

---

# Testimonial Justification under Epistemic Conflict of Interest

---

## Abstract

Can a hearer be rationally justified to have beliefs based on testimony alone when the source of his information is known to have conflicting epistemic goals? When it comes to belief justification, existing theories either recommend avoiding epistemic conflicts of interest or ignoring them. This is an important epistemological limitation. A theory that comes in degrees, capable of explaining what beliefs we are justified to hold and why, despite epistemic conflict of interest, is still lacking. Building on a game-theoretical approach, I suggest such a theory and argue that the hearer can justify some beliefs on testimony alone. This justification relies on an equilibrium concept, which is only reached in the long run. In addition, the hearer's justified beliefs will always be less accurate than those held by the original source. For instance, assume the speaker is a climate scientist who has good reasons to believe that a 2°C increase in temperature will lower the current global GDP by 10 percentage points. Under epistemic conflict of interest, a hearer will typically be justified to a belief close to that value, but not equal to it. The smaller the epistemic conflict of interest, the closer, on average, the hearer's and speaker's belief. These results highlight the importance of scientific norms which, in practice, are the embodiment of these equilibrium mechanisms and thus of scientific credibility.

**Keywords:** Social Epistemology, Game Theory, Epistemic Conflict of Interest, Philosophy of Testimony.

## 1 Introduction

Because our collective understanding of the world has never been so vast, we individually believe we understand a large part of it. But, most of the time, our justifications for these beliefs comes from the expertise of others. Yet those on whom we rely might have different epistemic goals than ourselves. They might, for instance, have different validation

---

standards in order to believe a proposition or simply different practical interests. Consider the example of beliefs regarding potential side effects of COVID-19 vaccines relying on mRNA technology. When first commercialised, tests on these vaccines had only been conducted by the pharmaceutical companies who produced and sold them, without time for independent replication or deepening on rare effects (Tanveer et al., 2022). What were we then justified to believe regarding the safety of these vaccines? More generally, what are we justified to believe when we have diverging interests from those on whom this justification relies?

Such settings have been described as cases of epistemic conflict of interest (Henderson (2020), Müller (2022) for a systematic review). They are challenging situations when it comes to belief justification, especially in social contexts and when the justification relies entirely on expert testimony. As a matter of fact, when it comes to belief justification, existing theories either recommend avoiding epistemic conflicts of interest or ignoring them<sup>1</sup>. Arguably, this is an important epistemological limitation. Assuming that it is always unjustified (or justified) to form any belief in case of epistemic conflict of interest seems unsatisfactory. It seems on the contrary intuitive that, while epistemic conflict of interest should have an impact on justification, that impact should not be all or nothing. Take the example of the COVID-19 vaccines mentioned earlier. Considering that it was fully unjustified to believe these vaccines were safe, without any graded notion of justification, seems unsatisfying. Such radical positions would open the door to a variety of conspiracy theories. Yet a theory that comes in degrees, capable of explaining what beliefs we are justified to hold and why, despite epistemic conflict of interest, is still lacking.

The reason why providing such a theory is so difficult is that epistemic conflict of interest creates situations which are strategic: agents know that their interests conflict and that trust is not granted a priori. This paper aims at addressing the question of the justification of a hearer's beliefs based on a speaker's testimony only, when there is epistemic conflict of interest between them. It will do so through game theory and specifically by introducing one of its central object: *equilibrium concepts*. Equilibrium concepts are assumptions on the beliefs agents hold regarding the behaviour of other agents. For instance, in a Nash equilibrium, all agents believe that actions are stable if no agent can benefit from a unilateral deviation. It follows that, at equilibrium, the behaviour of players is predictable. Therefore, even under epistemic conflict of interest, information provided at equilibrium cannot be deceiving.

---

<sup>1</sup>I review existing theories in detail in section 2.

