\alpha(\mathtt{B}_i)$ approaches the upper bound of α in the limit of this sequence. Reasoners can execute different inferences from each possible pair of $\mathtt{INPUT}_i$ and $\mathtt{B}_i$, but there is the inference that a given reasoner would 'in fact' execute from each given pair of $\mathtt{INPUT}_i$ and $\mathtt{B}_i$. Let a reasoner's *reasoning sequence* be the sequence of belief-sets $\mathtt{B}_0, \mathtt{B}_1, \ldots, \mathtt{B}_i, \ldots$ that results from those inferences that the reasoner would 'in fact' execute from the available information *if her cognitive resources were sufficient* (memory and time)[31]. In this context, I propose the following notion of α-maximality:

[28] The use of sentences in this model should not be seen as a commitment to specific objects of beliefs (see fn. 11), but as *modeling* the objects of beliefs (e.g. propositions). The assumption that agendas are denumerable is substantive because sets of propositions can have higher cardinalities, but it is weaker than the restriction to finite agendas. I will talk informally of propositions as the objects of beliefs.

[29] Informally, reasoners implicitly believe the logical consequences of their explicit beliefs. In the model, reasoners implicitly believe the logical consequence of their belief-set. Informally, reasoners have the accessible belief that ϕ iff they explicitly believe that ϕ after some amount of reasoning (Konolige, 1986, p. 19). In the model, the accessible beliefs of a reasoner are the sentences in $\pi(\mathtt{INPUT}, \mathtt{B}, i)$ for some i.

[30] In this case, the supremum of α would be a maximum. However, with some mathematical juggling, one could argue that the new function fulfills r1-r3 (e.g. that "$\infty + 1 = \infty$", or something similar).

[31] Function π determines a sequence $\mathtt{B}_0, \mathtt{B}_1, \ldots, \mathtt{B}_i, \ldots$, where $\mathtt{B}_0 = \mathtt{B}$ is the reasoner's initial belief-set and $\mathtt{B}_{i+1} = \pi(\mathtt{INPUT}_{i+1}, \mathtt{B}_i, i+1)$. This sequence could be used to represent the reasoner's reasoning sequence,

Definition 2 (α-maximality). A reasoner $\mathcal{R}$ with reasoning sequence $\mathtt{B}_0, \mathtt{B}_1, \ldots, \mathtt{B}_i, \ldots$ is α-maximal iff $\lim_{i\to\infty} \alpha(\mathtt{B}_i) = sup(\alpha)$.

An α-maximal reasoner is a reasoner who approaches the upper bound of function α in the limit of her reasoning sequence. Broadly speaking, function α in this notion of α-maximality can be seen as measuring the α-value of the (explicit) beliefs that a reasoner adopts at some stage of her reasoning sequence and are not withdrawn subsequently. Let those be *stable beliefs* ('stable' in the sense that, once adopted, these beliefs are not withdrawn)[32]. The notion of stable belief is related to that of psychological certainty, i.e. the property of a belief that is incorrigible (e.g. even in the face of additional reasoning or evidence)[33]. In this case, the set of stable beliefs (i.e. $\mathtt{B}_\omega$) may be interpreted as the set of all the (explicit) beliefs that a reasoner is psychologically certain of at some stage or other of her reasoning sequence. In the following, 'belief' refers to explicit beliefs and 'belief$_\omega$' to stable beliefs.

To show that definition 2 is satisfiable in the face of blindspots, I will construct a reasoner $\mathcal{R}^*$ who approaches α-maximality without believing$_\omega$ any of her blindspots. This construction relies on a sequence of `INPUT`s, which is related to a posteriori reasoning. Neither the notion of α-maximality nor the problems of blindspots are necessarily related to a posteriori reasoning. In addition, there are a priori reasoners who approach α-maximality without believing$_\omega$ their blindspots[34]. I have chosen this particular construction for two reasons. The first is that the sequence of `INPUT`s provides an ordering in which propositions are considered, which is relevant in dealing with complex blindspots. This ordering could also be provided by the numeric parameter of π, but the exposition is more natural using `INPUT`s. The second is that $\mathcal{R}^*$ is used in Section 4, where I use the α-model in alethic norms and evaluation methods. The construction of $\mathcal{R}^*$ highlights the

although nothing in the following depends on this choice. The condition 'if her cognitive resources were sufficient' is necessary (for finite reasoners) because reasoning sequences are infinite sequences.

[32] The notion of stable belief is related to that of P-stability in Leitgeb (2014, p. 140), where a belief is P-stable when it is sufficiently probable given any compatible proposition that is available for the reasoner (e.g. as evidence). There are two differences, though. The first is that P-stability is concerned with the relation between full beliefs and credences. The second is that the notion of stable beliefs considers the order in which these other propositions are considered (which is crucial in dealing with complex blindspots). In the model, the stable beliefs of a reasoner $\mathcal{R}$ with reasoning sequence $\mathtt{B}_0, \mathtt{B}_1, \ldots, \mathtt{B}_i, \ldots$ are in the set $\mathtt{B}_\omega = \bigcup_i \bigcap_{j\geq i} \mathtt{B}_j$. If $\mathtt{B}_\omega$ is infinite, $\alpha(\mathtt{B}_\omega)$ is undefined. That's why stable beliefs provide only a "very rough" interpretation of the measure $\lim_{i\to\infty} \alpha(\mathtt{B}_i)$, which is not the same as $\alpha(\mathtt{B}_\omega)$.

[33] This is the notion of psychological certainty: "A belief is *psychologically* certain when the subject who has it is supremely convinced of its truth. Certainty in this sense is similar to *incorrigibility*, which is the property a belief has of being such that the subject is incapable of giving it up" (Reed, 2011, p. 2). A reasoner is incapable of giving up her beliefs$_\omega$ in the sense that these are the beliefs that she 'in fact' does not give them up even in the face of all available evidence.

[34] For example, a reasoner who completely ignores her `INPUT`s and has reasoning sequence $\mathtt{B}_0 = \varnothing, \mathtt{B}_1 = \{\phi \vee \neg\phi\}, \mathtt{B}_2 = \{\phi \vee \neg\phi, (\phi \vee \neg\phi) \vee \neg(\phi \vee \neg\phi)\}$, etc is α-maximal without believing$_\omega$ her blindspots, but this strategy for approaching α-maximality is problematic with regard to finite reasoners (see sec. 3.3).

epistemic conditions in which a rational reasoner does not believe$_\omega$ her blindspots despite having evidence that they (or their components) are currently true.

