(This is a penultimate version. The final version is due to be published in *Theoria: A Swedish Journal of Philosophy and Psychology*. Please refer to the published version).

Belief, Credence and Statistical Evidence

DAVIDE FASSIO (ZHEJIANG UNIVERSITY)
JIE GAO (ZHEJIANG UNIVERSITY)

Abstract
According to the Rational Threshold View, a rational agent believes *p* if and only if her credence in *p* is equal to or greater than a certain threshold. One of the most serious challenges for this view is the *problem of statistical evidence*: statistical evidence is often not sufficient to make an outright belief rational, no matter how probable the target proposition is given such evidence. This indicates that rational belief is not as sensitive to statistical evidence as rational credence. The aim of this paper is twofold. First, we argue that, in addition to playing a decisive role in rationalizing outright belief, non-statistical evidence also plays a preponderant role in rationalizing credence. More precisely, when both types of evidence are present in a context, non-statistical evidence should receive a heavier weight than statistical evidence in determining rational credence. Second, based on this result, we argue that a modified version of the Rational Threshold View can avoid the problem of statistical evidence. We conclude by suggesting a possible explanation of the varying sensitivity to different types of evidence for belief and credence based on the respective aims of these attitudes.

Keywords: Belief, Credence, Statistical evidence, Rational Threshold View, Aims of credence.

1. Introduction

Philosophers frequently distinguish between two types of doxastic attitude. On the one hand, there is *outright belief* (hereafter simply "belief"). Beliefs are categorical attitudes, which one can either have or fail to have with respect to a given proposition, and which involve dispositions to take the believed content for granted in theoretical and practical reasoning and to sincerely assert it. On the other hand, there is *credence*. Credences come in degrees measuring the subject's strength of confidence in a proposition. Degrees of credence are commonly represented by probability functions. Each degree of credence is associated with a real number in the interval between 0 and 1, where cr(*p*)=1 represents absolute certainty that *p* and cr(*p*)=0 represents absolute certainty that not-

p. Many also hold that a subject's credence in a proposition *p* is rational only if it is proportional to the degree of evidential support that that subject has in favor of *p*.[1] Rational credence functions ought to conform to probability axioms, and should be updated on new evidence by using conditionalization rules.

What is the relationship, if any, between belief and credence? According to a popular view, the so-called *Rational Threshold View*, a normative relation obtains between these attitudes. This view holds that a quite high degree of rational credence reaching a certain threshold is both necessary and sufficient for rational belief. A standard formulation of the view is the following:

(RTV) There is a threshold t $(0.5<t<1)$ such that a rational agent believes that *p* if and only if $\mathrm{cr}(p) \geq t$.[2]

One motivation for (RTV) stems from the following set of considerations. First, in order to rationally believe that *p* it seems that our credence in *p* should at least be higher than the credence in not-*p*. Second, if my rational credence in *q* is higher than that in *p*, and the degree of credence in *p* is sufficient for rational belief, then intuitively it is also rational for me to believe that *q*. Third, it seems obvious that in order to rationally believe something we do not have to be absolutely certain of it. Credence 1 is a too demanding condition for belief (at least if we stick to the interpretation of credence 1 as maximal confidence).

Though quite intuitive, (RTV) is not free from problems. One of the most serious is what we may call the *problem of statistical evidence*.[3] According to standard views in epistemology, rational credence is insensitive to the distinction between different types of evidence. If statistical evidence is the only type of evidence present in a circumstance, this straightforwardly determines the degree of rational credence in a proposition—e.g., credence in a proposition *p* should match the observed frequency of *p* in a relevant reference class. However, a number of intuitive cases seem to show that rational belief is not as sensitive to statistical evidence as rational credence. In particular, the cases indicate that statistical evidence alone is often not sufficient to justify an outright belief, no matter

[1] Easwaran and Fitelson (2015), Foley (2009), Silins (2005).

[2] This view is also often called *Lockean Thesis* (e.g., Foley, 1993, Ch.4; 2009; Leitgeb, 2014; Locke, 2014; Weatherson, 2005). The present formulation is from Buchak (2014, p. 289). The Rational Threshold View should not be confused with what some philosophers call the Threshold View, the descriptive thesis according to which there is a threshold *t* such that an agent believes *p* if and only if $\mathrm{cr}(p) \geq t$. The latter is a descriptive claim about the relationship between the two doxastic attitudes, and it is often interpreted as a metaphysical reductivist claim of belief to credence.

[3] For other problems for (RTV), see, e.g., Frankish (2009, pp. 80-81), Buchak (2014), Genin (2019). This paper will not be concerned with these further problems.

how probable the target proposition is given such evidence.[4] The consequence is a straightforward violation of (RTV).

The aim of this paper is twofold. First, we argue that, besides playing a decisive role in justifying outright beliefs, non-statistical evidence also plays a preponderant role in rationalizing credence. More precisely, we provide an analogical argument to the effect that, when both types of evidence are available in the same context, statistical evidence plays a significantly reduced role in rationalizing credence compared to non-statistical evidence. This result has important consequences for how we should conceive rational credence and its relation to evidence. The standard view holds that rational credence encodes all sorts of evidential grounds indiscriminately. Furthermore, according to this view, credence should be updated on new evidence using typical methods of conditionalization which are insensitive to non-formal features of evidence. As a result, qualitative differences in evidence cannot make any difference to the result of the update. Our result shows that this traditional view is problematic. An account of the influence of evidence on rational credence must take into account the asymmetrical rationalizing role of statistical and non-statistical evidence.

Second, based on the previous result, we show that a modified version of the Rational Threshold View can avoid the problem of statistical evidence. Roughly, our suggestion is that the threshold for rational belief is on the degree of rational credence provided that such credence is at least partially grounded in non-statistical evidence. The presence of non-statistical evidence ensures a reduced weight of statistical evidence in rationalizing credence, which in turn explains the reduced sensitivity of rational belief to the latter kind of evidence.

The paper is organized as follows. In §2 we present the problem of statistical evidence in more detail. In §3 and §4 we argue respectively for the two main claims introduced above. In §5 we conclude by suggesting a possible explanation of the varying sensitivity to different types of evidence for belief and credence. Our suggestion is that in order to understand this sensitivity we should look at the specific aims and functions of these attitudes. More precisely, we suggest that while belief is strongly sensitive to non-statistical evidence because it aims at knowledge, credence has a variable sensitivity to different types of evidence because it has two aims: its primary aim is to reach a level sufficient to constitute a rational outright belief, while its secondary aim is to minimize the risk of error.

Before proceeding further, an important clarificatory remark about the scope and targets of our paper is in order. RTV has been widely discussed in formal epistemology as a promising way of accommodating outright beliefs within a Bayesian formal apparatus. However, the debate on RTV has traditionally developed within less formal domains of philosophy, such as the philosophy of

[4] A terminological note: in this paper we will use 'justified belief' and 'rational belief' as interchangeable terms referring to what a reasonable person ought to believe in their circumstances. For a similar use see Buchak (2014, pp. 287-288).

mind and traditional epistemology. Philosophical theorizing about the RTV presupposes a gradable notion of rational credence, but doesn't rely on the presupposition of a probabilistic (Bayesian) representation of this gradability. While it is quite natural to interpret degrees of rational credence as probabilities, this interpretation is neither a precondition of RTV, nor a necessary assumption in the main arguments given in its support. That said, we would like to stress that the proposals we are going to defend in this article—our account of rational credence in §3 and the modified RTV in §4—are consistent with a Bayesian framework in which degrees of credence are conceived as probability functions and are updated using standard conditionalization methods. We will come back to this point in §3.1, where we will say more on how our proposal could be framed within a Bayesian picture.

2. The problem of statistical evidence

The problem of statistical evidence for (RTV) consists in the fact that rational belief is not as sensitive to statistical evidence as rational credence. While mere statistical considerations directly determine the degree of credence that a rational subject should have, they are not sufficient alone to justify an outright belief, no matter how strong the probabilistic support of such considerations to the target proposition is. As a result, there are possible cases in which statistical evidence supports an extremely high degree of credence in a proposition *p*, but the subject is not justified to believe that *p*, and, conversely, cases in which a subject is justified to believe that *p* though the degree of rational credence is significantly lower.