---

Following Zollman (2021)'s suggested use of game theory for epistemic problems, and by further introducing the behavioural beliefs to lead to the Nash equilibrium, I will argue that it can be justified to ground belief on a speaker's testimony alone, even when her epistemic goals diverge from those of the hearer. Yet this belief will always be *more imprecise*, in a specific sense, than the speaker's original belief. For instance, assume the speaker is a climate scientist who has good reasons to believe that a 2°C increase in temperature will lower the current global GDP by 10 percentage points. Under epistemic conflict of interest, a hearer will typically only be justified to believe that that the reduction is between 5 and 15 percentage points. Importantly, this justification only emerges *at equilibrium*. Equilibrium behaviours appear in the long run, after a long repetition of trial and error. In practice, these equilibrium mechanisms materialise as social and scientific norms and are embodied through institutions (Bicchieri (2016), Fehr et al. (2002), Henrich and Muthukrishna (2021)). My results thus contribute to highlighting the importance of these social elements for science to be an influential and reliable basis of beliefs.

## 2 Existing views and limitations

Can I rely on the testimony of others to acquire justified belief? This problem has been called the Source Problem for testimonial entitlement. As argued by Hardwig (1985), for a hearer to be justified to ground a belief fully on a speaker's word, two conditions need to be reached:

- (a) The speaker must be an intellectual authority on the subject
- (b) The hearer must be justified in deferring to the speaker regarding her claim

While most philosophers agree upon condition (a), two sides have been taken regarding (b). In the tradition of Kant's imperative for epistemic independence, authors such as Chisholm (1989) have defended the idea that beliefs are justified only through the verification abilities of perception or logic. Therefore, this side has been called *reductionist*. Yet others have argued that testimony could be a primary source of justification without the need to appeal to other warrants. For proponents of this anti-reductionist side, the central idea is that the *default* epistemic position regarding a testimonial source should be trust. It is clear that the latter

---

position is controversial and demands strong arguments in its favour. Mainly, two objections have been commonly raised against it:

- (i) Agents can have different practical, moral and epistemic motivations
- (ii) Agents can lie

Several distinct positions have been held on the anti-reductionist side regarding the Source Problem. I will review them and show that they all leave the case of justification under epistemic conflict of interest aside. They are *prima facie* theories arguing that it is justified in believing *p*, absent defeat such as epistemic conflicts of interest. I will further argue that, in doing so, they fail to explain important cases where testimonial justification seems to hold. An *ultima facie* theory offering a deciding upon belief justification in presence of an important defeater such as epistemic conflict of interest is, I believe, still lacking.

**Ruling out epistemic conflicts of interest**—Although a variety of anti-reductionist theories have been proposed in recent years, most of them support the idea that testimony can be a justification for belief only if the speaker has no systematic interest in lying. That is, avoiding objection (ii) above.

A first approach to deal with objection (ii) has been made in Goldberg (2010) and Goldberg (2014). Essentially, Goldberg offers a fallibilist argument: lies can be accounted for as simple mistakes in the justification process. What matters for justification is that speaker and hearer employ the same cognitive process for the formation of their belief. In response to this view, authors like Ross (1986), Hinchman (2005) and Hinchman (2014) have argued that testimony is not evidence of the same kind that our senses can give us. Memorial failure is a random process which can sometimes occur, whereas insincerity is a voluntary one, which can be, at least to some extent, systematically identified. This view has been called the assurance view. It stresses the importance of interpersonal relationships in justification. According to this view, when a hearer acquires testimonial justification for believing a proposition on the basis of a speaker's claim, his belief is at least in part justified by the speaker's assurance. This assurance, by nature, is non-evidentialist. Moran (2006), Moran (2018), Faulkner (2007), Faulkner (2011), Fricker (2012), Zagzebski (2012) and McMyler (2011) further stress the importance of the speaker's intentions, when informing a hearer about a proposition. In the particular case of

---

epistemic conflict of interest, the fallibilist approach is arguably unsatisfactory: lies are not the result of mistakes, but of clear intentions.