Let $\mathcal{R}^*$ be a reasoner such that: (i) $\mathcal{R}^*$ is in a perfect epistemic situation (i.e. has access to all and only the true evidence)[35]; (ii) $\mathcal{R}^*$ forms beliefs according to the evidence; (iii) $\mathcal{R}^*$ has an adequate access to her beliefs (see Egan and Elga, 2005, p. 83); and (iv) $\mathcal{R}^*$ minimally revises her beliefs. Let $\mathcal{L}$ be composed of the sentences $\phi_1, \phi_2, \ldots, \phi_i, \ldots$ and let $\#\phi_i = i$ be the index of ϕ_i. That (i) is modeled as $\texttt{INPUT}_{i+1} = \{\phi_{i+1}\}$ if ϕ_{i+1} is true and $\texttt{INPUT}_{i+1} = \varnothing$ otherwise. That (ii) is modeled as function π being such that if $\phi \in \texttt{INPUT}_{i+1}$, then $\phi \in \texttt{B}_{i+1}$. That (iii) is modeled as $\mathcal{R}^*$ being able to decide the truth-value of beliefs when it depends only on her holding some belief or not. That (iv) is modeled as function π being such that if she can decide that $\phi \in \texttt{B}_i$ is false, then there is a $j > i$ such that $\phi \notin \texttt{B}_j$. For simplicity, I assume that $j = i + 1$, but 'minimum revision' in (iv) means that she could perform the revision as gradually as she wants (i.e. at *any* subsequent stage of the reasoning sequence). Let $\texttt{B}^*_0 = \varnothing$. Informally, $\mathcal{R}^*$ does as follows at each stage $i+1$ of her reasoning sequence: $\mathcal{R}^*$ inspects every $\phi \in \texttt{B}_i$; if she can decide that ϕ is false, then she withdraws the belief that ϕ from $\texttt{B}_{i+1}$; then $\mathcal{R}^*$ considers $\texttt{INPUT}_{i+1}$; if some $\phi \in \texttt{INPUT}_{i+1}$, then she adds ϕ to $\texttt{B}_{i+1}$.

$\mathcal{R}^*$ does not believe$_\omega$ her blindspots. Suppose that $\#bs1 = i$. Sentence $bs1$ is true in stage i and $bs1$ added to $\texttt{B}^*_i$. But $bs1$ is false in stage $i+1$ and $bs1$ is withdrawn from $\texttt{B}^*_{i+1}$. Since $bs1$ is not considered during subsequent stages, $\mathcal{R}^*$ does not believe$_\omega$ that $bs1$. The situation is similar to other blindspots. I am not presupposing that $\mathcal{R}^*$ believes$_\omega$ the 'easy' logical consequences of her beliefs$_\omega$. However, even if this was the case, she would not believe$_\omega$ that $bs2$ as a logical consequence of her believing$_\omega$ that ϕ and that ψ because she never believes$_\omega$ both that ϕ and that ψ. Suppose that $\#\phi = i$ and $\#\psi = j$. Either $i < j$ or $i > j$. If $i < j$, then ϕ is true at stage i and ϕ is added to $\texttt{B}^*_i$, but ψ is false at stage j and ψ is not added to $\texttt{B}^*_j$. If $i > j$, then ψ is true at stage j and ψ added to $\texttt{B}^*_j$; ϕ is also true at stage i and ϕ is added to $\texttt{B}^*_i$. But $\mathcal{R}^*$ can decide that ψ is false at stage $i+1$ and ψ is withdrawn from $\texttt{B}^*_{i+1}$. The situation is similar for other complex blindspots.

$\mathcal{R}^*$ approaches α-maximality. It is the case that $\lim_{i\to\infty} t = \infty$ because $\mathcal{R}^*$ is in a perfect epistemic situation and forms beliefs according to the evidence. It would be the case that $\lim_{i\to\infty} \alpha(\texttt{B}_i) \neq sup(\alpha)$ only if $\lim_{i\to\infty} f = \infty$. But the value of f is only higher than 0 when the number of blindspots with all of their components in $\texttt{B}^*_i$ is higher than 0, but, whenever this happens at stage i, one of these components is withdrawn at stage $i+1$, such that $\lim_{i\to\infty} f = 0$ (i.e. $\mathcal{R}^*$ does not believe$_\omega$ her blindspots).

[35] I am assuming that it is possible for $\mathcal{R}^*$ to form infinitely many true beliefs$_\omega$ (and, consequently, to approach α-maximality) from the available evidence in such a perfect epistemic situation. That assumption would not hold if, for example, the situation were entirely composed of (true) blindspots for $\mathcal{R}^*$. However, such would be a case of skeptical rather than of a perfect epistemic situation (see fn. 38).

3.3 Finite reasoners

Informally, a finite reasoner has cognitive limitations such as finite input (perception, etc transmits only a finite amount of information), finite memory (memory can store only a finite amount of information), and finite computational power (reasoning can execute only finitely many operations, such as performing an inferential step, in a finite time interval). In the model of a reasoner, a finite reasoner is as follows:

Definition 3 (Finite reasoner). A reasoner $\mathcal{R} = \langle \mathcal{L}, \mathtt{B}_0, \pi \rangle$ is finite iff $\mathcal{L}$ is recursively enumerable, $\mathtt{B}_0$ contains finitely many sentences, and π is a recursive function, whose inputs ($\mathtt{INPUT}_i$ and $\mathtt{B}_i$) can contain only finitely many sentences.

That $\mathcal{L}$ is recursively enumerable (r.e.) is a requirement for $\mathcal{L}$ being learned by a finite reasoner[36]. If $\mathcal{L}$ is r.e., then $\mathcal{L}$ is finitary, i.e. all sentences in $\mathcal{L}$ have a finite length (which is related to finite memory). The $\mathtt{B}_i$ being finite models finite memory. The $\mathtt{INPUT}_i$ being finite models finite incoming of information. Function π being recursive is related to finite computational power. If π is recursive, then every belief that the reasoner holds is associated with the execution of at most finitely many inferential steps, which models the reasoner being able to execute only finitely many operations in a finite time interval.

Finite reasoners can approach α-maximality without believing$_\omega$ their blindspots. For example, $\mathcal{R}^*$ is a finite reasoner if we presuppose that $\mathcal{L}$ is r.e., which I will presuppose[37]. This result is somewhat unsatisfactory because most often finite reasoners (e.g. humans) are not in perfect epistemic situations. However, a finite reasoner $\mathcal{R}$, who is in an imperfect epistemic situation (but has the features ii-iv), can approach α-maximality without believing$_\omega$ her blindspots as well if: (i') $\mathcal{R}$ is in a very good epistemic situation where a very good epistemic situation is one in which it is possible to believe$_\omega$ infinitely many truths from the evidence but at most finitely many falsehoods[38]. In this case, $\lim_{i\to\infty} t = \infty$ and $\lim_{i\to\infty} f = x$, where x is finite, which entails that $\lim_{i\to\infty} \alpha(\mathtt{B}_i) = sup(\alpha)$. Considerations about the blindspots of $\mathcal{R}^*$ also hold for $\mathcal{R}$ because those rely only on features ii-iv.

