Logical Reasoning in Social Settings

edited by
Mario Piazza
Matteo Tesi
Pietro Vigiani

EDIZIONI
DELLA
NORMALE

Workshop
Reasoning with Imperfect Information
in Social Settings
26-28th October 2023
Scuola Normale Superiore, Pisa

First edition: June 2025
e-Isbn 978-88-7642-806-7
doi https://doi.org/10.2422/978-88-7642-806-7

Contents

From expert testimony to lay belief: a Bayesian view

1. *Introduction*

Modern societies are crucially dependent on the opinion of experts. The asymmetric relationship between experts and non-experts poses many socially impactful problems that have been subjected to the lens of different disciplines, including philosophy of science, social epistemology, argumentation theory, cognitive science, and psychology. In this paper, we focus on one fundamental problem that lies at the intersection of these approaches: how lay reasoners should update their beliefs in some hypothesis or claim H given that some expert asserts that H. We will approach this question from the perspective of Bayesian epistemology, following a recent trend that applies Bayesian models to represent, clarify, and manipulate traditional concepts in all of the above-mentioned areas [3,9,21,29,33,37]. More precisely, we shall treat the assertion of H by the expert as a case of testimony, and study how a Bayesian layman should update his probabilistic beliefs in H given this evidence. Bayesian confirmation theory [8, 13] will be helpful in modeling the notion of expert reliability and how it affects laymen's reasoning.

We proceed as follows. In section 2 we introduce the elements of the Bayesian approach to uncertain reasoning and confirmation of hypotheses. In section 3, we model expert opinion as evidence from testimony, developing a Bayesian model of both the experts' reliability and the evidentiary impact of their testimony in terms of Bayesian confirmation measures. In section 4, we show how our model impacts the discussion concerning expert opinion both in social epistemology and argumentation theory. As for the former, we show how the idea of "epistemic deference" amounts to ignoring the fallibility of real experts and fails as a general strategy of belief updating in the face of expert opinion. As for the latter, we elaborate on the recent discussion of *ab auctoritate* reasoning to clarify and evaluate this much-disputed argumentative strategy. By adopting a model of expert reliability based on Jeffrey conditionalization, we argue for a new characterization of *ab auctoritate*, more faithful to the complex epistemic interplay between laymen and experts. A short conclusion and some ideas for future research are offered in section 5.

2. *Bayesian reasoning, confirmation, and evidence*

In the following, we will treat both laymen and experts as ideal epistemic agents in the Bayesian sense. In this section, we offer a quick overview of the Bayesian framework, which will be applied in the next sections to the problem of expertise and testimony.

Updating beliefs on evidence. If Ann is a Bayesian agent, her beliefs about a given hypothesis H (such as "It will rain tomorrow") are represented by the (subjective) probability $p(H)$ she assigns to H. When Ann receives some piece of evidence E relevant to H (like "Tonight the sky is cloudy"), Ann will change her belief in H by moving from $p(H)$ to the conditional probability $p(H|E)$ of H given E. The updating rule is governed by the Bayes theorem:

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)} \tag{2.1}$$

where $p(H)$ is often called the "prior" probability of H and $p(H|E)$ its "posterior" probability (viz. before and after receiving evidence E). Bayes rule makes clear that the strength of Ann's posterior belief in H should be proportional both to that of her prior belief in H and to the so-called *likelihood* of H, *i.e.*, the probability $p(E|H)$ expressing how well H "explains" E or, better, how much E is expected in the light of H.[1]

For most applications, it is convenient to express the denominator of the Bayes formula, *i.e.*, the probability of the evidence, in its "unpacked" form using the rule of total probability. The resulting updating rule, which is equivalent to eq. (2.1) above, will look like this:

$$p(H|E) = \frac{p(E|H)p(H)}{p(E|H)p(H) + p(E|\neg H)p(\neg H)} \tag{2.2}$$

This new formulation of Bayes theorem makes clear that $p(H|E)$ must be a function of three different probabilities: the prior $p(H)$ of the hypothesis, its likelihood $p(E|H)$, and the likelihood of its negation, *i.e.*, $p(E|\neg H)$. Intuitively, H is the more probable in the light of E the more H was probable before observing E and the more E was expected given H instead of $\neg H$.

[1] For an introduction to Bayesian reasoning see [20]; see also [23] and [33] for more advanced discussions.