Before illustrating the problem with specific cases, it is helpful to provide a rough characterization of the distinction between statistical and non-statistical evidence. Let's call *statistical evidence* all evidence grounded on merely statistical considerations, e.g., considerations about the frequency that a certain fact or event has been or will be the case. For example, that the weather forecast said that tomorrow there is a 90% chance of rain is statistical evidence that it will rain; and that there are one thousand tickets in a fair lottery is statistical evidence that my ticket will not be the winner. *Non-statistical evidence* is based on considerations that are not merely statistical. That I see bright light out of the window is evidence that it is sunny. The propositions that it is sunny and that I see the light are related in a way that is not merely statistical. A possible way of putting this connection is in counterfactual terms: the counterfactual "if it hadn't been sunny, I wouldn't have seen the light out of the window" is true if it is sunny, but the counterfactual "if I had won the lottery, there would not have been one thousand lottery tickets" is false, and this is regardless of

whether I won the lottery or not. Another way is to attribute to evidence a causal or explanatory connection with the proposition it supports: while the fact that it is sunny can explain why I see the light outside the window, that my lottery ticket is not the winner cannot explain the fact that there are one thousand lottery tickets.[5]

Martin Smith (2016, §2.3) proposes a useful test for discriminating statistical and non-statistical evidence. His idea is that only non-statistical evidence provides a specific type of support, that he calls *normic*. According to Smith "a body of evidence E normically supports a proposition P just in case the circumstance in which E is true and P is false requires more explanation than the circumstance in which E and P are both true" (2016, p. 40). If non-statistical evidence supports a certain proposition *p*, in all normal circumstances *p* is also the case, and if *p* turns out to be false there should be some explanation for why *p* is not the case. On the contrary, if mere statistical evidence supports *p*, cases in which *p* is false, even though less probable, are as normal as the ones in which *p* is true. If *p* turned out to be false despite statistical evidence to the contrary, no special explanation of this error would be needed. For example, suppose I believe that Matthew is on holiday based on the fact that he told me so and I know he is normally very trustworthy. If this belief turned out to be false then, given the evidence upon which it is based, there needs to be some explanation for why Matthew is not on holiday despite the evidence—for example, due to an emergency situation at the office, Matthew's boss asked him to push back his holidays for some days. However, if based on mere statistical evidence I believe that my lottery ticket is a loser and this turns out to be false, there would be no need of explanation for the error. The possibility that my ticket is the winner is as normal a possibility as the one in which it is a loser (though a less probable one).

A couple of further remarks are in order here. First, from a formal point of view, statistical correlations are indistinguishable from other non-statistical properties such as modal, causal and explanatory relations.[6] A consequence is that, as some authors have observed, we are not going to be able to read off the type of evidence from purely formal features of one's credal state.[7] Instead we need to look at other (modal, causal or explanatory) features of the world in order to establish whether a piece of evidence is statistical or non-statistical. Second, saying that non-statistical evidence doesn't rely on mere statistical or frequency-based considerations doesn't mean that non-statistical evidence for *p* doesn't provide probabilistic support for *p*. Obviously, that I see bright light

[5] Buchak (2014, p. 294) considers counterfactual and causal accounts of non-statistical evidence. See also Bird (2007, pp. 101-102); Nelkin (2000, p. 397). See Belkoniene (2019) for a recent explanationist approach to the problem of statistical evidence. For similar accounts in the legal domain see, for example, Di Bello (2013); Enoch et al. (2012); Thompson (1986).

[6] Here with 'formal' we roughly mean that can be defined in terms that are probabilistic or broadly logical (Douven & Williamson, 2006, p. 758). For a similar use of 'formal' see also Buchak (2014); Staffel (2016).

[7] E.g., Buchak (2014, p. 294); Spirtes et al. (1996); Staffel (2016, p. 1733).

out of the window makes it likely for me that it is sunny—or at least, assuming that I previously ignored whether it was sunny, my seeing bright light makes it more likely than before. We suggest conceiving non-statistical evidence as involving further features *in addition to* probabilistic support. In this respect, non-statistical evidence is something *more* than mere statistical evidence.[8]

The problem of statistical evidence is well illustrated by paradigmatic examples showing the lack of correlation between high degrees of rational confidence and outright rational belief. Consider the following example (Nelkin, 2000, pp. 388-389; Smith, 2010, pp. 13-14):[9]

> **(Background color)** Suppose that I have a computer program that chooses a background color for my computer screen by randomly selecting a number between one and one million each time I turn on my computer. Associated with one of the numbers is a red screen. Associated with 999,999 numbers is a blue screen. One day I turn on my computer and then go into the next room to attend to something else. In the meantime Bruce, who knows nothing about how my computer's background color is determined, wanders into the computer room and sees that the computer is displaying a blue background.

Intuitively, I ought not to outright believe that the background is blue—even though I should believe that it's overwhelmingly likely that the background is blue. But it seems perfectly rational for Bruce to believe that it is blue. However, given Bruce's perceptual evidence, the probability that the computer is displaying a blue background would be nowhere near as high as how likely the proposition is given my evidence (i.e. 99.9999%). The chance that he could be hallucinating or have been struck by colour blindness is surely higher than my risk of error.

Buchak (2014, pp. 290-291) considers pairs of scenarios in which practical stakes and other contextual factors are identical, but the subject possesses different types of evidence:[10]

> **(Blue Bus)** It is late at night and an individual's car is hit by a bus. This individual cannot identify the bus. 95% of the buses circulating in the area at the time of the accident are operated by the Blue Bus Company, 5% are operated by the Green Bus Company, and there are no buses in the vicinity except those operated by one of these two companies. Each of the other elements of the case – negligence, causation, and, especially, the fact and the extent of the injury – is either stipulated or established to a virtual certainty.

---

[8] See Buchak (2014, p. 294), Smith (2016, p. 7). Thanks to Martin Smith for encouraging us to clarify these points.

[9] For other examples, see Buchak (2014, pp. 290-295); Nelkin (2000); Smith (2010; 2016); Williamson (2000, pp. 98-99).

[10] The examples reported in the text are simplifications of the cases originally reported by Buchak (2014). For a discussion see also Staffel (2016). Buchak borrows these cases from classic examples in philosophy of laws used to outline the so-called 'problem of naked statistical evidence'.

(Green Bus) It is late at night and an individual's car is hit by a bus. This individual cannot identify the bus. There is an eyewitness who identifies the bus as belonging to the Green Bus Company (the two bus companies operate busses with distinctive shapes). It is night-time, and so her vision is not ideal: let us say that in conditions like that in which the accident took place she makes mistakes 10% of the time.[11] All of the other elements of the case remain the same.

According to Buchak, upon reflection on the evidence available in these cases, in (Blue Bus) we don't have sufficient evidence to form a rational belief about whether a Blue Bus hit the car, but in (Green Bus) we do. Intuitively, even though the chance of being mistaken given the available evidence is greater in (Green Bus) than in (Blue Bus), it is not rational to believe that a Blue Bus hit the car in (Blue Bus), while in (Green Bus) it seems rational to believe that a Green Bus hit the car.[12] However, if the rational credence of a subject in a proposition *p* should exactly match the frequency or propensity that *p* in the relevant reference class, rational credence in the proposition supported by evidence should be higher in (Blue Bus) than in (Green Bus).

Generalizing from the above examples, it is possible to build pairs of cases A and B such that in A the degree of rational credence is higher than in B, but the subject is rational to believe in B but not in A (or at least more rational to believe in B than in A). These cases constitute a direct refutation of (RTV).[13] Buchak's diagnosis of this and other similar cases is that statistical evidence can give rise to a high degree of rational credence, and maybe it can also rationalize beliefs about the chance of a proposition being true, but it is often not enough by itself to justify a belief in the target proposition. In most cases, some non-statistical evidence is necessary. She concludes that the Rational Threshold View fails because rational belief and rational credence are sensitive to different features of evidence (*ibid.*, pp. 295-296). As the above cases show, while standards of rationality for credence don't distinguish between statistical and non-statistical evidence, standards of rationality for belief give a straightforward priority to non-statistical evidence over statistical evidence.

---

[11] While we have followed other presentations of the case, it is worthwhile to dispel potential ambiguities in the sentence "in conditions like that in which the accident took place she makes mistakes 10% of the time". What this sentence means is not that the probability that the bus involved really was a Green Bus bus, given the witness testimony, is as high as 90% ($Pr(GB|E)=0.9$), but that the witness correctly identifies the color of the bus 90% of the time ($Pr(E|\sim BB) = 0.9$) and incorrectly identifies the opposite color 10% of the time ($Pr(E|BB) = 0.1$). Observe also that probabilities in (Blue Bus) and (Green Bus) range over the same reference class. This excludes approaches to the problem based on principles of direct inference distinguishing between singular and generic probabilities (e.g., Niiniluoto, 1981). Thanks to Richard Pettigrew for directing our attention to these important details about the cases.

[12] If you don't share this intuition, feel free to change the percentages in the two examples to a degree that is enough to make reasonable a belief in (Green Bus).

[13] Buchak rightly observes that this conclusion is valid also for variable threshold views, since her cases do not involve differences in stakes or other practical factors that may influence a variation of the threshold. The cases are identical under all practical respects.