[36] Davidson (1965, p. 387) argues that a finite reasoner can only learn a language if it is constructive, in the sense of having compositional syntax and semantics. Davidson himself requires those languages to contain finitely many semantic primitives, but we only need to require that it is r.e. (see Haack, 1978).

[37] The $\mathtt{INPUT}_i$ of $\mathcal{R}^*$ are composed of at most one sentence. $\mathcal{R}^*$'s initial belief-set is empty and all her other $\mathtt{B}^*_i$, which are such that $max(|\mathtt{B}_{i+1}|) = |\mathtt{B}_i| + 1$, are also finite. $\mathcal{R}^*$'s function π is recursive because it executes only finitely many basic operations at each stage $i+1$: $\mathcal{R}^*$ reviews finitely many beliefs in $\mathtt{B}^*_i$, withdraws finitely many beliefs from $\mathtt{B}^*_{i+1}$, then considers at most one input in $\mathtt{INPUT}_{i+1}$ and adds at most one belief to $\mathtt{B}^*_{i+1}$.

[38] We could talk about a regular epistemic situation, where $\mathcal{R}$ can form infinitely many true and infinitely many false beliefs$_\omega$ from the evidence, and of a skeptical situation, where it is possible for $\mathcal{R}$ to form infinitely many false, but at most finitely many true beliefs$_\omega$ from the evidence. It is possible to approach α-maximality in a regular epistemic situation, but there is no infallible way to do so (from the evidence). It is impossible to approach α-maximality (from the evidence) in a skeptical situation.

Finite reasoners can approach α-maximality only deriving logical truths at every stage of their reasoning sequence (see fn. 34 for an example), but I think that there is something epistemically defective about this pattern of inference. Here is a sketch of an argument to be developed in another paper. Function α is such that every new true belief adds an α-value lower than the previous one (e.g. $\alpha(t+2, f) - \alpha(t+1, f) < \alpha(t+1, f) - \alpha(t, f)$). Finite reasoners would be in tension between increasing their α-value and decreasing their *alethic potential* (the first derivative of α, see fn. 20). A consequence of this tension is that they would be required to delay adopting true beliefs when they can adopt them later. Supposedly, reasoners can always delay the derivation of logical truths (and logical consequences of beliefs), but not the adoption of empirical beliefs because the relevant $\mathtt{INPUT}_i$ may not be available later. In this case, finite reasoners would be required to delay indefinitely the derivation of logical truths (and logical consequences of beliefs) unless there is non-alethic (e.g. practical) value attached to it. When developed, this would be a veritistic argument for the principle of clutter avoidance (Harman, 1986, p. 12). See Section 4.1 and footnote 4 for more about memory cluttering and finite reasoners.

4 Discussion

Function α may be used to generate norms and evaluation methods that can be applied to finite reasoners and that do not exhibit the problems discussed so far.

4.1 α and norms

In Section 2.1, I have surveyed several alethic norms, which exhibit problems with blindspots and, consequently, are unsatisfactory for our purposes of defining the notion of α-maximality. While the 'may' norms underdetermine this notion, the 'ought' norms yield contradictions when confronted with blindspots. I will focus on 'ought' norms because the 'may' norms inevitably underdetermine the notion of α-maximality. Norms n1-n3 exhibit the general problem of blindspots. This problem is avoided by norms n4-n5 (n4'-n5'), but these norms still have the problem of complex blindspots (see fn. 40). The 'ought' norms also have the *problem of finite reasoning*: they require those reasoners to hold infinitely many (explicit?) beliefs, which they cannot do. The approaching of the upper bound of function α in the limit of a reasoning sequence (i.e. $\lim_{i\to\infty} \alpha(\mathtt{B}_i) = sup(\alpha)$) is a property of (the reasoning sequence of) a reasoner, which is not reducible to changes in the momentary sets of beliefs of that reasoner. Consequently, the epistemic value of having a pattern of inference such that $\lim_{i\to\infty} \alpha(\mathtt{B}_i) = sup(\alpha)$ does not directly ground norms regarding the holding or not of specific beliefs. Nevertheless, the use of function α provides some insights into this normative project. For example, consider the following norms:

(n8) $\mathcal{R}$ ought to believe that ϕ iff $\alpha(\mathtt{B}+\phi) > \alpha(\mathtt{B}-\phi)$,

(n9) $\mathcal{R}$ ought to believe$_\omega$ that ϕ iff $\alpha(\mathtt{B}+\phi) > \alpha(\mathtt{B}-\phi)$.

where $\mathtt{B}+\phi$ means $\mathtt{B}\cup\{\phi\}$, and $\mathtt{B}-\phi$ means $\mathtt{B}\smallsetminus\{\phi\}$.

The left-right direction of n8 states that a reasoner ought to hold a belief only if it increases the α-value of her set of beliefs. Since the value of α strictly increases wrt t and strictly decreases wrt f, this direction states that a reasoner only ought to believe propositions that are true when believed by her. The right-left direction of n8 exhibits the problem of finite reasoning because there are infinitely many ϕ such that $\alpha(\mathtt{B}+\phi) > \alpha(\mathtt{B}-\phi)$. This problem is avoided by n9. The left-right direction of n9 states that a reasoner ought to be certain only of propositions that increase the α-value of her set of beliefs. The right-left direction of n9 avoids the problem of finite reasoning because a finite reasoner can hold infinitely many beliefs$_\omega$ (e.g. $\mathcal{R}^*$). However, n9 has a residual problem of finite reasoning because there are infinitely many *uninteresting* truths such that $\alpha(\mathtt{B}+\phi) > \alpha(\mathtt{B}-\phi)$. This is the *problem of memory cluttering* (see Harman, 1986, p. 12 and my fn, 4).