To make things more concrete, suppose that Ann is going to take a COVID-19 test in order to see whether she is ill (H) or not ($\neg H$). Before taking the test, she may believe she has a 5% probability of being ill, perhaps on the basis of her knowledge about the prevalence of the disease in her area. This fixes her prior probability $p(H)$ to 0.05 (and her prior probability of *not* being ill to $p(\neg H) = 1 - p(H) = 0.95$). Consulting the test's leaflet, Ann can also find the two likelihoods required by Bayes theorem. The "sensitivity" of the test, or, in other words, its true positive rate, expresses the probability $p(E|H)$ that the test will turn out positive (E) if Ann were actually ill; let's assume that $p(E|H) = 0.7$. The "specificity" $1 - p(E|\neg H)$ of the test is instead related to its false positive rate, expressing the probability that Ann gets a positive result even if she is actually not ill; let's assume that the specificity is 90%, and hence $p(E|\neg H) = 0.1$. Putting these numbers into eq. (2.2) gives us:

$$
\begin{aligned}
p(H|E) &= \frac{p(E|H)p(H)}{p(E|H)p(H) + p(E|\neg H)p(\neg H)} \\
&= \frac{0.7 \times 0.05}{(0.7 \times 0.05) + (0.1 \times 0.95)} \simeq 0.27
\end{aligned}
$$

Thus, Ann has a 27% probability of having COVID-19 given that the test result was positive. Note that this figure is significantly lower than 50%, meaning that it is still more probable that Ann is fine than she is ill. The reason is that we assumed that being ill was quite improbable (5%) in the first place; still, such probability has much increased (more than five times) after observing the test result. This increase in probability signifies how much the evidence given by the positive result "supports" or "confirms" the hypothesis of illness. Bayesian confirmation theory aims at making this intuition precise.

Confirmation as evidential support. Another interesting formulation of Bayes theorem is obtained expressing the relevant probabilities, following the gambling use, in terms of "odds". In statistical jargon, the odds of a hypothesis H refers to the ratio $o(H) = \frac{p(H)}{p(\neg H)}$ of the probability of H to that of its negation. Re-writing Bayes theorem in terms of these ratios leads to its "odds" form:

$$
\frac{p(H|E)}{p(\neg H|E)} = \frac{p(E|H)}{p(E|\neg H)} \times \frac{p(H)}{p(\neg H)} \tag{2.3}
$$

In this new form, the theorem tells us that the ratio of the posterior probabilities of H and its negation is given by the ratio of the prior probabilities multiplied by the so-called "likelihood ratio", *i.e.*, the ratio of the likelihoods of H and of its negation. Adopting the odds notation and writing "$LR(H,E)$" for the likelihood ratio, we can re-write eq. (2.3) as follows:

$$o(H|E) = LR(H,E) \times o(H) \tag{2.4}$$

That is to say, the posterior odds of H given E are just its prior odds multiplied by the likelihood ratio.[2]

The advantage of this odds formulation is that it makes very clear the role of the evidence E in changing Ann's belief in H. Let's consider again our previous example relative to Ann taking a COVID-19 test. Initially, the odds of Ann in favor of H (the hypothesis of being ill) are very low: $o(H) = \frac{0.05}{0.95} \simeq 0.053$. However, when Ann takes into account the positive test result, she can calculate a likelihood ratio of $LR(H,E) = \frac{p(E|H)}{p(E|\neg H)} = \frac{0.7}{0.1} = 7$. This number tells Ann how much the evidence received by the test speaks in favor of H rather than of its negation: multiplying her prior odds $o(H)$ by this factor gives her posterior odds in favor of H:

$$o(H|E) = LR(H,E) \times o(H) = 7 \times 0.053 \simeq 0.37$$

Thus, H is now much more likely than before, indeed seven times more.[3]

More generally, the likelihood ratio LR is a possible measure of how strongly a piece of evidence E supports or undermines a given hypothesis H, *i.e.*, of how much H is confirmed or disconfirmed by the evidence.[4] In Bayesian confirmation theory, we say that H is confirmed by E iff the probability of H is increased in the light of E: *i.e.*, $p(H|E) > p(H)$; that H is disconfirmed by E iff the probability of H is decreased in the light of E: *i.e.*, $p(H|E) < p(H)$; and that E is neutral for H iff the probability of H is left untouched by learning E: *i.e.*, $p(H|E) = p(H)$. It is now not difficult to prove that $LR(H|E) > 1$ iff E confirms H; that $LR(H|E) < 1$ iff E disconfirms H; and that $LR(H|E) = 1$ iff E is neutral for H. In the practice of many scientific disciplines, for instance medicine, the likelihood ratio is routinely employed (often under the name of "Bayes factor") as a measure of how well some evidence (*e.g.*, the result of a medical test)

---

[2] Note that some authors – like [23, 21] and [3, 17] – call "likelihood ratio" the reciprocal of $LR(H,E)$, *i.e.*, $\frac{p(E|\neg H)}{p(E|H)}$. We follow here the standard terminology in contemporary Bayesian confirmation theory [8, 33].