Though we agree that the problem of statistical evidence undermines classical unqualified versions of the Rational Threshold View, we think that revised versions of this principle can be maintained. We also partially disagree with Buchak's conclusion that rational belief and rational credence are systematically sensitive to different features of evidence. In the next section we argue that in certain contexts also rational credence is sensitive to the difference between statistical and non-statistical evidence.

3. Confidence and statistical evidence

Our argument starts from the consideration of a range of intuitive cases. Such cases are in all relevant respects analogous to those generating the problem of statistical evidence, except that they concern credence rather than outright belief. The cases trigger the intuition that rational credence is sensitive to the difference between statistical and non-statistical evidence. More precisely, they suggest that in contexts in which there is both statistical and non-statistical evidence, the former plays a significantly reduced role in determining the degree of rational credence compared to the latter. The argument proceeds by analogy: given that differences between the two sorts of cases are unsubstantial, if from belief-cases considered in §2 it is licit to infer that rational belief is asymmetrically sensitive to the two types of evidence, it is also licit to infer a similar conclusion for rational credence from analogous credence-cases. The conclusion is that for credence, as for belief, non-statistical evidence plays a prominent role in determining which attitude it is rational to have.

In §3.1 we introduce the argument. We also briefly outline a specific account of how statistical and non-statistical evidence combine to determine rational credence, which is naturally suggested by a consideration of the intuitive cases. In §3.2 we consider an important objection to the argument and provide some replies to it.

3.1. The argument

As anticipated in the introduction, according to standard views in epistemology, rational credence that p encodes all sorts of evidential grounds in support of or against p, indiscriminately treating all types of evidence on a par. Credence probability functions update new evidence using typical methods of conditionalization which are insensitive to qualitative, non-formal features of evidence such as counterfactual, causal or explanatory properties specific to non-statistical evidence. As a result, qualitative differences in the evidence don't make any difference to the result of the update.

Reflection on a specific range of cases suggests that this standard view is incorrect.[14] When we consider circumstances in which both statistical and non-statistical evidence are available, it becomes apparent that the two types of evidence don't contribute evenly in determining rational credence. In such cases, non-statistical considerations seem to have a dominant weight in determining the degree of rational credence.

Consider the following case adapted from the Blue Bus/Green Bus examples:

(Blue Bus Updated) The scenario is initially the same as in (Blue Bus): there is statistical evidence that 95% of the buses in the area at the time of the accident are operated by the Blue Bus Company, while only 5% are operated by the Green Bus Company. However, at a later time new independent evidence becomes available: an eyewitness identifies the bus as belonging to the Green Bus Company. As in (Green Bus), the reliability of the eyewitness's vision is 90% (viz., in relevantly similar circumstances the witness incorrectly identifies the type of bus 10% of the time and correctly identifies it 90% of the times).

If we update new evidence following Bayesian conditionalization, we have the following result—where E is the new eyewitness's testimony, S is the statistical evidence, BB is the claim that the bus that caused the accident was a Blue Bus:[15]

$$\mathrm{cr}(\mathrm{BB}\,|\,\mathrm{E\&S}) = \frac{\mathrm{cr}(\mathrm{E}|\mathrm{BB\&S}) \times \mathrm{cr}(\mathrm{BB}|\mathrm{S})}{\mathrm{cr}(\mathrm{E}|\mathrm{BB\&S}) \times \mathrm{cr}(\mathrm{BB}|\mathrm{S}) + \mathrm{cr}(\mathrm{E}|\sim\mathrm{BB\&S}) \times \mathrm{cr}(\sim\mathrm{BB}|\mathrm{S})} = \frac{0.1 \times 0.95}{0.1 \times 0.95 + 0.9 \times 0.05} = \frac{0.095}{0.14} \approx 0.68$$

According to this result, the rational credence that the bus that caused the accident was a Blue Bus given the new evidence is approximately 0.68. Thus our overall rational confidence should quite neatly favor BB over GB. However, this result seems very counterintuitive. Rather, when presented with the above case, people tend to react with the same sort of intuition as in the original Bus cases: it seems much more rational to trust the eyewitness's testimony than mere statistical considerations.[16] Consider a court case trying to establish which bus company caused the accident.

[14] It is important to stress here that our aim in this paper is not to provide an argument against Bayesianism and its methods of belief updating. We will say more on this below.

[15] Buchak (2014, pp. 294-295 and fn. 23) and Smith (2016, §4.3) discuss the same case and use the same updating method. An assumption in the calculation, also implicit in Buchak and Smith's presentations, is that the agent knows and assigns maximal credence to the proposition that the statistical evidence is true. This implies that cr(E|BB&S) = cr(E|BB) and cr(E|~BB&S) = cr(E|~BB).

[16] Such intuitive judgments have been repeatedly confirmed by experimental studies. See for example Kahneman & Tversky (1972); Lyon & Slovic (1976). While these psychologists have interpreted such data as symptomatic of a base rate fallacy, many other psychologists and philosophers disagree on such interpretation of the data. See, for example, Cohen (1981); Koehler (1996); Levi (1981). See §3.2 for considerations against identifying such cases with cases of base rate fallacy.

It seems clear that, if the court is in a situation of forced decision or must provide their own opinion about which company should pay for the car damage, it will trust the eyewitness more than statistical considerations. Consider then other dispositions and responses commonly associated with confidence, such as suspicion, inclination and conjecture: in (Blue Bus Updated) intuitively it seems perfectly appropriate to suspect that the bus that caused the accident was a Green Bus, and to be inclined to conjecture accordingly. All these considerations indicate that, intuitively, in (Blue Bus Updated) the degree of credence that it is reasonable to have in GB should be higher than in BB, even though the probabilistic support provided by statistical evidence is stronger than that provided by non-statistical evidence.

Consider another case:

> **(Background color-2)** Suppose that I have a computer program that chooses a background color for my computer screen by randomly selecting a number between one and one thousand each time I turn on my computer. Associated with one of the numbers is a green screen. Associated with 999 numbers is a blue screen. One day I turn on my computer and then go into the next room to attend to something else. In the meantime Bruce, who knows nothing about how the computer's background color is determined, wanders into the computer room and looks at the screen. Then he comes into my room and tells me that the computer is displaying a green background. Though Bruce is a quite reliable person, we can assume that in such situations his testimony will be unreliable at least once every 100 times (i.e., he will incorrectly report the color 1% of the time and correctly report 99% of the times).

By updating the new evidence (again using Bayesian conditionalization), we should conclude that my rational credence that the screen is blue should be much higher than that it is green. More precisely, it should be ≈ 0.9,[17] which is just slightly lower than the initial rational credence. This result, again, sounds absurd. Maybe Bruce's testimony is insufficient to rationalize an outright belief that the screen is green, but intuitively it seems clear that my rational confidence should lean more in favor of the claim that the screen is green than the contrary. In any case, after Bruce's testimony it seems at least rationally permissible to stop being very confident that the screen is blue and become almost equally unconfident in the two possibilities, to claim that I have no idea now what the color of the screen is, and to avoid conjectures about it.

[17] Where BB is that the background is blue, S is the statistical evidence, T is Bruce's testimony: $cr(BB|T\&S) = cr(T|BB\&S) \times cr(BB|S) / [cr(T|BB\&S) \times cr(BB|S) + cr(T|{\sim}BB\&S) \times cr({\sim}BB|S)] = (0.01) \times (0.999) / [(0.01) \times (0.999) + (0.99) \times (0.001)] = 0.00999 / 0.01098 \approx 0.91$

We could introduce here other examples eliciting similar intuitive judgments. Instead, we suggest a general method to construct such cases. The method is quite simple: take any of the examples in the literature used to illustrate the problem of statistical evidence and reformulate them making sure that (i) the two kinds of evidence support opposite conclusions, and (ii) both statistical and non-statistical evidence are present in the same circumstance and available to the same subject.[18] In addition to providing a general template for new cases illustrating the point, the present method also indicates important similarities between these cases and those considered in §2. Indeed, such similarities are so deep that we can conceive some of these cases as a subset of those used to illustrate the problem of statistical evidence. The main difference is that we now look at our intuitive judgments about the subject's rational confidence rather than about her rational belief. The resulting intuitions are exactly the same: non-statistical evidence seems to have a heavier weight than statistical evidence in the determination of both rational belief and rational credence.