Norm n9 does not exhibit the general problem of blindspots. The blindspot $bs1$, for example, is true iff $\mathcal{R}$ does not believe that $bs1$, which means that $bs1$ is true iff $bs1 \notin \mathtt{B}$. Therefore, it is false that $\alpha(\mathtt{B}+bs1) > \alpha(\mathtt{B}-bs1)$ and $\mathcal{R}$ does not ought to believe that $bs1$, independently of whether $\mathcal{R}$ currently believes that $bs1$ or not. The same holds for any other blindspot. Norm n9 also does not have the problem of complex blindspots: for example, a reasoner does not ought to believe that $bs2$ as a logical consequence of her beliefs that ϕ and that ψ because it is not the case that she ought to believe that ϕ while believing that ψ and vice versa[39]. If $\mathcal{R}$ believes that ϕ, then ψ is false and $\mathcal{R}$ does not ought to believe that ψ (it is false that $\alpha(\mathtt{B}+\psi) > \alpha(\mathtt{B}-\psi)$). If $\mathcal{R}$ believes that ψ, then ϕ is still true, but believing that ϕ turns the belief that ψ false and $\alpha(\mathtt{B}+\phi) > \alpha(\mathtt{B}-\phi)$ is false[40]. However, norm n9 has a *residual problem of complex blindspots.* The avoidance of the problem of complex blindspots involves a norm prescribing beliefs for some (but not all) components of a complex blindspot[41]. Such a prescription should have a principled

[39] In order to satisfy some 'ought', a (finite) reasoner must adopt or withdraw beliefs (one at a time), but each change in her set of beliefs generates a new set of 'oughts' that are 'active' for her. This dynamical character of n8-n9 circumvents the problem of complex blindspots (compare with that of n3).

[40] Suppose that $\mathtt{B}$ has an amount t of truth (including ψ), an amount f of falsehood, and that the truth-value of the other beliefs remain fixed. Since $\mathtt{B}+\phi$ has the same amount $t+x-x=t$ of truth as $\mathtt{B}-\phi$ ($+x$ for ϕ, $-x$ for ψ) and an amount $f+y>f$ of falsehood larger than $\mathtt{B}-\phi$ ($+y$ for ψ), it is false that $\alpha(\mathtt{B}+\phi) > \alpha(\mathtt{B}-\phi)$. Then $\mathcal{R}$ does not ought to believe ϕ. Norms n4, n4', n5, and n5' (see fn. 9) are correct when $\mathcal{R}$ starts with ϕ, but if $\mathcal{R}$ starts with ψ, then ϕ would still be truly-believable/true were $\mathcal{R}$ to believe that ϕ.

[41] A norm cannot prescribe belief for all components of complex blindspots because it would be prescribing belief for the blindspot as an easy logical consequence of its components. The norm cannot

way for distinguishing which components are to be believed and which are not, but n9 distinguishes those cases based only on relative order (maintain what you believed first).

Norm n10 does not have any of the problems discussed so far (residual included):

(n10) $\mathcal{R}$ ought to believe$_\omega$ that ϕ iff $\phi \in \mathtt{B}^*_\omega$.

Norm n10 may be interpreted as stating that a reasoner ought to be certain of a belief at some stage of her reasoning sequence iff it is ideally warranted[42]. This norm does not have the problem of finite reasoning for the same reason as n9. Norm n10 also does not exhibit either the general problem or the problem of complex blindspots because $\mathtt{B}^*_\omega$ does not contain a blindspot or all components of a complex blindspot, although it may contain some such components (see sec. 3.2). Norm n10 partially avoids the problem of memory cluttering. There are infinitely many beliefs in $\mathtt{B}^*_\omega$, but these beliefs are not uninteresting in at least one sense: these are beliefs that are based on the total (true) evidence. Since n10 does not require reasoners to believe all the elements of $\mathtt{B}^*_\omega$ (but only to believe$_\omega$ them), in some sense, this norm only requires a reasoner to adopt beliefs when urged by the evidence. Norm n10 also does not have the residual problem of blindspots because the components of complex blindspots that are left out from $\mathtt{B}^*_\omega$ are always the self-referential ones. This is a principled way for distinguishing which components of blindspots are to be believed and which are not.

Norm n10 is arguably not purely alethic because the construction of $\mathcal{R}^*$ depends on non-alethic considerations, such as '$\mathcal{R}^*$ forms beliefs according to the evidence'. Since this is the only norm I know that avoids the problems of blindspots and finite reasoning, I am skeptical about the project of providing (purely) alethic norms because of blindspots. Carr (2020) draws similar conclusions from propositions that are closely related to blindspots. Despite relying on non-alethic considerations, n10 is still an objective norm (see fn. 7). The idealization that $\mathcal{R}^*$ is in a perfect epistemic situation is indispensable for an *alethic* norm based on her beliefs$_\omega$ because it is this idealization that provides alethic import to the norm[43]. Regular reasoners most often are not in a perfect epistemic situation

prescribe the absence of belief for all components because it would fail to prescribe belief for harmless truths.

[42] Consider Pollock's notion of ideal warrant: "Ideal warrant has to do with what a reasoner should believe if it could produce all possible relevant arguments and then survey them" (Pollock, 1995, p. 133). In some sense, $\mathcal{R}^*$ produces all the relevant conclusions because she has access to all and only the *true* evidence and forms beliefs accordingly. In this sense, $\mathtt{B}^*_\omega$ is the set of true beliefs that are warranted in a perfect epistemic situation (which is closely related to ideal warranty). The reasoner $\mathcal{R}^*$ can be seen as an 'epistemic counterpart' of a given rational regular reasoner $\mathcal{R}$. Then $\mathtt{B}^*_\omega$ can also be seen as a counterfactual expansion of $\mathcal{R}$'s pattern of inference: these are the beliefs $\mathcal{R}$ would hold if she had sufficient cognitive resources, was in a perfect epistemic situation, and fulfilled the conditions ii-iv in Section 3.2.

[43] Consider the $\mathtt{B}_\omega$ of a reasoner $\mathcal{R}$ who differs from $\mathcal{R}^*$ only because $\mathcal{R}$ is in an imperfect epistemic situation (e.g. $\mathcal{R}$ in Section 3.3). Norms defined in terms of $\mathtt{B}_\omega$ simply lack alethic import. Since we don't

and usually do not have control over this, but n10 is insensitive to the actual epistemic situation of the reasoner. Consequently, n10 is usually not deliberative, except for cases where the truth-value of propositions is transparent for the reasoners.

4.2 α and evaluations

In Section 2.2, I have surveyed two comparative evaluation methods, both of which present problems with blindspots. The main problem was that they fail in generating a notion of rationality that is a transitive and connex relation, where, for all reasoners $\mathcal{R}$, $\mathcal{R}'$ and $\mathcal{R}''$, if $\mathcal{R} \succeq \mathcal{R}'$ and $\mathcal{R}' \succeq \mathcal{R}''$, then $\mathcal{R} \succeq \mathcal{R}''$ and either $\mathcal{R} \succeq \mathcal{R}'$ or $\mathcal{R}' \succeq \mathcal{R}$. This failure results in 'rationality' being ambiguous between incommensurable notions. The α model can be used to measure the α-value in two ways. The first is to measure the α-value of momentary sets of beliefs. The comparison between the α-values of momentary sets of beliefs may, in some cases, generate deliberative evaluations of the rationality of holding or not some beliefs. The second is to measure the α-value associated with a reasoner as the limit of the α-value of her momentary sets of beliefs as she advances her reasoning sequence. With some qualifications, this second form of measurement can generate (non-comparative and comparative) methods for evaluating the rationality of reasoners that avoid the problem of incommensurable notions of rationality.