[3] If one has the odds o in favor of H, one can calculate its probability p as follows: $p = \frac{o}{o+1}$. As one can check, with this formula the just calculated odds of 0.37 gives us the 27% probability of section 2.

[4] Note that there are many, non-equivalent, confirmation measures, and many arguments in favor or against each of them [8, 14, 15, 33].

supports or undermines a given hypothesis (*e.g.*, a diagnosis). In Section 3 we shall return to the role played by likelihood ratios or Bayes factors in the confirmation of hypotheses.

Updating beliefs on uncertain evidence. Above, we employed Bayes theorem to say how Ann has to rationally update her belief in the face of incoming evidence. In other words, we assumed that Ann had a certain degree of belief $p(H)$ in the truth of H at some moment t_0, after which she receives some evidence E. At the new time t_1, Bayesian conditioning prescribes that Ann's degree of belief in the truth of H changes from her old credence $p(H)$ to her new credence $p'(H)$, which must be equal to $p(H|E)$ as given by eq. (2.1). Crucially, we assumed here that E is veridical and certain, so that, after receiving E, $p'(E) = 1$. This means, for instance, that there are no doubts in the reading of the COVID-19 test result, which is certainly positive in our example above.

Now suppose that just before the result is displayed on the test, a power failure occurs in Ann's room. In the dark, Ann cannot be completely confident about the outcome appearing on the test. Compared with the previous case, where Ann knew with certainty whether the test was positive or not, there is now an added dimension of uncertainty concerning the interpretation of evidence. This kind of uncertainty cannot be addressed by standard Bayesian conditioning. Instead, Ann's belief updating will have to follow "Jeffrey Conditioning" [24, 25]:

$$p'(H) = p(H|E)p'(E) + p(H|\neg E)p'(\neg E) \tag{2.5}$$

In words, the new probability Ann assigns to H is the sum of her old probabilities in H given E and in H given $\neg E$, respectively weighted by the (new) probability that E is actually true.[5]

For example, suppose Ann is fairly confident that she sees a double line (*i.e.*, a positive result). Let us set her subjective probability to $p'(E) = 0.8$. Her belief updating on this uncertain evidence will then be as follows:

$$\begin{aligned} p'(H) &= p(H|E)p'(E) + p(H|\neg E)p'(\neg E) \\ &= 0.27 \times 0.8 + 0.017 \times 0.2 = 0.22 \end{aligned}$$

Not surprisingly, Ann's degree of belief in H is now lower than in the previous case (22% as opposed to 27%), due to the uncertainty of the evidence.

[5] Note that if Ann is certain of her evidence, *i.e.*, $p'(E) = 1$ and hence $p'(\neg E) = 0$, then Jeffrey conditionalization given by eq. (2.5) immediately reduces to standard Bayesian conditionalization as in eq. (2.1).

This concludes our survey of the Bayesian framework employed in this paper, which is applied to the problem of expert testimony in the next section.

3. *A Bayesian model of expert testimony and reliability*

When Bob, a layman, listens to the opinion of some recognized expert, Bob's beliefs will typically change in one way or another. Such change is the result of a complex interaction between the two, which usually involves both epistemic and non-epistemic aspects (*e.g.*, trust), that are being studied across different disciplines (including social epistemology and philosophy of science, social and cognitive psychology, political science, and so on). Here, we shall focus on the purely epistemic dimension of the layman-expert relationship.

More precisely, we shall assume that Bob is a Bayesian agent and study how his (probabilistic) beliefs in some hypothesis H should change when some recognized expert testifies that H is true. In general, we can think of a (scientific) expert as a particular kind of witness, who has privileged epistemic access to a particular state of the world that falls within her domain of expertise. Intuitively, Bob's credences will depend, at least in part, on the *reliability* of the expert. The more the expert is reliable, the more the expert testimony should impact on Bob's belief in H. In this section, we apply the framework introduced in Section 2 to analyze these issues.[6]

Expert testimony as evidence. Suppose that Bob, who has some prior beliefs concerning H, learns that an expert has testified that H is true. The expert's opinion is the evidence on which Bob updates his beliefs. To reflect this in the notation, we shall write "E_H" for "expert E testifies that H is true". For the sake of simplicity, we shall assume that the expert's testimony is "categorical": *i.e.*, it corresponds to a clear-cut affirmation or negation of the hypothesis under evaluation, without considering degrees of certainty. In other words, we only consider cases where the expert says, for instance, that it will rain (for sure) tomorrow (E_H), but not cases where the expert says that it is likely that it will rain tomorrow, or that it will rain with, say, 80% probability. Thus, we shall focus on Bob's posterior assessment $p(H|E_H)$ of the probability of H given the expert testimony in favor of H.