Importantly, the differences between the two types of case do not seem substantive enough to motivate different explanations of the respective intuitive judgments. On the contrary, the sorts of intuitions triggered in the two types of case are sufficiently similar to suggest a common explanation. Of course, separate accounts are possible—for example an account relying on the nature of belief for cases considered in §2 and one appealing to ambiguity or under-specification in the formulation of cases involving mere credence.[19] However, given the substantial similarities, if a common type of explanation were available this would be by far preferable, more parsimonious, intuitive and less *ad hoc*. A common explanation would also avoid the risk of radically divorcing the conditions for rational belief and rational confidence. Such a divorce would have counterintuitive consequences such as the possibility of rationally believing p while having low reasonable confidence that p (we will come back to this and other undesirable consequences in §3.2).

Such a common explanation is available. We can apply to the present cases an explanation similar to the one that other philosophers have considered for the cases discussed in §2: like rational belief, also rational credence is sensitive to differences between statistical and non-statistical evidence. For credence, as for belief, non-statistical evidence plays a prominent role in determining which attitude it is rational to have in the relevant cases.

According to the suggested explanation, in contexts in which both statistical and non-statistical evidence are present, the latter has a relatively heavier weight compared to statistical evidence in determining the degree of credence of a rational subject. However, this doesn't exclude a rationalizing role of statistical evidence for credence. On the one hand, when only statistical

[18] Several cases in the literature already happen to satisfy these two conditions. See Gardiner (2019) for examples.

[19] For explanations appealing to a lack of specificity in the information relevant to determine base rates see Levi (1981); Colyvan et al. (2001).

evidence for a proposition is available, this evidence seems to directly determine the degree of rational credence that a subject should have in that proposition.[20] For example, if the only available evidence is that there is 60% probability that it will rain, it is rational to form credence .6 that it will rain. On the other hand, when non-statistical evidence is also available, statistical evidence can either weaken or reinforce the rational confidence of the subject. For example, that a DNA test indicates that the suspect was on the crime scene, in combination with some other non-statistical evidence (e.g., a potential motive to commit the crime) may lead to a higher degree of rational credence (and eventually to a belief) in the suspect's guilt.[21]

How do different types of evidence interact in determining rational credence? Our previous discussion suggests that the support of statistical evidence to a proposition is contextually variable. When both types of evidence are available in a context, the support provided by statistical evidence is partially defeated or screened off by that of non-statistical evidence. In these circumstances, rational credence is the result of a complex function that gives a heavier weight to the support of non-statistical evidence over that of statistical evidence. How precisely the respective support of the two types of evidence should be weighed in the determination of rational credence is a difficult question that unfortunately we cannot answer here. Importantly, this asymmetric weight function would apply only when both types of evidence for the same proposition are present. If there is only statistical evidence for *p*, this evidence will determine the degree of rational credence in *p* without that its support be screened off or diminished.[22]

Before proceeding further, we would like to add a clarification and address a potential worry. First, the clarification: while our present proposal diverges in important ways from standard accounts of rational credence in formal epistemology, our aim in this paper is not to provide an argument against Bayesianism and its methods of credence updating. On the contrary, we would like to stress that our proposal is fully consistent with the standard Bayesian formal apparatus. Nothing in our account hinders a probabilistic representation of degrees of credence. The proposal is also compatible with updating rational credence using standard conditionalization methods. The only point of disagreement with the standard Bayesian picture is that we do not think that credence updating should be uniquely sensitive to statistical features of evidence. Our suggestion is that the

[20] Cohen (1981, p. 329) makes the same point when he says that when no causal propensity is at issue, the only basis for estimating the required probability is the relative frequency. See Lyon & Slovic (1976, p. 294) and Tversky & Kahneman (1980, p. 63) for empirical studies confirming this tendency. This is also the intuition behind Lewis' Principal Principle according to which rational agents conform their credences to chances.

[21] We are here assuming that the result of the DNA test provides mere statistical evidence. We are aware that this is contentious. The reader is free to replace the example with one she or he finds more appropriate. For a defense of the claim that DNA profiling is properly characterized as statistical evidence see Ross (2019, §3).

[22] We do not exclude that evidential support can vary contextually depending on practical factors. For example, where a lot turns on whether *p* (e.g., in a court case deciding whether to condemn a person to a life sentence), it may be rational to rely less on statistical considerations than in more relaxed and ordinary contexts.

standard Bayesian machinery should be supplemented by additional constraints that take into account specific qualitative features of evidence.[23] These constraints do not concern the methods for belief updating but the specific weight attributed to different features of evidence. As explained above, the evidential input in conditionalization should be a complex factor in which statistical and non-statistical components of evidence should receive different weights in different contexts.[24] However, as long as this input is probabilistically coherent, rational credence can be updated by conditionalization in the light of new evidence, and the result will be a probability function.

The potential worry is the following.[25] In assessing confidence in the above cases we mainly focused on phenomenological and behavioral dispositions commonly associated with this type of attitude, such as introspective appraisals, suspicion, conjecture, and tendencies in forced choice. However, we didn't consider another type of disposition traditionally associated with confidence, namely, betting behaviors. If we consider our betting dispositions in the above cases, our intuitions are unclear. For instance, it is not at all clear that in (Blue Bus Updated), if one is offered a bet on which bus caused the accident, the right response would be to bet on a Green Bus. After all, from a statistical perspective, it is more probable that a Blue Bus caused the accident. Thus we would be more likely to win the gamble by betting on the latter option.

In response, we disagree that explicit betting dispositions constitute a good test to assess our (rational) degrees of confidence. The early Bayesian idea that degrees of belief can be read off a subject's betting dispositions reflects the influence of historical views such as logical positivism and behaviourism. However, more recently the thought that explicit betting behavior gives a reliable measure of one's credence has been the target of important criticisms. Some have argued that betting dispositions are heavily affected by factors other than credences, such as risk aversion, motives besides money, and the format of the bet (e.g., Earman, 1992; Weatherson, 1999; Christensen, 2001; Eriksson & Hájek, 2007). Others have pointed out that the very practice of gambling involves prudential and moral dimensions that inevitably misrepresent our actual degree of confidence (Fassio, forthcoming, Salas, 2019). Moreover, several authors have argued that betting scenarios involve features that almost inevitably affect one's doxastic states and evidence, thereby modifying one's initial epistemic position. A consequence is that the degrees of (rational) confidence that guide an agent's choices in gambling cannot reflect the ones that the agent has toward a given

[23] In a footnote, Buchak (2014, fn. 21) briefly suggests that one could try to add more to the structure of credence functions in order to implement in it qualitative features such as the difference between causation and correlation. She says that "If one wants to take these escape route, it will be an interesting upshot of the argument here that rational agents need to have much more complex credences than is ordinarily supposed". This is precisely what the implementation of our proposal in a Bayesian framework would look like.

[24] A possible way to implement the above proposal is to use Jeffrey conditionalization method for calibrating the weighted support of different types of evidence for a hypothesis. See Lehrer (1986, pp. 13-20) for a similar use.

[25] We thank an anonymous reviewer for pressing us to address this important worry.

proposition in normal contexts (Dodd, 2017; Eriksson and Rabinowicz, 2013; Fassio, forthcoming, Hacking, 1965, pp. 206-207; Salas, 2019).

In our view, betting dispositions are the expression of what Martin Smith (2016, §4.3) calls *probability estimates*. According to Smith, these are "one kind of specialised judgment that we can choose to make about a proposition, but they are not a foundational aspect of our intellectual lives, and are as separate from degrees of confidence and degrees of belief as they are from outright belief" (2016, p. 84, fn. 6). We can conceive probability estimates as beliefs about what one takes to be probable given mere statistical considerations. Our betting dispositions would then be guided by a specific kind of probability estimate, concerning what we take to be statistically probable conditional on a certain kind of bet being proposed. Betting scenarios are contexts in which we are asked to provide a statistical evaluation of the circumstances and then take a gamble. For instance, when presented with a bet on whether the bus that caused the accident was Blue or Green, we are asked to estimate the statistical distribution of the respective events occurring (conditional on that bet being proposed). In providing such estimate, our focus is on statistical features of evidence only. Non-statistical features are irrelevant in that context.[26]

3.2. An objection and replies

Someone may object that our intuitive diagnosis of the cases commits a *base rate fallacy*.[27] This fallacy is a specific error of probabilistic reasoning consisting of neglecting base rate information when estimating how probable a hypothesis is made by a body of evidence.[28] Suppose, for instance, that the incidence of a particular disease in the population is 1 in 10,000. And suppose we know that the test for it has a 1% false positive rate—i.e., it delivers a wrong prediction that the person has the disease one in 100 times. Now suppose that Kate goes to a clinic where the test is administered. She tests positive. What should be our credence that she has the disease? In similar cases most people tend to fallaciously ignore the initial statistical evidence (namely, knowledge of the incidence of the disease) and pay much more attention to the result of the test, which presumably is counterfactually

[26] In our view, there is a principled reason why our betting behavior should be guided by probability estimates rather than other doxastic attitudes sensitive to non-statistical features of evidence. As we shall argue in §5, sensitivity to non-statistical evidence (for both credence and belief) is aimed at acquiring the causal and counterfactual support necessary for knowledge. By contrast, when we want to place a bet, we do not aim at knowledge, but exclusively at reducing the risk of error. This goal is better served by probability estimates exclusively sensitive to statistical considerations. Probability estimates are also the relevant kind of attitudes we should consider when we focus on a range of probabilistic fallacies, such as the Base Rate Fallacy. We will be back to this issue in the next subsection.