Following the assurance view, Graham (2010) and Graham (2015) argue that in practice, what a hearers oughts to do is to set up a filtering process that narrows the range of entitlement conferred by testimonial exchanges to situations where no epistemic conflict of interest is present. Yet, as argued by Simion (2021), one limit to this argument is that our deception detection abilities are, in most contexts, very limited and mostly unable to provide us with a fine tuned filter. In other words, we are bad at detecting problematic situations. Consider the vaccine and climate change examples I introduced above. For a non-expert audience, it is almost impossible to detect problematic claims in technical domains with which they are not familiar. This is particularly true of new or unexpected situations such as the global COVID-19 pandemic. Then, it was very hard for a non-expert audience to distinguish between diverging expert claims. At least in these situations, a conscious filtering process seems insufficient to grant justification to beliefs acquired through testimony under epistemic conflict of interest.

A more radical line of defence has been held regarding anti-reductionism, essentially by Burge (Burge (1993), Burge (1997), Burge (2013)). For Burge, we are a priori prima facie entitled to take intelligible affirmation at face value (Burge, 1993, 472). Burge's argument is mostly about rationality. Its function is to generate true content. So, if agents are rational, testimony about knowledge should by default generate justified beliefs. Recently, Simion (2021) supported this view by arguing that it is because of the presence of a social contact between the speaker and the hearer that the hearer is prima facie entitled to form beliefs based on speaker's assertions. However, for both Burge and Simion, testimony can only be a source of justification for belief in *non-problematic* situations, namely when no systematic epistemic conflict of interest exists. In Simion's words: "*If I know [...] that you stand to gain from lying to me, I am not entitled to expect norm conformity anymore.*"<sup>2</sup> I agree with Burge and Simion that, without further argumentation, a social contract does not suffice to explain belief justification in situations of epistemic conflict of interest. However, I believe that there are situations where testimony provides such a justification, despite the presence of epistemic conflict of interest.

---

<sup>2</sup>Simion (2021) p.910

---

**Two classes of problematic cases**—Excluding all situations where a speaker could use deception to her advantage from the range of testimony-based justified beliefs seems very restrictive. The literature on epistemic conflicts of interest has brought up numerous cases where testimony still seems a plausible justification for belief<sup>3</sup>. Consider the two following generic examples<sup>4</sup>:

1. An obvious case is one where the speaker cares about the belief of the hearer and of its truth, but also has other practical motivations<sup>5</sup>. Prominent cases include the COVID-19 vaccines relying on mRNA technology. Clearly, the pharmaceutical companies that have produced them are experts in these technologies (condition (a)). When they were first adopted during the pandemic, the urgency of the situation compelled healthcare institutions to adopt them before normal peer reviewed testing and certification could be done in order to assess their reliability. So the confidence we had in our belief regarding the safeness of these vaccines had to rely fully on their producers' word. In addition, pharmaceutical companies also had an interest in selling their vaccines and potentially in downplaying any side effect or disadvantage they may have had. So it was unclear if condition (b) was holding. For instance consider the following proposition:

$p_\omega$ : " The probability that the vaccine has negative side effects is at most  $\omega$ "

The vaccine producer might believe that  $p_{0.1}$  is true but would prefer, in order to remain below the maximum tolerated level of health risk, to convince the hearer that  $p_{0.05}$  is true.

2. A second important case is one where the speaker is a pure epistemic agent, but has different validation standards than the hearer. Consider a climate scientist reporting on the impact of a given global temperature increase on GDP. She has to state a proposition such as:

$q_\omega$ : " A 2°C increase in temperature will lower the current global GDP by  $\omega$  percentage points"

---

<sup>3</sup>Müller (2022) provides a useful taxonomy on the subject.

<sup>4</sup>Notice that in all cases, I consider a continuous class of proposition  $p_\omega$ , where  $\omega$  is an element of an interval. This has an importance for the corresponding model I'll develop.

<sup>5</sup>Then, the speaker is not a pure epistemic agent, as defined by Joyce (1998), because she does not care only about reaching the truth.