[6] A problem we are not going to discuss here, and that has attracted much attention in the recent philosophy of science literature, starting at least from [3], is whether multiple pieces of evidence coming from independent sources of information confirm a hypothesis more than a less "varied" *corpus* of evidence (the so-called Variety of Evidence Thesis). For Bayesian analyses of this problem see, *e.g.*, [5, 6, 22], and references provided there.

From eq. (2.1) and eq. (2.2), we know how Bob should update his beliefs in H:

$$\begin{aligned} p(H|E_H) &= \frac{p(E_H|H)p(H)}{p(E_H)} \\ &= \frac{p(E_H|H)p(H)}{p(E_H|H)p(H) + p(E_H|\neg H)p(\neg H)} \end{aligned} \tag{3.1}$$

In words, Bob's posterior belief in H, given the expert's testimony that H, is a function of his previous belief in H, represented by the prior $p(H)$, and of the two likelihoods associated to the testimony: the probability $p(E_H|H)$ that the expert (correctly) reports H when H is true, and the probability $p(E_H|\neg H)$ that the expert (mistakenly) reports H when H is false.

Interestingly, this way of construing the testimony of a well-informed witness as a piece of evidence in hypotheses evaluation is common in the legal domain [1, 16]. For example, in a murder trial, a judge might be unsure whether the suspect was at the crime scene when the murder happened (H). By asking a witness, the judge might gain useful insights and update his confidence in H accordingly. A positive report by the witness will likely increase the judge's subjective probability of the suspect being at the crime scene, given that the witness is reliable (*e.g.*, he was a bystander and does not have any interest in reporting the false). Conversely, a negative report by a reliable witness will decrease the judge's confidence in H. The more reliable the witness, the stronger the evidential support of his testimonial report. As we shall see, this way of modeling testimony is useful also in studying the layman-expert relation.

Expert reliability as confirmation. As we have noted, the strength of expert testimony will depend on the expert's reliability. But what does it mean for an expert witness to be reliable from a Bayesian perspective? To express expert reliability in our framework, we shall assume what we may call the "diagnostic-test model" of expert witnesses *cf.* [3, Chapter 3]. In other words, we shall assume that an expert, like the COVID-19 test considered in Section 2, also has rates of true and false positives, expressed, respectively, by the likelihoods $p(E_H|H)$ and $p(E_H|\neg H)$. In the case of medical tests, such probabilities are readily available in the test leaflet; in the case of experts, it is instead difficult to quantify precisely their "sensitivity" and "specificity". Still, it is a useful idealization to construe the reliability of an expert as a function of those likelihoods.[7]

[7] Moreover, one can argue that there are (fallible) indicators, discussed at length in the expertise literature [7, 18, 30], that may help in evaluating at least approximately the reliability of experts.

Accordingly, we shall define the reliability $Rel_E(H)$ of expert E relative to hypothesis H as the relevant likelihood ratio $LR(H, E_H)$:[8]

$$Rel_E(H) = LR(H, E_H) = \frac{p(E_H|H)}{p(E_H|\neg H)} \tag{3.2}$$

This is to say that the expert testimony is more reliable the more likely it is that the expert testifies that H is true when H is in fact true, and the less likely it is that the expert testifies that H is true when H is instead false. Note that this amounts to expressing the expert's reliability in terms of the confirmation that the expert report E_H confers on H [13]. This becomes clear if one applies Bayes theorem in its odds form to the case at hand, *i.e.*, eq. (2.4):

$$o(H|E_H) = Rel_E(H) \times o(H) \tag{3.3}$$

This is to say that, after receiving the expert report E_H, Bob will revise his prior odds by multiplying them by the reliability of the expert. This means that the more reliable the expert is, the more his report will confirm H. In general, we may say that the expert is "reliable" (relative to H) when $Rel_E(H)$ is greater than 1; and "unreliable" if $Rel_E(H)$ is smaller than or equal to 1, since in this latter case his report will not increase, but possibly decrease, the probability that H is true.[9]

Again, the legal domain provides a useful illustration of this idea. Consider, as a toy example, a murder case: during a crowded football match, a supporter gets shot and dies. The murder weapon is readily found, but nobody has seen the shooter. A man, Mr. Guy Hapless, is seen walking briskly towards the exit. The authorities stop him and get his fingerprints. The judge assigned to the case then asks a forensic dactyloscopy expert to compare them with those found on the gun. The expert reports that the two fingerprints match. How should a Bayesian judge evaluate such evidence?