[27] Thanks to Richard Pettigrew and Julien Dutant for bringing this possible objection to our attention. The example considered in the text has been suggested to us by Pettigrew. A similar case study was first considered by Hammerton (1973).

[28] See, for example, Kahneman & Tversky (1972).

related to the disease. In this case, people tend to consider it highly probable that Kate has the disease. Even though this is what people typically say in response to this example, it is usually accepted that they are committing a fallacy, and they should instead update new evidence using standard statistical methods such as Bayes conditionalization and set their credence below 1%.[29]

One problem with this objection is that it assumes that the above example is analogous to the cases considered in §3.1. More precisely, the objection assumes that because there is a counterfactual correlation between the positivity to the test and the disease, this evidence should be considered non-statistical. However, the counterfactual correlation is not to the proposition that Kate has the disease, but to the higher probability that this proposition is true (99%).[30] Once we realize this, it becomes clear that the new evidence cannot be treated as non-statistical like that in (Blue Bus Updated) and (Background color-2). Consider Smith's test for discriminating between statistical and non-statistical evidence. If non-statistical evidence supports a certain proposition p, more explanation is needed if p ends up to be false than if it ends up to be true. Normal conditions require less explanation than abnormal conditions do. We can easily see that the evidence provided by the medical test doesn't fulfill this condition: the positivity of the test just indicates that the person is more likely to have the disease, but if she ended up not having the disease, no special explanation would be needed. The case in which the person does not have the disease is as normal as one in which she has the disease—even though relatively less probable. Therefore, evidence in such paradigmatic examples of base rate fallacy is all purely statistical, and thus when the new evidence is updated the weight of the old evidence shouldn't be reduced. In this respect, such cases are importantly different from those considered in §3.1, which involve non-statistical evidence. Our account of the latter cases is compatible with predicting a base rate fallacy in the medical test example as well as in all other paradigmatic base rate fallacy cases.[31]

One may insist that the base rate fallacy does not only affect cases involving mere statistical evidence, such as the medical test example, but also cases such as (Blue Bus Updated), in which the updated evidence is non-statistical. Indeed, early psychological studies testing the fallacy used

[29] More precisely, by conditionalization (using Bayes theorem), we obtain that the probability that Kate has the disease is ≈ 0.0098 (assuming that the false negative rate is zero for the same result). It is worth mentioning that the reality of such a fallacy has been seriously challenged in recent years. See for example Koehler (1996), who challenges studies supporting the fallacy's existence from empirical, methodological and normative standpoints.

[30] Krantz (1981, p. 341) correctly observes that in this type of case both pieces of information are simple population frequencies.

[31] Incidentally, our account can also provide an explanation of the biased intuitions in base rate fallacy cases. Even though the evidence provided by the medical test is merely statistical, it is counterfactually related to a certain probabilistic distribution. The higher frequency of illness if the test is positive can be explained by this counterfactual correlation. The cause of the bias may be due to a confusion of counterfactual correlations of frequencies with counterfactual correlations of non-probabilistic facts. While the latter type of correlation is a condition for non-statistical evidence (and can be used to test whether a certain piece of evidence is non-statistical), the former correlation is not.

precisely this type of case.[32] However, there are reasons to resist equating such cases to less contentious cases of base rate fallacy. First, several psychologists recognize that cases updating on non-statistical evidence involve factors not specific to typical base rate fallacy cases, which can taint the experimental results. Koehler (1996, pp. 8-10) notes that such cases are sensitive to the relative diagnosticity of the information and the reliability of the information source, factors concerning the qualitative content or the source of information rather than its form. Sensitivity to these factors is deemed by most psychologists to be perfectly appropriate even from a normative point of view, manifesting a genuine aspect of human rationality (e.g., Barbey & Sloman, 2007; Cohen, 1981; Koehler, 1996; Stanovich & West, 2000).[33] Importantly, such factors are also indicative of explanatory and modal correlations characteristic of non-statistical evidence.

Second, psychological studies have confirmed that in cases like (Blue Bus Updated), intuitive judgments are much more robust and impervious to variations in the topic, numerical details and sequential formulations of the story told to the subjects than in other cases such as the medical test example.[34] These differences can be easily explained if we avoid trying to reduce all the cases to a common fallacy and we start recognizing the important role of qualitative evidential differences in the rationalization of belief and credence.[35]

Third, in the previous section we introduced a distinction between probability estimates and credences. Martin Smith (2016, §4.3) has argued that the base rate fallacy doesn't affect attitudes such as belief and credence, but exclusively probability estimates. The fallacy is a specific kind of error occasionally affecting our probabilistic reasoning when we estimate probabilities (e.g., statistical frequencies) of certain events. According to Smith, forming a certain credence or belief about a given proposition "is not at all the same as trying to figure out how to assign probabilities to

[32] Bar-Hillel (1980); Kahneman & Tversky (1972); Lyon & Slovic (1976). More recently, in their tests psychologists have tended to avoid cases such as (Blue Bus Updated) for several reasons, some of which are outlined below.

[33] For example, according to Kohler, "[h]igh diagnostic information should have a greater impact on predictions, beliefs and attributions than less diagnostic information" and "[r]eliable evidence should have a greater impact on judgments than less-reliable evidence" (1996, pp. 8-9).

[34] See, e.g., Lyon & Slovic (1976).

[35] A reviewer proposes an alternative explanation of these data. The difference between (Blue Bus Updated) and Kate's cases could be explained by an error theory. Assuming Buchak (2014)'s view that blame and attribution of responsibility are appropriate only when an agent has an outright belief, when presented with the (Blue Bus Updated) case people would think about who is to blame and reason as if they were asked to determine what to outright believe. For this reason, they would be more sensitive to non-statistical features of evidence in their interpretation of the case. We are not convinced by this explanation. For one thing, in the studies conducted by psychologists the questions posed to people unambiguously concern a partial doxastic attitude such as confidence. Given how the question is framed, it is objectively hard to confuse this with the further question what to outright believe. Moreover, the description of the case (poor visibility conditions, distance) clearly conveys that epistemic support is not enough to form an outright belief in the circumstance. A second worry with this explanation is that people tend to have exactly the same sort of intuitive judgments about cases in which blame is clearly not involved, such as in the Background color example. Finally, even assuming that this proposal provided an adequate explanation of the data, other things being equal, considerations of simplicity and unity would recommend avoiding an error theory postulating a systematic over-projection and questions' misinterpretation in favor of alternative explanations not positing any judgmental error on the part of the tested subjects.

propositions" (ibid., p. 90). Credences and probability estimates involve different regulation processes and respond to different rationality standards. Even though there are important differences between our and Smith's accounts of rational credence, Smith's response to the base rate fallacy objection could apply equally well to our view.

While the above considerations may be sufficient to bring home our point, we also think there is a further, more serious reason to resist attributing our intuitive judgments about cases considered in §3.1 to a systematic fallacy. This attribution would lead us to a dilemma. Either we provide completely different accounts of similar cases involving belief and credence (e.g., Blue-Green Bus in §2 and Blue Bus Updated in §3.1), or we adopt a common treatment. If we take the first horn, then we accept that qualitative differences in evidence matter for rational belief but not for rational credence. But then we are forced to accept the counterintuitive consequence that it is possible to completely divorce rational belief from rational confidence. This consequence, besides excluding the viability of any version of a Rational Threshold View, leads to many absurd results: first, it allows the possibility of rationally believing p while having a substantially low confidence that p, or even a stronger confidence that not-p.[36] Second, assuming that it is rationally permissible to sincerely assert what one reasonably believes, it would be reasonable to sincerely assert odd-sounding sentences like "Jack is the murderer but I am not confident that he is". Third, if justified belief involves a rational disposition to act on the believed proposition, then it may be perfectly rational for someone to act on a proposition that she reasonably takes to be false with a high degree of confidence.