---

Such predictions rely on a great number of prospective simulations, all of which differ in the underlying hypothesis or the mechanisms accounted for (see for instance those reported in Meinshausen et al. (2009)). These simulations can be ranked by a continuous confidence level, for instance by a measure of its predictive accuracy. How accurate a simulation has to be to be considered reliable might differ between speaker and hearer. For instance, it might be known that the speaker is more cautious than the hearer and would consider that even less accurate simulations should be considered. This would cause her to support a more pessimistic proposition. That is, the speaker would report  $q_\omega$  while the hearer, in her position, would have reported  $q_{\omega'}$ , where  $\omega > \omega'$ .

In both these cases, is it always unjustified for the hearer to believe any speaker's statement about mRNA vaccines or climate change? Answering yes to these questions implies that a variety of conspiracy theories cannot be avoided by testimony alone. I do not share this intuition. In my view, in the cases described above we are justified to hold at least some belief regarding the COVID-19 vaccine side effects, or the effect of climate change on GDP, based on testimony alone. Yet a theory capable of explaining what belief we are justified to hold and why testimony can provide this justification despite the epistemic conflict of interest is necessary, but still lacking.

### 3 Game theory for social epistemology

#### 3.1 Epistemic agents

Traditionally, game theory has been used to study strategic interactions between agents concerned with pragmatic criteria. Most famous examples such as the prisoner's dilemma or coordination games are illustrated by players concerned with escaping prison (Poundstone, 1993) or choosing a Saturday night leisure activity (Luce and Raiffa, 1989). Yet Joyce (1998) successfully argued that utility functions should also be used to capture purely epistemic motivations of agents concerned partially or fully with learning the truth. This program has been pushed further in many directions (Pettigrew, 2016). While most of this research has been oriented towards decision theory (Bradley, 2017), Zollman (2021) made a particular case in favour of game theory. He argued that agents interested in epistemic values can find

---

themselves involved in game theoretical situations, even when they have similar interests.

In his paper, Zollman considers agents who are epistemically altruistic, that is, who care only about reaching the truth. From a formal perspective, this is captured by utility functions called scoring rules. My starting point will be similar to Zollman. Let me introduce a slightly more general formal framework<sup>6</sup> to illustrate my point. Let  $\Omega = [0, 1]$  be the set of possible states. The set of states represents all the possibilities agents agree upon. The agent is unsure about the true state and holds a credence function over the possible states, formally:  $c \in \Delta(\Omega)$ .  $c$  is a probability distribution over  $\Omega$ . Let's call  $b_i \in [0, 1]$  agent  $i$ 's belief and assume that, for him, this is the state expected to be the true one. This approach to the connection between belief and credences is very close to Weatherson (2016). For our agent, his belief  $b_i$  is his action variable. An epistemically altruistic agent's utility function is maximised when his belief corresponds to the true state and, the more distant it is from that state, the more it decreases. A simple way of illustrating this is by the following utility function:

$$u_i(b_i, \omega) = -(b_i - \omega)^2$$

where  $\omega_0 \in \Omega$  is the true state. Assume our agent is Bayesian, that is that he conforms himself to expected utility theory (Savage, 1972). When asked what the single most likely state is, our agent answers the question by maximising  $b$  in the following function:

$$\mathbb{E}(u_i(b, \omega)) = \int_{\omega \in \Omega} c(\omega) u_i(b, \omega) d\omega$$

In other words, an agent purely concerned with finding the truth will maximise the expected utility of the scoring rule  $u_i$ , given his credence, in order to report which single state is, for him, the most likely.  $u_i$  is thus called a *proper scoring rule* as it maximises the expected epistemic utility function by honestly announcing his belief  $b$  instead of an alternative one  $b'$ .

Now, let's go back to our speaker/hearer problem. Assume both are epistemic agents. The hearer is a purely altruistic epistemic agent, concerned only about the truth in the manner we

---

<sup>6</sup>Zollman considers only binary states situations. A natural and important extension is to any finite number of states or to an infinite one. Here I consider the latter case.