The α-value of a momentary set of beliefs should not be interpreted as a measure of the rationality of a reasoner for two reasons. The first reason is that the α-value of a momentary set of beliefs varies depending on specific functions α[44]. In this context, it is not clear whether $\alpha(\mathtt{B}) = x$ means anything in particular (a real quantity). On the other hand, comparisons between the α-values of sets of beliefs are more robust: for example, they are invariant up to positive linear transformations (if $\alpha(\mathtt{B}) > \alpha(\mathtt{B}')$, $\alpha'(t, f) = \alpha(a * t + b, c * f + d)$, and $a, b, c, d > 0$, then $\alpha'(\mathtt{B}) > \alpha'(\mathtt{B}')$). The second reason is that the α-value of a momentary set of beliefs is simply a poor indicator of a reasoner's overall rationality. For example, a reasoner with mostly false initial beliefs (low α-value) can approach α-maximality by using the available evidence for correcting her beliefs. This reasoner should be regarded as at least as rational as a reasoner who was lucky enough to have mostly true initial beliefs (high α-value), but who was incompetent in approaching α-maximality using the available (and, possibly, mostly veridical) evidence.

The comparison between the α-values of momentary sets of beliefs may, in some cases, generate deliberative evaluations of the rationality of holding some beliefs, but this only happens when the result of the comparison is transparent for the reasoners. This is the

know much about $\mathcal{R}$'s epistemic situation, any belief can be in $\mathtt{B}_\omega$ (excluding those related to blindspots and contradictions).

[44] For example, if $t = 2$ and $f = 1$, $\alpha 4(t, f) \approx .6667$ and $\alpha 5(t, f) \approx .3333$ (where c, d are close to 0). See footnote 27.

case, for example, for blindspots. A rational reasoner should not believe that $bs1$ because, for any $\mathtt{B}$, $\alpha(\mathtt{B} + bs1) < \alpha(\mathtt{B} - bs1)$. A rational reasoner should also not believe that $\neg bs1$ because, for any $\mathtt{B}$ of a *rational* reasoner, $\alpha(\mathtt{B} + \neg bs1) < \alpha(\mathtt{B} - \neg bs1)$[45]. Consequently, a rational reasoner should suspend judgment about $bs1$. Kroon (1993) argues that this is the correct prescription for this case because both $bs1$ and $\neg bs1$ fail both the evidential and the coherence conditions for rational belief[46]. How could a rational reasoner be permitted to withhold judgment about a proposition that she is in the position to know to be true? The answer is related to my results in Sections 2.1 and 4.1 and in Carr (2020): it is not reasonable to require a reasoner to believe all true propositions because, in the case of $bs1$, she is in the position to know that believing $bs1$ or $\neg bs1$ results in decreasing her α-value.

Momentary evaluations can also be deliberative in the case of inconsistent sets of propositions. Since function α strictly decreases wrt f, a rational reasoner should never believe (they should disbelieve) a contradiction because $\alpha(\mathtt{B} + \phi \wedge \neg\phi) < \alpha(\mathtt{B} - \phi \wedge \neg\phi)$. Now, a contradiction is a unitary inconsistent set. The situation is not so straightforward for larger inconsistent sets, where the result depends on factors that usually are not transparent for reasoners[47]. This problem may be avoided, in some cases, using a different strategy. Consider a lottery composed of n tickets, where exactly one is the winner. Consider the inconsistent set composed of 'ticket 1 is not the winner', ..., 'ticket n is not the winner', and 'exactly one ticket is the winner'. Independently of the other beliefs of a reasoner, as n grows, the α-value of believing all the propositions in this inconsistent set dominates that of suspending judgment or disbelieving them because, for all t and f, $\lim_{n\to\infty} \alpha(t + n, f + 1) > \alpha(t, f) > \alpha(t + 1, f + n)$. In this case, if n is large enough, a rational reasoner should believe all the propositions in this inconsistent set[48]. However,

[45] A rational reasoner $\mathcal{R}$ should not believe that $bs1$ for the reasons stated above. The proposition $\neg bs1$ is equivalent to the proposition that $\mathcal{R}$ believes that $bs1$. Then, for a rational reasoner, $\neg bs1$ should be false and $\alpha(\mathtt{B}+\neg bs1) < \alpha(\mathtt{B}-\neg bs1)$. In general, if $bs1 \notin \mathtt{B}$ and $\neg bs1 \notin \mathtt{B}$, then $\alpha(\mathtt{B}+bs1) = \alpha(\mathtt{B}+\neg bs1) < \alpha(\mathtt{B} + bs1, \neg bs1) < \alpha(\mathtt{B})$.

[46] This is the problem of $bs1$ with the "coherence condition": "The agent doesn't accept p [$bs1$] because he recognises that if he does accept p, and continues to be a good reasoner, then he is in a position to run a modus ponens argument to not-p and thereby come to recognise unconditionally that the proposition he accepts is in fact false" (Kroon, 1993, p. 383). The other three arguments in Kroon (1993) are on page 382 and in his footnote 8.

[47] For some factors, the more true beliefs a reasoner holds, the less attractive it is for her to believe the propositions in an inconsistent set; the larger the proportion of truth in the set, the more attractive it is for her to believe the propositions in the set; etc. For example, consider the inconsistent set $\{\phi, \neg\phi\}$, function $\alpha5$ with c close to 0 and a $\mathtt{B}$ that does not contain ϕ or $\neg\phi$. If $\mathtt{B}$ is such that $t = 1$ and $f = 0$, then $\alpha(\mathtt{B} \cup \{\phi, \neg\phi\}) < \alpha(\mathtt{B} \setminus \{\phi, \neg\phi\})$; but if $t = 0$ and $f = 1$, then $\alpha(\mathtt{B} \cup \{\phi, \neg\phi\}) > \alpha(\mathtt{B} \setminus \{\phi, \neg\phi\})$.

[48] This result also supports our intuition that it is epistemically less defective to hold a large body of beliefs that turns out to be inconsistent than a small inconsistent set of beliefs (e.g. an outright contradiction). For example, we blame Frege much less for subscribing to the inconsistent system in *The Foundations of Arithmetic* than a reasoner who believes an outright contradiction. My argument here appeals, not to the greater difficulty of spotting the inconsistency, but to alethic considerations alone.

a rational reasoner should not believe the conjunction of these propositions, which is a contradiction. This result is in line with the solution to the lottery paradox proposed in Kyburg (1970, p. 56), which relies on probabilistic considerations. Similar considerations apply to the preface paradox (see Makinson, 1965), where Kyburg's probabilistic solution is not directly available. This strategy needs to be further developed in another paper.