If, on the other hand, we take the second horn of the dilemma, we accept a co-variance of rational belief and credence in the relevant cases, in which case either both these attitudes are more sensitive to non-statistical than to statistical evidence or they are not. If the latter, then we have to assume that intuitive judgments in cases considered in §2, such as (Background Color) and the pair (Blue Bus)-(Green Bus), are systematically mistaken. Such kind of intuitive judgments do not have a purely academic interest. They are pervasive in many ordinary domains and sometimes play important roles in our lives and practices. For example, under prevailing legal practice and across a wide range of jurisdictions, mere statistical evidence is deemed insufficient to base a verdict of guilt or liability.[37] Buchak (2014, §4) and Littlejohn (forthcoming) make similar considerations about the prominent role of non-statistical evidence as a basis for moral assessments of responsibility and blame. It is objectively hard to put all these patterns down to prejudices and biases. Are courtroom

[36] As Staffel observes, "[p]lausibly, an agent cannot rationally believe p unless she is very confident that p is true" (2016: p. 1722). Some philosophers divorce rational probability estimates and doxastic justification in this way, (e.g., Smith, 2010, pp. 26-27; 2016, Ch.4; Cohen, 1981; Kaplan, 1995). However these views are far from standard.

[37] This has also been widely acknowledged by many philosophers, see e.g. Blome-Tillman (2015); Colyvan et al. (2001); Enoch et al. (2012); Moss (2018, Ch.10); Redmayne (2008); Smith (2010; 2016, §2.2); Thomson (1986). See Smith (2016 §2.2, fn. 10) for a long list of real court cases in support of this claim.

decisions systematically biased when they refuse to condemn someone merely based on statistical considerations? For example, is it legitimate for a shoplifter to escape conviction on the basis of considerations about the relative frequency of honest shoppers in that area (Cohen, 1981, p. 329)? Should we believe guilty, blame and condemn a person for the fact that she belongs to a certain reference class most of whose members are guilty? While such a sort of systematic judgmental error is possible, both philosophers and psychologists consider this possibility with skepticism.[38]

The alternative is to adopt explanations of the relevant cases preserving the special sensitivity of both rational belief and rational credence to non-statistical evidence. This latter option has the advantage of explaining the cases in ways compatible with our intuitive judgments. Furthermore, as we will see in the next section, this approach allows us to avoid the problem of statistical evidence for a modified version of the Rational Threshold View.

4. Solving the problem of statistical evidence

As we have seen in §2, the problem of statistical evidence undermines traditional versions of the Rational Threshold View. The problem emerges because, while a certain degree of rational credence may be sufficient to justify a belief in some cases, the same or higher degree is not sufficient in others. An important feature that such cases have in common is that, whenever evidence is insufficient to warrant an outright belief, non-statistical evidence is absent. In contrast, whenever it seems rational to form an outright belief, non-statistical evidence is available.

Our main contention with critics of the Rational Threshold View is with the claim that rational belief and rational credence are sensitive to different features of evidence. If, as argued in §3, also rational credence is sensitive to the difference between statistical and non-statistical evidence, a necessary connection between belief and credence can be maintained. In particular, while circumstances in which only statistical evidence is available are such that rational credence can be very high but it is irrational to form an outright belief, when non-statistical evidence is also available both rational belief and rational credence are systematically more sensitive to non-statistical evidence. We may thus hypothesize that the two attitudes are tightly, structurally related in circumstances in which non-statistical evidence is available.

Our proposal presupposes an account of the influence of different types of evidence on rational credence along the lines suggested in §3.1, involving a contextually variable sensitivity to statistical

[38] For similar conclusions in the psychological domain, see Barbey & Sloman (2007); Cohen (1981); Stanovich & West (2000) and open peer commentaries to these articles in the same volumes. To our knowledge, the only philosopher who explicitly endorses the systematic error hypothesis is Papineau (forthcoming).

evidence. The hypothesis is that the threshold is on the degree of rational credence provided that such credence is at least partially grounded in non-statistical evidence:

(MRTV) There is a threshold *t* such that, if there is at least some non-statistical evidence for *p* available to a rational agent A, A believes *p* if and only if cr(*p*) ≥ *t*.[39]

(MRTV) straightforwardly avoids the problem of statistical evidence. Assume that, as argued in §3, when both statistical and non-statistical evidence are available, statistical evidence receives a diminished weight in determining overall rational credence. In (Background Color), my credence that the laptop screen is blue is not grounded in non-statistical evidence, not even partially. Therefore, according to (MRTV), this credence cannot make an outright belief rational (no matter how high this credence is). On the contrary, Bruce, who has non-statistical evidence that the background is blue, can rationally believe it provided that its degree of rational credence reaches threshold *t*. According to (MRTV), since there is some relevant non-statistical evidence available to Bruce, and his overall degree of rational credence is quite high, he is rational to believe that the screen is blue. Similar explanations apply to other examples considered in §2.

An important feature of this account is that when non-statistical evidence is available, actual rational credence can be influenced by statistical evidence only to a minor extent. Hence in such circumstances rational belief and rational credence are equally weakly sensitive to statistical evidence, and a variant of the threshold view can be maintained. However, when non-statistical evidence is absent, rational credence and rational belief take separate paths: while rational credence becomes strongly sensitive to statistical evidence, belief cannot be rationalised by any degree of merely statistically grounded credence, no matter how high.

(MRTV) is compatible with statistical evidence exercising some influence on rational belief when both types of evidence are available. This is because (MRTV) sets the threshold on the degree of rational credence provided that such credence is at least partially grounded in non-statistical evidence, and, as suggested in §3, this degree is the result of a complex function taking into consideration both types of evidence, even though weighing them differently. This is a positive

[39] Smith (2010, p. 26; 2016, Ch.4) suggests a hybrid view according to which normic support specific of non-statistical evidence is a necessary but not sufficient condition for justification. This view would be compatible with the claim that there is some evidential probability requirement upon justification. However, Smith rejects that view on the basis that it allows for violations of multiple premise closure for justification in risk-aggregating preface paradox cases. We do not consider these instances of closure failure a serious problem for our view, also because the intuitions underlying the preface paradox have been considered by many as supporting a Rational Threshold View and a problem for closure rather than the contrary (e.g., Christensen, 2004; Fantl & McGrath, 2009). An important aspect that distinguishes our account from that suggested by Smith is that, while in his mixed view epistemic justification requires two separate properties (normic support and high probability), our view requires a unique property, viz., a high degree of credence rationally supported by some non-statistical evidence. In our account, normic support is already built into rational credence partially grounded in non-statistical evidence.

feature of our account. Intuitively, statistical evidence can sometimes influence the strength of justification of an outright belief.[40] For instance, statistical evidence confirming other available non-statistical evidence may reinforce the rational confidence of the subject, and relatedly it may strengthen the degree of justification of a belief. That a DNA test indicates that the suspect was on the crime scene (statistical evidence), in combination with some other non-statistical evidence, can justify an outright belief that the suspect is guilty. We can conceive similar cases in which statistical evidence can lower one's rational confidence even though this is supported by non-statistical considerations.

Our view is also compatible with cases in which non-statistical evidence can decisively contribute to justify an outright belief or the suspension of judgment. An example of the former: DNA found at a crime scene in combination with some other non-statistical evidence can justify an outright belief that the suspect is guilty, while the mere non-statistical evidence would have been insufficient to grant this justification.[41] An example of the latter is a case in which the odds are so much against a proposition that non-statistical evidence that normally would suffice to justify a belief in that proposition is not sufficient to do that in the given circumstance. In this case, the statistical evidence is so strong that even after being partially screened off in the weight with non-statistical evidence, it can still lower the degree of rational confidence to the point of justifying suspension of judgment.

(MRTV) sets the threshold on the degree of rational credence provided that such credence is at least partially grounded in non-statistical evidence. A potential problem for our account concerns cases in which there is very strong statistical evidence in favor of p and very weak non-statistical evidence against p. Even if the non-statistical evidence is weighted more heavily than the statistical evidence, if the difference between the two is sufficiently great, the rational credence in p could in principle reach the threshold for rational belief. However, intuitively, in such cases it doesn't seem rational to believe that p.[42] We can think of three different responses to this challenge: first, one may argue that this type of circumstance is possible in principle, but so rare that it is never the case in practice. In order to have a case in which strong statistical evidence for p outweighs non-statistical

[40] Though some philosophers don't share this intuition, e.g., Cohen (1981, p. 329). See Ross (forthcoming) for a recent defense of the rationalizing role of statistical evidence.

[41] In a similar vein, Gardiner (2019) and Ross (forthcoming, pp. 6-8) have recently argued that statistical evidence such as DNA profiling can play a decisive justificatory role in deciding legal verdicts, even though it may be insufficient in isolation. Similarly, Smith (2018, §4) suggests that DNA evidence should never be sufficient, in isolation, for a criminal conviction, but that it could make a decisive difference when combined with evidence of other kinds. However, we should also observe that other authors (Di Bello, 2019, §6; Krauss, 2020; Ross, forthcoming) have argued that DNA evidence could in isolation be sufficient to meet the criminal standard of proof. The present matter is complicated by two further issues: first, as we observed in fn.21, it is contentious whether DNA tests provide mere statistical evidence or also non-statistical evidence. Second, it is a matter of dispute whether the debate on standards for criminal conviction should be related to the epistemological issue of what it is rational to believe (Ross, forthcoming).