---

presented up to now. He holds credence  $c$  which leads him to belief  $b$ . The speaker is an expert in some subject, and knows the true state. She aims to inform the hearer about it and cares about his belief  $b$ . Yet, she is not purely altruistic in the sense that her expected epistemic utility function is maximised when  $b = \omega + m$ , where  $m \neq 0$  is called the misalignment factor. For simplicity, let us assume her utility function is the following:

$$u_S(b, \omega_0) = -(b - \omega_0 - m)^2$$

where  $\omega_0 \in \Omega$  is the true state.

### 3.2 Nash equilibrium

Consider a case where agents have an epistemic conflict of interest and thus that  $m \neq 0$ . What will happen when such a speaker and hearer interact? In particular, is the speaker's claim a strong enough justification for the hearer to believe it?

Of course, it might be that, whatever she reports, our speaker can provide some external proof that she is not lying to the hearer. For instance, researchers can provide statistical evidence or mathematical proofs supporting their claims, and if the hearer is able to understand them, he can check the results by himself. But then we have left the domain of anti-reductionist theories: ultimately the justification relies on logic and not on testimony. The interesting case for us is the one where justification relies only on the speaker's word.

So let us consider a case where the speaker reports about the state to the hearer, with no other argument in favour of his claim. If both agents are only maximising their utility function, I claim that the hearer will never be entitled to the same belief as the speaker. For the sake of the example let us assume this misalignment is due to different validation standards (example 2 above). The hearer is aware of the speaker's validation standards and thus of the misalignment of  $m$ . So, if the latter decides to report sincerely and state  $q_\omega$ , the hearer will believe that, given that the speaker is maximising her utility and would thus want to induce a belief which is  $m$  away from the true belief,  $q_{\omega+m}$  is true. But then, the speaker could anticipate the reaction of the hearer and, in order to maximise her epistemic

---

utility function, report  $q_{\omega-m}$ . First, note that in doing so, the speaker gives up on sincerity in the hope of inducing the right belief, given her standards. Will it work? Most probably not. Following the same logic, by hearing  $q_{\omega-m}$  the hearer can anticipate the behaviour of the speaker and consider that  $q_{\omega+m}$  might be true, or even  $q_{\omega+2m}$ . In fact, once speaker and hearer are trapped in this corrective logic, it is impossible to deduce from the speaker's word what she really knows.

Let us call this reasoning a case of *infinite correction*. By only assuming that agents are maximising their utility, situations of epistemic conflict of interest such as I have described lead to dead ends, such as *infinite correction*. In other words, the speaker has an incentive to lie and the hearer knows about it. The latter is thus clearly not justified to hold beliefs based on the former's claims. Should we reject justification through testimony under epistemic conflict of interest in general, as argued until now in the literature (section 2)?

I believe we should not. One way of exiting this dead end is to assume more regarding both agents' behaviour. Namely, I will assume their choice of strategies in the game I consider conforms to the logic of the Nash equilibrium. In the context of our problem, the Nash equilibrium translates as follows<sup>7</sup>:

**Nash equilibrium:** The hearer is entitled to believe a proposition conveyed by the speaker if, *given the resulting belief of the hearer*, the speaker has no interest in conveying another proposition.

Let us unpack this proposition carefully. What matters is that an agent will always decide *given the reaction* he can anticipate from the other. This is exactly why the speaker chooses not to say  $q_{\omega}$  but  $q_{\omega-m}$  and so on. By choosing *precise* claims such as  $q_{\omega}$ , anticipating the other's reaction is what led to the unstable regression presented above. Yet if the speaker had no interest in entering the *infinite correction* logic in the first place, the Nash equilibrium principle tells us that her claim will be justified in the speaker's ears. It would naturally follow that the hearer would be justified to believe the speaker's claim. Is this possible in the examples we are considering? The point we made in our example shows that it is not when the speaker restricts herself to propositions of the type  $q_{\omega}$ , with  $\omega \in \Omega$ . Precise claims are always *out of equilibrium* claims in our game and fail to lead to belief justification. Yet, in the

---

<sup>7</sup>A formal definition can be found in Kreps (1990)

---

following, I will show that the hearer can possess justification for his belief for a more general class of propositions. But before moving on it seems important to address an fundamental concern: why do we need to introduce the Nash equilibrium into the picture? Are we justified to do so?