[8] See also [3, 9, 13]. Note that $Rel_E(H)$ "measures the reliability of the witness *with respect to the report in question* [here, H] and not the reliability of the witness *tout court*", as [3, 17] put it. The latter reliability will depend on the expert's (average) performance on several different hypotheses in some field; reliability $Rel_E(H)$ as defined here only measures the expert's performance on a single hypothesis H. Note that much recent literature follows [3] in modeling reliability in a different way, using Bayesian networks; a minimal network will comprise three nodes, one for the hypothesis, one for the witness report (our E_H), and one for the reliability of the witness, which can be inferred *a posteriori* by manipulating the other nodes *cf.* [29, Chapter 10].

[9] Bovens and Hartmann [3], who define expert reliability as $1 - \frac{1}{LR(H,E_H)}$, consider as "fully unreliable" an expert (or, more generally, a witness) for whom $Rel_E(H) = 1$; here, we also consider the case of "misleading" witnesses, for which $Rel_E(H) < 1$, whose testimony actually disconfirms the hypothesis under consideration.

Suppose that, initially, the judge has no specific reason to suspect that "Mr. Hapless is guilty" (H), besides the fact that he was attending the match. Moreover, let's suppose that the judge is (rightly or wrongly) sure that the fingerprints found on the gun will certainly be the murderer's. This means that the prior probability of H is $p(H) = \frac{1}{10000}$ and its prior odds $\frac{1}{9999}$, both very low. The expert's report that Mr. Hapless' fingerprints match those found on the gun will of course confirm H. How much will depend on the expert's reliability: let's assume that our expert is highly reliable, detecting true positives in 92.5% of the cases, and reporting a false positive only 0.1% of the times.[10] Then, our expert's reliability $Rel_E(H)$ will be equal to 925 and, applying eq. (3.3), we obtain:

$$o(H|E_H) = Rel_E(H) \times o(H) = 925 \times \frac{1}{9999} \simeq 0.092$$

which corresponds to a posterior probability of H approximately equal to 8%. Not surprisingly, the new probability $p(H|E_H)$ is significantly higher than the prior (roughly 800 times so), since the expert testimony has strongly confirmed the hypothesis of guilt. Given the low initial probability, however, such strong confirmation is not enough to think that Mr. Hapless, on this only evidence, is more likely guilty than not.

In this connection, let us note that some conventions have been established in the literature on Bayes factors (following seminal work by Turing and Good) to qualitatively assess the "weight of evidence" in favor of H provided by different likelihood ratios. For instance, in our example above the weight of evidence provided by the expert is well above the highest threshold of "decisive" evidence according to table 1. Interestingly, such literature may provide guidance in developing conventional expertise thresholds for any given scientific domain in light of its epistemic status.

TABLE 1 Interpreation of likelihood ratios LR as a measure of the weight of evidence according to [26] and [27].

LR	Weight of evidence	
	[26]	[27]
$1 - 3.2$	Barely worth mentioning	Barely worth mentioning
$3.2 - 10$	Substantial	Substantial
$10 - 32$	Strong	Strong
$32 - 100$	Very strong	Strong
> 100	Decisive	Decisive

[10] Expert performance in recognizing matching fingerprints has been the object of several studies [28,34,35]. For the sake of our example, we have selected the results that show higher expert reliability, as reported in [35].

4. *Bayesian rationality, epistemic deference, and arguments from authority*

In this section, we show how the tools from Bayesian epistemology that we have presented so far allow us to shed light on two much-debated topics in social epistemology and argumentation theory: epistemic deference and *ab auctoritate* arguments.

Epistemic deference vs. Bayesian rationality. Social epistemology studies the social aspects of knowledge production and dissemination, including the role of testimony, trust, and communication in the pursuit of knowledge by individuals and groups [19]. A fundamental question in this area is how one should be guided by other people's beliefs. One strategy discussed in the literature is *complete epistemic deference*, which is applied to interactions with experts, *e.g.*, [12]. This strategy implies that, when Bob receives the expert opinion E_H, he will have to revise his beliefs by adjusting them to the expert's. According to [12, 480], this means that if $p_{exp}(H)$ is the probability that expert E assigns to H, then Bob should simply adopt such probability as his own, so that $p_{Bob}(H) = p_{exp}(H)$. Since we are only considering categorical opinions from experts, this would imply that, if the expert testifies that H, then Bob should plainly accept H as certain, *i.e.*, $p_{Bob}(H|E_H) = 1$.

Epistemic deference is of course an extreme updating strategy in the face of expert opinion. It amounts to treating the expert as a fully veridical and trustworthy oracle. From a Bayesian standpoint, this strategy cannot be justified in general; the only case where it could apply is when two highly idealized assumptions are made:

1. experts are completely reliable;
2. laypeople's prior beliefs on H should not be taken into account.