The rationality of a reasoner should be related to how much her pattern of inference is truth-conducive. This feature, in turn, is related to how the α-value of her momentary sets of beliefs changes as she advances her reasoning sequence. In definition 2, I propose that a reasoner $\mathcal{R}$ with reasoning sequence $\mathtt{B}_0, \mathtt{B}_1, \ldots, \mathtt{B}_i, \ldots$ is α-maximal iff $\lim_{i\to\infty} \alpha(\mathtt{B}_i) = sup(\alpha)$. Then it could be thought that $\lim_{i\to\infty} \alpha(\mathtt{B}_i)$ is a measure of rationality. I do not think that this is correct. The first problem with this measure is that it is not defined for all reasoners. Consider an undecided reasoner, who starts with no beliefs and, at every other stage of her reasoning sequence, adopts and withdraws the same belief. Her value of $\lim_{i\to\infty} \alpha(\mathtt{B}_i)$ is undefined because the α-value of her beliefs oscillates as she advances her reasoning sequence. The second problem is that this measure seems to return incorrect diagnoses sometimes. Consider a detached reasoner, who starts with no beliefs and, at each stage of her reasoning sequence, adopts two new true beliefs and withdraws her oldest belief. The α-value of her beliefs grows strictly monotonically and $\lim_{i\to\infty} \alpha(\mathtt{B}_i) = sup(\alpha)$, but I do not think that this reasoner should be seen as rational.

These problems are avoided using the following measure:

$$\vec{\alpha}(\mathcal{R}) = \lim_{i\to\infty} \alpha(\mathtt{B}_i \cap \mathtt{B}_\omega) \tag{3}$$

If $\phi \in \mathtt{B}_i \cap \mathtt{B}_\omega$, then $\phi \in \mathtt{B}_{i+j} \cap \mathtt{B}_\omega$ for every $j \geq 0$. Then $\alpha(\mathtt{B}_i \cap \mathtt{B}_\omega)$ grows monotonically and function $\vec{\alpha}$ always returns a value. Consequently, function $\vec{\alpha}$ can be used to evaluate any reasoner $\mathcal{R}$. The value of $\vec{\alpha}(\mathcal{R})$ is defined when both t and f approach a finite number. In this case, the result is obtained by simply substituting these values in function α. Specific results depend on specific functions α (e.g. if both t and f approaches the same x, then $\vec{\alpha}5(\mathcal{R}) = 0$), but comparisons are more robust. The value of $\vec{\alpha}(\mathcal{R})$ is also defined when the value of t or f (but not both) approaches infinity. These results are general: if t approaches infinity while f approaches a finite x, then $\vec{\alpha}(\mathcal{R}) = sup(\alpha)$; if f approaches infinity while t approaches a finite x, then $\vec{\alpha}(\mathcal{R}) = inf(\alpha)$, the infimum of α. Regular reasoners are most often in regular epistemic situations (see fn. 38), where often both t and f approach infinity. Function $\vec{\alpha}$ is able to discriminate those cases, depending on how t and f grow relatively to each other. Specific results depend on specific functions α, but comparisons are more robust: if t and f grow at the same rate, then $\vec{\alpha}5(\mathcal{R}) = 0$; if t grows twice as fast as f, $\vec{\alpha}5(\mathcal{R}) \approx .3333$; if f grows twice as fast as t, $\vec{\alpha}5(\mathcal{R}) \approx -.3333$[49].

[49] There are cases in which, when both t and f approach infinity, function $\vec{\alpha}$ returns its upper bound.

What does $\mathtt{B}_i \cap \mathtt{B}_\omega$ mean? Should function $\vec{\alpha}$ be seen as a measure of rationality? The members of $\mathtt{B}_i \cap \mathtt{B}_\omega$ are the beliefs that the reasoner is certain of at the stage i of her reasoning sequence. In this case, the function $\vec{\alpha}$ would tie up rationality and psychological certainty in a way that is related to the reflection principle (Goldstein, 1983)[50]. In this interpretation, oscillation (as in the undecided reasoner) has the same α-value as suspension of judgment, which may be correct. However, the function $\vec{\alpha}$ falls short of being a measure of rationality because a reasoner may maximize function $\vec{\alpha}$ by mere luck. For example, a stubborn conspirator, who forms certainties ignoring the evidence and happens to be in an epistemic situation where all of his conspiracies are true, may maximize function $\vec{\alpha}$, but were he in any other situation, his ill-formed certainties would be significantly mistaken. I do not think that this reasoner should be seen as rational. This problem is avoided if we average the results of $\vec{\alpha}(\mathcal{R})$ over a large number of situations. *This* would be an interesting objective evaluation method for rationality (see fn. 7).

This form of measurement is feasible, for example, in Computational Epistemology, where epistemologists have full access to the truth-values of their agents' beliefs. Examples of investigations in this field are Douven (2013), Trpin and Pellert (2019), and Olsson (2011). In a typical investigation, computational epistemologists design a computer simulation, where artificial agents (designed to exhibit the epistemic strategy to be evaluated) update their beliefs interacting with an environment (maybe, including other agents) that is randomly generated from fixed parameters. The goal is to investigate the impact of the epistemic strategy (e.g. Bayesian conditionalization) on the alethic value of the agents' beliefs. Computational epistemologists usually average the alethic impact of an epistemic strategy over many trials in the way stated above[51], although their measurements of alethic value often amount to functions other than α or $\vec{\alpha}$. For example, Trpin and Pellert (2019) use $\alpha 3$ with t and f measured using Brier scores. These other measures are only suitable for the task because epistemologists usually investigate agents who are opinionated regarding fixed and finite (often, unitary) agendas. Functions α and $\vec{\alpha}$ are more adequate measures of alethic value for the investigation of non-opinionated

For example, consider a reasoner $\mathcal{R}$ with a reasoning sequence such that, at stage 1, the value of f is 1 and the value of t is such that $\vec{\alpha}(\mathtt{B}_i \cap \mathtt{B}_\omega) \approx sup(\alpha) - 1/2$. In the subsequent stage, the value of f increases by 1 and the value of t is such that $\vec{\alpha}(\mathtt{B}_i \cap \mathtt{B}_\omega)$ is approximately $sup(\alpha) - 1/4$, $\vec{\alpha}(\alpha) - 1/8$, etc.

[50] The reflection principle states that you should defer your beliefs to those beliefs that you expect to hold in the future. The measure in equation 3 is related to the reflection principle in the sense that all the beliefs that are relevant to your rationality are beliefs that you have in the later stages of your reasoning sequence. In some sense, you only have rational beliefs that you hold in the later stages of your reasoning sequence. There are arguments in favor of the reflection principle in EUT (e.g. Easwaran, 2013).