[42] We thank an anonymous reviewer for directing our attention to this potential problem.

evidence against *p* to such an extent that it can reach the threshold for rational belief, we would need an absurdly high level of statistical support, one that we normally do not have. Second, we can avoid the problem if we conceive the weight between types of evidence as being such that the statistical support alone can never exceed the threshold for rational belief, no matter how large the disparity of evidential support between the two types of evidence. Third, we could slightly modify our proposal by assuming that it is not rational to believe *p* if there is not at least some (or preponderant) non-statistical evidence for *p*. The latter suggestion would fit well with how we have formulated (MRTV), as requiring at least some non-statistical evidence *for p* (not just some *available* non-statistical evidence).

5. The aims of credence

In §3, we have argued that rational credence is more sensitive to non-statistical evidence than to statistical evidence when both types of evidence are available. In §4, we have suggested a modified Rational Threshold View that can avoid the problem of statistical evidence by attributing a central role to non-statistical evidence in justifying outright belief. A natural question at this point is: why should belief and credence be more sensitive to non-statistical evidence than to evidence of a statistical kind? In this section we consider a possible answer to this question.

A satisfactory answer should take into account the specific nature of non-statistical evidence. This type of evidence possesses specific causal, explanatory or modal connections with the propositions it supports. But why should such connections be relevant to justify doxastic attitudes? A relatively straightforward answer has been provided for outright belief: belief aims at knowledge.[43] Considerations about the mere probability of a proposition cannot grant the kind of evidential support necessary for knowledge. Some causal, explanatory or modal connection between the evidence and the supported proposition is also required.[44] In this respect, a belief formed on the basis of mere statistical considerations is a somewhat defective belief, one that doesn't aim at the specific support required for knowledge. Since that belief is guaranteed to fail to achieve its aim, it lacks justification. The aim of belief at knowledge can explain the special relation that belief has with assertion and rational action. The non-statistical grounds necessary for belief also explain why

[43] The claim that knowledge is the aim or norm of belief has been defended by, amongst others, Adler (2002); Bird (2007); Huemer (2007); Littlejohn (2013; 2015); Peacocke (1999); Sutton (2007); Williamson (2000).
[44] For similar considerations, see e.g., Bird (2007, p. 102); Blome-Tillman (2017); Buchack (2014); Ichikawa (2014); Littlejohn (2015; forthcoming); Nelkin (2000); Smith (2016, §1.1); Staffel (2016).

this attitude can provide bases for reasonable responsibility, blame and praise attributions, while credence based on mere statistical evidence cannot.[45]

If an explanation is available for why belief should be more sensitive to non-statistical evidence, one is still missing for credence. The case of credence is more complex than that of belief, since for credence the stronger sensitivity to non-statistical evidence occurs only when both statistical and non-statistical evidence are available. When non-statistical evidence is absent, rational credence is straightforwardly determined by statistical evidence, exactly matching the frequency in the relevant reference class. How can we explain this variable sensitivity of credence to statistical evidence in different circumstances?

Our tentative suggestion is that credence has two aims or functions. The primary aim is to lead the subject to form beliefs. This can be achieved by reaching a level and quality of credence sufficient to constitute justified outright belief (and possibly knowledge). This goal motivates a qualitative selection of the information relevant for credence change, since belief requires a degree of credence meeting a certain threshold while at the same time being supported by evidence that maintains specific modal, causal or explanatory correlations with the supported proposition. This explains the asymmetrical role of statistical and non-statistical evidence in determining rational credence when both types of evidence are present. In short, credence aims at becoming a belief; belief aims at being knowledge; knowledge needs to be grounded in non-statistical evidence; therefore credence, in pursuing its primary aim, should be more sensitive to non-statistical evidence.

Sometimes, however, the primary aim of credence cannot be achieved. This happens in circumstances in which non-statistical evidence is not available. The secondary aim of credence becomes operational in such circumstances. This aim is to minimize the risk of error by providing the best possible grounds for decisions in situations in which the available information is merely statistical, and thus not qualitatively sufficient to support a justified belief. In order to satisfy its secondary aim, credence must be sensitive to statistical considerations. When only statistical evidence is available, the best way to minimize the risk of error is to form credence exactly matching the objective or epistemic probability of that proposition. This secondary function of credence would respond to specific needs of agents in situations of partial information. For example, if someone must decide whether to buy or sell certain stock market shares and all her available evidence is an estimated probability of the future shares' value based on frequency considerations, it is reasonable to rely in her decision on credence fully sensitive to the available statistical information.

[45] Buchak (2014); Frankish (2009); Littlejohn (2015; forthcoming).

Some philosophers would here object that attributing asymmetrical weights to different types of evidence might sometimes increase the likelihood of arriving at false judgments.[46] This might happen in some extreme cases in which strong statistical evidence contradicts non-statistical evidence. According to these philosophers, in such cases forming credences and beliefs in a way that strongly privileges non-statistical over statistical evidence would be irrational. We agree that in such extreme cases an asymmetrical weighting of different types of evidence might have detrimental consequences. If our hypothesis were correct, one would better avoid relying on her confidence and beliefs in matters in which it is particularly important to minimize risk—as suggested by some psychologists.[47] However, this doesn't imply that such an asymmetrical weighting of evidence is epistemically irrational and always (or even often) practically unreasonable; nor does it indicate that credence formation privileging non-statistical evidence over statistical evidence should be considered defective.

In our view, these philosophers commit the mistake of excessively focusing on the secondary aim of credence, tending to ignore the primary aim of such attitudes. They assume that the only or the most fundamental cognitive goal of credence is a certain type of risk minimizing accuracy, ignoring other fundamental cognitive and epistemic aims, most notably the aim at knowledge. For this reason, their approaches to doxastic rationality are only partially correct. While they legitimately apply to contexts in which knowledge is not an achievable goal, they are more doubtful when qualitative features of available evidence allow for hoping to achieve rational belief (and eventually knowledge). In such contexts, mere risk-minimization accuracy is no longer an appropriate means to achieve that goal. Since knowledge requires robust causal, explanatory and counterfactual connections between the evidence and the believed proposition, only evidence possessing such features would constitute an appropriate means to achieve that goal. This doesn't mean that we should sacrifice accuracy for other epistemic goals. Rather, we should recognize two different kinds of accuracy, each having its respective role in our doxastic life: a risk-minimization, quantitative accuracy, and a knowledge-centered accuracy, more sensitive to the safety and goodness of causal and explanatory grounds of the attitude.[48] Far from depicting an irrational model of our doxastic attitudes, an account of credence taking into account qualitative features of evidence would be in a position to explain most of our intuitive judgments about these attitudes,

[46] E.g., Papineau (forthcoming).

[47] For representative lists see references in Cohen (1981, p. 317) and Stanovich & West (2000, pp. 645-646).

[48] This doesn't contradict a characterization of accuracy in terms of "closeness to the truth" (e.g., Weisberg, 2015, p. 818). One may be closer to the truth in the sense of having attitudes more likely to be true; but in another, qualitative sense one is "closer" to the truth when one possesses that truth, i.e., knows it.

while at the same time preserving stable and intuitive connections between confidence, belief and knowledge.[49]

Acknowledgments. Earlier versions of this paper were presented at the 2015 Conference 'Belief, Credence and Action' in Geneva, and the 2017 BARSEA Epistemology Workshop in Cambridge (UK). Thanks to the audiences for their helpful feedback. We would like to thank in particular Martin Smith, Richard Pettigrew and two anonymous reviewers for helpful comments on earlier drafts of this paper, and Igor Douven, Julien Dutant and Bernard Salow for valuable discussions and feedbacks. The two authors have contributed equally to this work.

References

Adler, J. (2002). *Belief's Own Ethics*. MIT Press.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From statistical formats to cognitive structures. *Behavioral and Brain Sciences*, 30(3): 287–292.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, 44(3): 211–233.

Belkoniene, M. (2019). Normalcy, Understanding and the Problem of Statistical Evidence. *Theoria* 85(3): 202-218.

Bird, A. (2007). Justified judging. *Philosophy and Phenomenological Research*, *74*(1): 81–110.

Blome-Tillman, M. (2015). Sensitivity, causality and statistical evidence in courts of law. *Thought* 4: 102-112.