As to (1), in our framework a completely reliable expert would be one who cannot make mistakes, *i.e.*, for which $p(E_H|\neg H) = 0$; this would make $Rel_E(H)$ meaningless, or better tending to infinity as $p(E_H|\neg H)$ approaches 0. Since we are interested in modeling real, fallible experts, we exclude this case from consideration. As for (2), priors play of course a crucial role in the Bayesian treatment of how our layman Bob reacts to the expert testimony (*cf.* eq. (3.3)). Here, a natural objection to such treatment is the following: it is sensible to exclude the layman's prior beliefs from the picture, since the expert holds (by definition) more accurate beliefs, hence the layman can just rely on the expert and needs not consider his own priors when updating on an expert report.

While intuitively convincing, we think one should rebut this objection, as our fictitious murder example shows. In that case, the Bayesian judge is not "deferent" to the fingerprint expert at all, at least not in the sense suggested by [12]. For sure, the judge duly takes into account the expert report as a crucial piece of evidence, without doubting it. This is as it should be, since while the judge is an expert in the legal domain, he should be regarded as a layman when it comes to recognizing matching fingerprints – and this is why the opinion of the forensic dactiloscopy expert is required.[11] However, the judge does not conclude that Mr. Hapless is guilty, as his prior for such a hypothesis is very low. This would happen if, as per (2) above, the judge ignores his own priors, meaning that $p(H) = p(\neg H) = 0.5$. Then, since $o(H) = 1$, eq. (3.3) would correctly reduce to $o(H|E_H) = Rel_E(H)$, so that the expert's opinion fully determines the posterior beliefs of the judge. However, in most cases, this would amount to committing the so-called prosecutor's fallacy, an instance of the "base rate fallacy", where the epistemic agent tends to conflate the posterior probability for the likelihood $p(E|H)$ [1, 16, 17].

The importance of the layman's priors in updating on expert evidence highlights an important distinction between two concepts: the "weight of evidence" in favor of H and the "acceptance" of H. The former is provided by the expert's report and appropriately measured, as we argued, by his reliability $Rel_E(H)$; the latter crucially depends, in addition, also on the layman's priors. The idea of epistemic deference tends to obscure such distinction and should be resisted accordingly.

Arguments from authority and layman testimony. Argumentation theory studies the social practice of "giving and asking for reasons" [4], focusing on reasonable and fallacious uses of arguments in dialogical contexts [11, 22]. An important case is that of arguments from authority (*ab auctoritate*). This strategy is employed whenever an arguer appeals to the opinion of an authority to support her standpoint in a critical discussion. The kind of authority we are interested in here is an epistemic one – and therefore this argument is also known as *appeal to expert opinion* [36]. For example, Bob and Carl are discussing whether it is true that food that has fallen on the floor within 5 seconds will still be safe to eat (H). Carl is confident that H while Bob is convinced of the opposite. One week before the conversation, Carl overheard on the TV an expert reporting that H (*i.e.*, Carl came to know the evidence E_H). To persuade Bob, Carl appeals to the authority of that expert claiming that H. This means that Bob has received some evidence from Carl's testimony, which we will denote by

[11] More generally, the epistemic roles of "layman" and "expert" are always domain-dependent: the same person can be an expert in one domain and a layman in another (see [18]).

L_{E_H} (where L is for "layman"). How is Bob supposed to change his mind in light of this evidence?

In argumentation theory, this scenario has been studied mostly from a procedural perspective (rather than an epistemic one): *i.e.*, what is reasonable for a discussant to ask in such a scenario, rather than how they should update their beliefs. Walton [36] has come up with a set of "critical questions" that a lay discussant is always entitled to ask. Their function is to momentarily shift the burden of proof on the proponent of the argument, by identifying its potential weak spots.[12] While these questions carry a useful heuristic value, they do not inherently offer criteria for evaluation of the answers. As we argue, a Bayesian approach provides valuable insights into the assessment of arguments from authority.