[51] Computational epistemologists most often consider the mean alethic value (e.g. Trpin and Pellert, 2019; Olsson, 2011), but it may be interesting to consider other statistical averages, such as the mode (e.g. the "in most trials" of Douven, 2013, see fn. 1), the median, etc.

agents, agents with varying numbers of beliefs or with infinite agendas[52].

The measure of rationality using function $\vec{\alpha}$ does not directly generate a comparative evaluation method, but an obvious way of generating such a method would be as follows:

(m2) $\mathcal{R} \succeq \mathcal{R}'$ iff $\vec{\alpha}(\mathcal{R}) \geq \vec{\alpha}(\mathcal{R}')$,

where $\vec{\alpha}(\mathcal{R})$ would be averaged as stated above. Method m2 generates a notion of rationality that is a transitive and connex relation. The measure of $\vec{\alpha}(\mathcal{R})$ always returns a real number. For all real numbers x, y, and z, if $x \geq y$ and $y \geq z$, then $x \geq z$. Consequently, for all reasoners $\mathcal{R}$, $\mathcal{R}'$ and $\mathcal{R}''$, if $\mathcal{R} \succeq \mathcal{R}'$ and $\mathcal{R}' \succeq \mathcal{R}''$, then $\mathcal{R} \succeq \mathcal{R}''$. For all real numbers x and y, either $x \geq y$ or $y \geq x$. Consequently, for all reasoners $\mathcal{R}$ and $\mathcal{R}'$, either $\vec{\alpha}(\mathcal{R}) \geq \vec{\alpha}(\mathcal{R}')$ or $\vec{\alpha}(\mathcal{R}') \geq \vec{\alpha}(\mathcal{R})$. Therefore, method m2 avoids the problem of incommensurable notions of rationality. Most methods of evaluation fail in evaluating and comparing the rationality of reasoners with infinite sets of beliefs (see Treanor, 2013, p. 580)[53]. Function $\vec{\alpha}$ and method m2 are able to evaluate and compare the rationality of reasoners with infinitely many beliefs$_\omega$. In fact, they are able to discriminate different levels of rationality for reasoners with infinitely many true or false beliefs$_\omega$ (or both).

5 Conclusions

A conclusion of Section 4.1 was that the project of providing purely alethic norms for (finite) reasoners is defective because of blindspots. Nevertheless, the fulfillment of the truth-goal is still a source of epistemic value, whose measurement can be used in evaluations of rationality. The existence of blindspots also has consequences for this project, such as, for example, the difficulty in making sense of α-maximality (e.g. the ideal set of beliefs). There are two strategies for avoiding this problem. The first strategy is to adopt a weaker notion of α-maximality, where two reasoners who disagree in doxastic states regarding blindspots can be both α-maximal (e.g. the notion in def. 2). The second strategy is to forget about α-maximality and settle for purely comparative evaluations. The frameworks in EUT that do not rely on the ideal set of beliefs (e.g. Leitgeb and Pettigrew, 2010a; Easwaran, 2013) can follow this second strategy, provided that they offer a solution to the problem of incommensurable notions of rationality, either by restricting

[52] Trpin and Pellert's measure can be used to deal with non-opinionated agents with a varying number of beliefs because function $\alpha 3$ fulfills most requirements in Section 3.1 (more specifically, r1 and r2), but their measure may collapse to $\pm\infty$ when dealing with infinite agendas because $\alpha 3$ does not fulfill r3.

[53] For example, some frameworks in EUT measure the (in)accuracy of a regular set of beliefs as its distance to the ideal set of beliefs (Fitelson and Easwaran, 2015; Pettigrew, 2016). This distance may collapse to $+\infty$ when dealing with infinite agendas. For this reason, these frameworks are often restricted to finite agendas (e.g. Leitgeb and Pettigrew, 2010a; Fitelson and Easwaran, 2015; Pettigrew, 2016), where Easwaran (2013) is an exception. Method m1, on the other hand, does not exhibit this problem.

the framework to situations without blindspots (e.g. Pettigrew, 2016), by subscribing to a specific theory about the objects of belief (e.g. Fitelson and Easwaran, 2015, p. 86), or something else.

The frameworks in EUT highlight the subjective character of reasoning by considering all situations that the reasoner finds possible (e.g. in measures of expected inaccuracy); the α-model highlights the procedural character of reasoning by considering how beliefs are adopted and withdrawn (e.g. in limiting evaluations). This difference in focus endows these approaches with different affinities. EUT seems to be better equipped to support deliberative norms from analytical arguments. The α-model, on the other hand, is more easily conjoined with procedural investigations that use measures of actual epistemic value (e.g. Computational Epistemology). Analytical arguments for deliberative norms may use synchronic or diachronic evaluations. The synchronic evaluations of EUT produce, for example, arguments for Probabilism (e.g. Joyce, 1998). The synchronic form of evaluation is hampered in the α-model by the existence of different functions α. The diachronic evaluations in EUT produce, for example, arguments in favor of Bayesian conditionalization (e.g. Leitgeb and Pettigrew, 2010b). The diachronic form of evaluation in the α-model may be used in arguments for deliberative norms, but only in specific cases, such as, for example, in the analysis of the lottery paradox in Section 4.2. This form of argumentation needs to be further investigated. There is also the possibility of using the α-model 'instantaneous' measures in arguments, as in the sketch of the argument for the principle of clutter avoidance in Section 3.3. This form of argumentation also needs to be further investigated. The limiting evaluations of the α-model are difficult to implement in analytical arguments. For example, the choice of an 'actual situation' would be rather ungrounded in an analytical argument.

The analytical strategy is much less applicable to the evaluation of complex epistemic strategies and situations. This is an example from Social Epistemology, where the goal is to investigate how the adoption of epistemic practices affects the aggregated alethic value of groups of agents: which practice is epistemically more defective, lying and bullshitting? Frankfurt (2005, p. 61) claims that "bullshit is a greater enemy of the truth than lies are", but I do not think that there are analytical arguments for either conclusion. The computational approach is more fruitful in those cases. Olsson (2011), for example, constructs computer simulations where agents update their credences from the interaction with other agents and compare the alethic values of practices averaged over randomly generated social networks. In Olsson's framework, alethic value is measured using the V-value (Goldman, 1999, sec. 3.4). This measure, which is related to t alone, is only suitable for the task because the framework only deals with agents who are opinionated regarding a fixed (unitary) agenda. The limiting evaluations of the α-model would be more adequate measures of alethic value for the investigation of non-opinionated agents, agents with a varying number of beliefs, or agents with infinite agendas. The limiting aspect of these measures can be estimated by simulating the target process with inputs of increasing sizes and by extrapolating analytically the behavior of the resulting function.