Blome-Tillmann, M. (2017). 'More Likely Than Not' - Knowledge First and the Role of Statistical Evidence in Courts of Law. In A. Carter, E. Gordon & B. Jarvis (eds.), *Knowledge First - Approaches in Epistemology and Mind*, pp. 278–292. Oxford: Oxford University Press.

Buchak, L. (2014). Belief, Credence, and Norms. *Philosophical Studies*, 169(2): 285–311.

Christensen, D. (2001). Preference-based arguments for probabilism. *Philosophy of Science* 68 (3): 356-376.

Christensen, D. (2004). *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford: Oxford University Press.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4(3): 317–370.

Colyvan, M., Regan, H. M., & Ferson, S. (2001). Is it a crime to belong to a reference class. *Journal of Political Philosophy*, 9(2): 168–181.

de Finetti, B. (1990). *Theory of probability* (first edition 1974, Vol. I). New York: Wiley.

Di Bello, M. (2013). *Statistics and Probability in Criminal Trials: The Good, the Bad, and the Ugly*. Stanford University.

Di Bello, M. (2019). Trial by statistics: Is high probability of guilt enough to convict? *Mind*, 128(512): 1045-1084.

Dodd, D. (2017). Belief and certainty. *Synthese*, 194(11): 4597-4621.

Douven, I., & Williamson, T. (2006). Generalizing the Lottery Paradox. *British Journal for the Philosophy of Science*, 57(4): 755–779.

[49] For recent insightful discussions and criticisms of risk minimization approaches to justification see Littlejohn (2015), Smith (2016). For views advocating a distinct role for outright belief and degrees of credence, see for example Jackson (2019), Weisberg (2013; 2020). Differently from these authors, we not only advocate different roles for belief and credence, but also different contextually-relative roles for credence.

Earman, J. (1992). B*ayes or Bust: A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.
Easwaran, K. and Fitelson, B. (2015). Accuracy, Coherence, and Evidence. In T. Gendler and J. Hawthorne (ed.), *Oxford Studies in Epistemology*, Vol. 5, pp. 61-96. Oxford: Oxford University Press.
Enoch, D., Spectre, L., & Fisher, T. (2012). Statistical Evidence, Sensitivity, and the Legal Value of Knowledge. *Philosophy and Public Affairs*, 40(3): 197–224.
Eriksson, L., & Hájek, A. (2007). What are degrees of belief? *Studia Logica*, 86(2): 185–215.
Eriksson, L., & Rabinowicz, W. (2013). The interference problem for the betting interpretation of degrees of belief. *Synthese* 190(5): 809-830.Fantl, J., & McGrath, M. (2009). *Knowledge in an Uncertain World*. Oxford: Oxford University Press.
Fassio, D. (forthcoming). Moderate Skeptical Invariantism. *Erkenntnis*.
Foley, R. (1993). *Working Without a Net: A Study of Egocentric Epistemology*. Oxford: Oxford University Press.
Foley, R. (2009). Beliefs, Degrees of Belief, and the Lockean Thesis. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of Belief*, pp.37–47. Springer.
Frankish, K. (2009). Partial Belief and Flat-out Belief. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of Belief*, pp.75–93. Springer.
Gardiner, G. (2019). Legal burdens of proof and statistical evidence. In David Coady and James Chase (Eds.), *The Routledge handbook of Applied Epistemology*. New York: Routledge.
Genin, K. (2019). Full and Partial Belief. In R. Pettigrew and J. Weisberg (eds.), *The Open Handbook of Formal Epistemology*. URL:<https://philpapers.org/archive/PETTOH-2.pdf>.
Hacking, I. (1965). *Logic of statistical inference*. Cambridge: Cambridge University Press.
Hammerton, M. (1973). A case of radical probability estimation. *Journal of Experimental Psychology*, 101(2): 252–254.
Huemer, M. (2007). Moore's Paradox and the Norm of Belief. In *Themes from G.E. Moore.*, pp.142–157. Oxford: Oxford University Press.
Ichikawa, J. J. (2014). Justification is potential knowledge. *Canadian Journal of Philosophy*, 44(2): 184–206.
Jackson, L. (2019). How Belief-Credence Dualism Explains Away Pragmatic Encroachment. *The Philosophical Quarterly*, 69(276): 511-533.
Jeffrey, R. (2004). *Subjective Probability-The Real Thing*. Cambridge: Cambridge University Press.
Kahneman, D., & Tversky, A. (1972). On prediction of judgment. *Oregon Research Institute Research Bulletin*, *12*(4).
Kaplan, M. (1995). Believing the improbable. *Philosophical Studies*, 77(1): 117–146.
Koehler, J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioural and Brain Sciences*, 19: 1–53.
Krantz, D. H. (1981). Improvements in human reasoning and an error in L. J. Cohen's. *Behavioral and Brain Sciences*, 4(3): 340.
Krauss, S. (2020). Against the alleged insufficiency of statistical evidence. *Florida State University Law Review*, 47.
Lehrer, K. (1986). The coherence theory of knowledge. *Philosophical Topics*, 14(1): 5–25.
Leitgeb, H. (2014). The Stability Theory of Belief. *Philosophical Review*, 123(2): 131–171.
Levi, I. (1981). Should Bayesians sometimes neglect base rates? *Behavioral and Brain Sciences*, 4(3): 342.
Lewis, D. (1974). Radical interpretation. *Synthese*, 27(3/4): 331–344.
Littlejohn, C. (2015). Who Cares What You Accurately Believe? *Philosophical Perspectives*, 29(1): 217-248.
Littlejohn, C. (forthcoming). Truth, Knowledge, and the Standard of Proof in Criminal Law. *Synthese*.
Locke, D. T. (2014). The Decision-Theoretic Lockean Thesis. *Inquiry*, 57(1): 28–54.
Lyon, D., & Slovic, P. (1976). Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psyehologica*, 40: 287–298.

Maher, P. (1993). *Betting on Theories*. Cambridge: Cambridge University Press.
Moss, S. (2018). *Probabilistic Knowledge*. Oxford: Oxford University Press.
Nelkin, D. K. (2000). The Lottery Paradox, Knowledge, and Rationality. *Philosophical Review*, 109(3): 373–409.
Nesson, C. (1985). The evidence or the event? On judicial proof and the acceptability of verdicts. *Harvard Law Review*, 98(7): 1357–1392.
Niiniluoto, I. (1981). L. J. Cohen versus Bayesianism. *Behavioral and Brain Sciences*, 4(3): 349–349.
Papineau, D. (forthcoming). The disvalue of knowledge. *Synthese*.
Ramsey, F. (1926). Truth and probability (Reprinted from in R. B. Braithwaite (ed), *The Foundations of Mathematics and Other Logical Essays*, 1950, pp. 156-198. New York: The Humanities Press).
Redmayne, M. (2008). Exploring the proof paradoxes. *Legal Theory*, 14: 281–309.
Reichenbach, H. (1949). *The Theory of Probability*. Berkeley: University of California Press.
Ross, L. (forthcoming). Rehabilitating Statistical Evidence. *Philosophy and Phenomenological Research*.
Salas, Javier González de Prado (2019). Extreme Betting. *Ratio*, 32 (1): 32-41.
Silins, N. (2005). Deception and Evidence. *Philosophical Perspectives* 19: 375-404.
Smith, M. (2010). What Else Justification Could Be. *Noûs*, 44(1), 10–31.
Smith, M. (2016). *Between Probability and Certainty*. Oxford: Oxford University Press.
Smith, M. (2018). When does evidence suffice for conviction? *Mind*, 127(508): 1193-1218.
Spirtes, P., Glymour, C., & Scheines, R. (1996). Causation, Prediction, and Search. *British Journal for the Philosophy of Science*, 47(1), 113–123.
Staffel, J. (2016). Beliefs, buses and lotteries: Why rational belief can't be stably high credence. *Philosophical Studies*, 173(7): 1721-1734.
Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23(5): 645–665.
Thompson, J. (1986). Liability and individualized evidence. In W. Parent (ed.), *Rights, Restitution and Risk*. Harvard University Press.
von Mises, R. (1957). *Probability, Statistic and Truth* (revised English edition). New York: Macmillan.
Weatherson, B. (1999). Begging the question and bayesianism. *Studies in History and Philosophy of Science*, 30: 687–697.
Weatherson, B. (2005). Can we do without pragmatic encroachment? *Philosophical Perspectives*, 19(1): 417–443.
Weisberg, J. (2020). Belief in Psyontology. *Philosophers' Imprint*, 20(11): 1-27.
Weisberg, J. (2013). Knowledge in Action. *Philosophers' Imprint*, 13(22): 1-23.
Weisberg, J. (2015). You've Come a Long Way, Bayesians. *Journal of Philosophical Logic*, 44: 817–834.
Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.