While some authors have explored this idea [9,21], what is still needed is a framework that specifically takes into account the significance of second-order testimony, viz. Bob's assessment of the evidence provided by Carl's report of an expert testimony. This, we submit, is the crucial aspect of interest in the analysis of arguments from authority, which distinguishes such case from that of "standard" expert testimony. In fact, this scenario adds a further dimension of uncertainty, compared to the fingerprint example in the previous paragraph. Not only is the link between H and E_H uncertain because of the fallibility of the expert providing the testimony; but, crucially, this evidence itself is also uncertain because it is not gained first-hand, but reported by a (lay) witness. Many things could go wrong and lead Carl to misreport what he heard – for example: Carl's bad faith, confirmation bias, motivated reasoning, failure of memory, hearing impairment, lack of understanding of the general context where the testimony was provided, *etc.*

To deal with this additional layer of uncertainty, we suggest proceeding as in Ann's example: like the sudden darkness of the room introduced an element of disturbance and opacity in the reading of the COVID-19 test, here Bob will have to deal with Carl's opaquely reported evidence. This is where Jeffrey Conditionalization comes back into the picture. Building on eq. (2.5), we can represent belief updating on a layman's report of expert testimony as follows:

$$p'(H) = p(H|L_{E_H}) = p(H|E_H)p'(E_H) + p(H|\neg E_H)p'(\neg E_H) \quad (4.1)$$

In words, Bob's new probability for H, given Carl's report L_{E_H} that an expert testified that H is true, will depend on the posterior probability of H

[12] They are: 1. How credible is expert *E* as an expert source?; 2. Is *E* an expert in the field *F* that assertion *A* is in?; 3. What did *E* assert that implies *A*?; 4. Is *E* personally reliable as a source?; 5. Is *A* consistent with what other experts assert?; 6. Is *E*'s assertion based on evidence?

being true given that the expert really testified that H – $p(H|E_H)$ – and, at the same time, on the probability of the hypothesis still being true in the light of the absence of the expert testimony – $p(H|\neg E_H)$.[13] Each of these values will have to be adjusted, respectively, for the probability that layman Carl reports the expert testimony correctly – $p'(E_H)$ – or incorrectly – $p'(\neg E_H)$. To assess such probabilities, Bob cannot rely on some objective assessment of Carl's reliability, contrary to what the judge could do, in our previous example, with the fingerprint expert. Indeed, one could see the difference between expert and layman witnesses precisely in this: that only for the former it is possible to estimate, at least approximately, the corresponding reliability. For the latter, Bob can only rely on his own subjective assessment of the reliability of Carl correctly reporting the testimony he has witnessed. It remains to be explored how Walton's critical questions can be useful for refining such an assessment.

5. *Conclusion*

In this paper, we addressed the epistemic problem of expertise through the lenses of Bayesian epistemology. First, we showed how a layman should update his beliefs on the basis of that particular kind of evidence provided by expert testimony, and how experts' reliability can be construed as the confirmation provided to the relevant hypothesis at issue. Then, we contrasted our Bayesian model with the strategy of complete epistemic deference as discussed in social epistemology, and argued that it can also shed light on arguments from authority. Future work on these lines can move in different but complementary directions. As to social epistemology, the Bayesian approach seems promising when applied to problems of peer-expert disagreement [31]. In argumentation theory, Bayesian insights can help in bridging the gap between procedural and epistemic approaches to reasonable dialogues [2]. Finally, the Bayesian standard offers a benchmark against which both experts' and laymen's beliefs can be tested, shedding light on potential cognitive biases in their interaction, as studied in cognitive science and psychology [10]. The promise is that of a formally sound and empirically grounded approach that is better equipped than existing ones to deal with some central problems in social and formal epistemology.

Acknowledgements. We thank Lorenzo Casini, Davide Coraci, Daria Mikhaylova and the other members of the MInD research group at IMT

[13] We can conceive of the absence of a report that H in two ways: either the expert actually testified $\neg H$—*i.e.*, $\neg E_H$ will be equivalent to $E_{\neg H}$; or the expert did not assert anything about H at all.

Lucca for useful discussions. We acknowledge financial support from the MUR-PRO3 project "Understanding public data: experts, decisions, epistemic values" (CUP: D67G22000130001). Gustavo Cevolani acknowledges financial support from the Italian Ministry of University and Research (MUR) through the PRIN 2022 project "Reasoning with hypotheses: Integrating logical, probabilistic, and experimental perspectives" (grant n. 2022ARRY9N; CUP: D53D23009570006), funded by the European Union (Next Generation EU).

References

[1] C. Aitken, F. Taroni and S. Bozza, "Statistics and the Evaluation of Evidence for Forensic Scientists", Wiley, Hoboken, NJ, third edition edition, 2021.

[2] G. Betz, *Evaluating dialectical structures with Bayesian methods*, Synthese **163** (2008), 25–44.

[3] L. Bovens and S. Hartmann, "Bayesian Epistemology", Oxford University Press, Oxford, 2003.

[4] R. Brandom, "Making It Explicit: Reasoning, Representing, and Discursive Commitment", Harvard University Press, Cambridge, Mass., 1994.

[5] L. Casini and J. Landes, *Confirmation by robustness analysis: a Bayesian account*, Erkenntnis **89** (2024), 367–409.