Some differences between the α-model and the frameworks in EUT result from how these frameworks interpret the fulfillment of the truth-goal. For example, the truth-goal may be interpreted as requiring every belief update to minimize (expected) inaccuracy (e.g. Leitgeb and Pettigrew, 2010b), but it also may be interpreted as demanding a pattern of inference to be part of the long-term achievement of having the most comprehensive and accurate set of beliefs about the world (i.e. maximizing actual α-value)[54]. There is another difference in this regard, which, I think, favors the α-model. The truth-goal has two parts: believe the truths (comprehensiveness) and do not believe the falsehoods (accuracy). The α-model considers both sub-goals: comprehensiveness is measured by t, accuracy is measured by f, their results being integrated by the function α. The Brier score alone, as used in EUT, is a measure of (in)accuracy, but not comprehensiveness; the V-value is a measure of comprehensiveness, but not of accuracy. These sub-goals are redundant when we are dealing with reasoners who are opinionated regarding a fixed agenda (and have the same number of beliefs), but this is not the case when, for example, we are dealing with agents who explore the environment and form beliefs about new propositions. As Easwaran (2013, p. 122) puts it: "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". The α-model provides a way of dealing precisely with these cases.

References

Boghossian, Paul (2003). "The Normativity of Content". In: *Philosophical Issues* 13.1, pp. 31–45.

Bykvist, Krister and Anandi Hattiangadi (2007). "Does Thought Imply Ought?" In: *Analysis* 67.296, pp. 277–285.

— (2013). "Belief, Truth, and Blindspots". In: *The Aim of Belief*. Ed. by Timothy Chan. Oxford: Oxford University Press. Chap. 6.

Caie, Michael (2013). "Rational Probabilistic Incoherence". In: *Philosophical Review* 122.4, pp. 527–575.

Carr, Jennifer (2020). *Should You Believe the Truth?* URL: `http://philosophyfaculty.ucsd.edu/~j2carr/research.html`.

Dantas, Danilo (2017). "No Rationality Through Brute-Force". In: *Philosophy South (Filosofia Unisinos)* 18.3, pp. 195–200.

[54] Douven (2013) exploits these different interpretations (see fn. 1).

Davidson, Donald (1965). “Theories of Meaning and Learnable Languages”. In: *Proceedings of the International Congress for Logic, Methodology, and Philosophy of Science*. Ed. by Yehoshua Bar-Hillel. Amsterdam: North-Holland, pp. 3–17.

Douven, Igor (2013). “Inference to the Best Explanation, Dutch Books, and Inaccuracy Minimisation”. In: *The Philosophical Quarterly* 63.252, pp. 428–444.

Easwaran, Kenny (2013). “Expected Accuracy Supports Conditionalization - and Conglomerability and Reflection”. In: *Philosophy of Science* 80.1, pp. 119–142.

Egan, Andy and Adam Elga (2005). “I Can’t Believe I’m Stupid”. In: *Philosophical Perspectives* 19.1, pp. 77–93.

Fitelson, Branden and Kenny Easwaran (2015). “Accuracy, Coherence and Evidence”. In: *Oxford Studies in Epistemology, Volume 5*. Ed. by Tamar Gendler and John Hathorne. Oxford: Oxford University Press.

Frankfurt, Harry (2005). *On Bullshit*. Princeton: Princeton University Press.

Goldman, Alvin (1999). *Knowledge in a Social World*. Oxford: Oxford University Press.

Goldstein, Michael (1983). “The Prevision of a Prevision”. In: *Journal of the American Statistical Association* 78.384, pp. 817–819.

Haack, R. J. (1978). “Davidson on Learnable Languages”. In: *Mind* 87.346, pp. 230–249.

Harman, Gilbert (1986). *Change in View: Principles of Reasoned Revision*. Cambridge: The MIT Press.

Joyce, James (1998). “A Nonpragmatic Vindication of Probabilism”. In: *Philosophy of Science* 65.4, pp. 575–603.

Kolmogorov, Andrei (1950). *Foundations of Probability*. New York: Chelsea Publishing Company.

Konolige, Kurt (1986). *A Deduction Model of Belief*. San Francisco: Morgan Kaufmann Publishers Inc.

Kroon, Frederick (1993). “Rationality and Epistemic Paradox”. In: *Synthese* 94.3, pp. 377–408.

Kyburg, Henry (1970). “Conjunctivitis”. In: *Induction, Acceptance and Rational Belief*. Amsterdam: Springer Netherlands, pp. 55–82.

Leitgeb, Hannes (2014). “The Stability Theory of Belief”. In: *The Philosophical Review* 123.2, pp. 131–171.

Leitgeb, Hannes and Richard Pettigrew (2010a). “An Objective Justification of Bayesianism I: Measuring Inaccuracy”. In: *Philosophy of Science* 77.2, pp. 201–235.

— (2010b). “An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy”. In: *Philosophy of Science* 77.2, pp. 236–272.

Makinson, David (1965). “The Paradox of the Preface”. In: *Analysis* 25, pp. 205–207.

Olsson, Erik (2011). “A Simulation Approach to Veritistic Social Epistemology”. In: *Episteme* 8.2, pp. 127–143.

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

— (2019a). “Epistemic Utility Arguments for Probabilism”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2019. Metaphysics Research Lab, Stanford University.

— (2019b). “Veritism, Epistemic Risk, and the Swamping Problem”. In: *Australasian Journal of Philosophy* 97.4, pp. 761–774.

Pollock, John (1995). *Cognitive Carpentry: a Blueprint for How to Build a Person*. Cambridge: The MIT Press.

Raleigh, Thomas (2013). "Belief Norms and Blindspots". In: *Southern Journal of Philosophy* 51.2, pp. 243–269.

Reed, Baron (2011). "Certainty". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2011. Metaphysics Research Lab, Stanford University.

Sorensen, Roy (1988). *Blindspots*. Oxford: Oxford University Press.

Steele, Katie and Orri Stefansson (2016). "Decision Theory". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2016. Metaphysics Research Lab, Stanford University.

Treanor, Nick (2013). "The Measure of Knowledge". In: *Noûs* 47.3, pp. 577–601.

Trpin, Borut and Max Pellert (2019). "Inference to the Best Explanation in Uncertain Evidential Situations". In: *British Journal for the Philosophy of Science* 70.4, pp. 977–1001.

Wedgwood, Ralph (2015). "Doxastic Correctness". In: *Aristotelian Society Supplementary Volume* 87.1, pp. 217–234.

Whiting, Daniel (2010). "Should I Believe the Truth?" In: *Dialectica* 64.2, pp. 213–224.