[6] F. Claveau and O. Grenier, *The variety-of-evidence thesis: a Bayesian exploration of its surprising failures*, Synthese **196** (2019), 3001–3028.

[7] H. Collins and R. Evans, "Rethinking Expertise", University of Chicago Press, Chicago, 2019.

[8] V. Crupi, *Confirmation*, In: "The Stanford Encyclopedia of Philosophy", E. N. Zalta (ed.), Metaphysics Research Lab, Stanford University, Spring 2021 edition, 2021.

[9] C. Dahlman and L. Wahlberg, "Appeal to Expert Testimony – A Bayesian Approach", Springer, Cham, 2015, 3–18.

[10] I. E. Dror, *Cognitive and human factors in expert decision making: six fallacies and the eight sources of bias*, Analytical Chemistry **92** (2020), 7998–8004.

[11] F. H. van Eemeren, "Argumentation Theory: A Pragma-Dialectical Perspective", Springer, Cham, 2018.

[12] A. Elga, *Reflection and disagreement*, Noûs **41** (2007), 478–502.

[13] R. Festa, *Testimonianze esperte e probabilità delle ipotesi*, L&PS – Logic & Philosophy of Science **XII** (2014), 3–39.

[14] R. Festa and G. Cevolani, *Unfolding the grammar of Bayesian confirmation: likelihood and antilikelihood principles*, Philos. Sci. **84** (2017), 56–81.

[15] B. Fitelson, *The plurality of Bayesian measures of confirmation and the problem of measure sensitivity*, Philos. Sci. **66** (1999), S362–S378.
[16] P. Garbolino, "Probabilità e Logica della Prova", Giuffrè Editore, Milano, 2014.
[17] G. Gigerenzer, "Calculated Risks: How To Know When Numbers Deceive You", Simon and Schuster, New York, 2002.
[18] A. I. Goldman, *Experts: Which ones should you trust?*, Philosophy and phenomenological research **63** (2001), 85–110.
[19] A. I. Goldman, "Why Social Epistemology is *Real* Epistemology", A. Millar A. Haddock and D. Pritchard (eds.), 2010, 1–29.
[20] I. Hacking, "An Introduction to Probability and Inductive Logic", Cambridge University Press, New York, 2001.
[21] U. Hahn, M. Oaksford and A. J. L. Harris, *Testimony and argument: a bayesian perspective*, In: "Bayesian Argumentation: the Practical Side of Probability", 2013, 15–38.
[22] C. L. Hamblin, "Fallacies", Vale Press, Newport News, Va., 1970.
[23] C. Howson and P. Urbach, "Scientific Reasoning: the Bayesian Approach", Open Court Publishing, Chicago, 2006.
[24] R. C. Jeffrey, "The Logic of Decision", University of Chicago Press, New York, 1965.
[25] R. C. Jeffrey, "Subjective Probability: The Real Thing", Cambridge University Press, Cambridge and New York, 2002.
[26] H. Jeffreys, "The Theory of Probability", Oxford University Press, Oxford, 1998.
[27] R. E. Kass and A. E. Raftery, *Bayes factors*, J. Amer. Statist. Assoc. **90** (1995), 773–795.
[28] P. J. Kellman, *et al.*, *Forensic comparison and matching of fingerprints: using quantitative image measures for estimating error rates through understanding and predicting difficulty*, PLOS ONE **9** (2014), 1–14.
[29] D. A. Lagnado, "Explaining the Evidence: How the Mind Investigates the World", Cambridge University Press, Cambridge, 2021.
[30] C. Martini, *Experts in science: a view from the trenches*, Synthese **191** (2014), 3–15.
[31] T. Mulligan, *The epistemology of disagreement: why not Bayesianism?*, **18** (2021), 587–602.
[32] B. Osimani and J. Landes, *Varieties of error and varieties of evidence in scientific inference*, British J. Philos. Sci. **74** (2023), 117–170.
[33] J. Sprenger and S. Hartmann, "Bayesian Philosophy of Science", Oxford University Press, Oxford and New York, 2019.
[34] J. M. Tangen, M. B. Thompson and D. J. McCarthy, *Identifying fingerprint expertise*, Psychological science **22** (2011), 995–997.

[35] B. T. Ulery *et al.*, *Accuracy and reliability of forensic latent fingerprint decisions*, Proc. Nat. Acad. Sci. India Sect. A **108** (2011), 7733–7738.
[36] D. Walton, "Appeal to Expert Opinion: Arguments from Authority", Penn State Press, University Park, PA, USA, 2010.
[37] F. Zenker, ed., "Bayesian Argumentation", Springer, Dordrecht, 2013.

Piero Avitabile
Gustavo Cevolani