In most cases, one cannot predict the outcome of a game without assuming more about the joint behaviour of players—as the infinite regression case illustrates. Although the Nash equilibrium is the most popular one among game theorists and economists, many others exist. Fourny (2020) offers a comprehensive overview and discusses their respective normative appeal. In particular, the Nash equilibrium has often been considered as capturing a very selfish conception of social interactions and its normative attractiveness has been challenged. I would not go as far as saying that our agents ought to interact following its logic. However, in many situations, the Nash equilibrium has a strong positive appeal. In particular, in the context of the class of games we are interested in here (generally called cheap-talk games in the game-theory literature), experimental studies such as Battaglini and Makarov (2014) have given evidence in favour of the Nashian logic. It seems that, in practice, subjects, when confronted with situations similar to the ones I described above, behave according to the predictions of the Nash equilibrium. So, assuming this equilibrium concept in order to be able to predict the outcome of the game we consider has positive ground for its support. The claim about rationality this paper aims to make is not about how agents should behave when facing strategic situations. I take this aspect as given by empirical evidence. The normative aim of this paper is to establish what agents are justified to believe, when their only source of justification is testimony.

Relying on the Nash equilibrium as a warrant for justified belief is shifting the weight of justification from potential evidences the hearer can obtain through perception of logic to an equilibrium behaviour. The latter puts some behavioural structure on the strategic setting both speaker and hearer find themselves in.

### 3.3 Bayesian updating at equilibrium

Until now we have restricted ourselves to the class of proposition  $p_\omega$ , with  $\omega \in \Omega$ . As we saw with the *infinite correction* case, as long as there is some misalignment, even if the speaker

---

knows the state, in our problem, the hearer is never justified in believing  $p_\omega$ . But what if we could consider any class of proposition of the form  $p_S$ , where  $S$  is any subset of  $\Omega$ ?

For the speaker to be able to state some proposition  $p_S$  as being part of a Nash equilibrium, it must be that, given the resulting belief of the hearer, the speaker has no interest in conveying another proposition than  $p_S$ .  $p_S$  would then be called an equilibrium message. But we still need to specify one element in our definition. Given an equilibrium message  $p_S$ , what should be the resulting credence of the hearer? Given that our hearer is Bayesian, I will make the natural assumption that he will update  $c$ , his credence function over  $\Omega$ , following Bayes' rule. That is: being justified to believe that the state is in  $S$ , he will simply re-scale his credence from  $\Omega$  to  $S$ . Formally, this is:

$$c(\omega|p_S) = \begin{cases} \frac{c(\omega)}{\int_{\omega \in S} c(\omega) d\omega} & \text{if } \omega \in S \\ 0 & \text{if not} \end{cases}$$

The hearer then maximises his expected utility function  $\int_{\omega \in \Omega} c(\omega|p_S) u_H(b, \omega) d\omega$  in function of  $b$ . On the contrary, if the hearer is told a proposition  $p'$  which is not an equilibrium message (as in the *infinite correction* example), it seems fair to assume that he will ignore it and stick to his credence<sup>8</sup>.

Notice that a consequence of assuming that both parties follow the Nash equilibrium logic is that once they reach equilibrium, strategies are known to all. This means in particular that an equilibrium claim of the speaker can not be deceiving, in the sense that  $\omega_0$  will always be in the support of  $c(\omega|p_S)$ . However, hearer and speaker can end up with different beliefs, given that the former one's if only  $\operatorname{argmax}_{b \in \Omega} \int_{\omega \in \Omega} c(\omega|p_S) u_H(b, \omega) d\omega$

Now that all elements necessary to express my claim have been defined, let me explicitly formulate the version of anti-reductionism I support in this paper.

### **Equilibrium Anti-Reductionism:**

---

<sup>8</sup>Combined with the Nash equilibrium assumption, these two updating assumptions lead to a natural refinement of the Nash equilibrium in the context of uncertainty we will adopt here: the Perfect Bayesian equilibrium, first proposed by Fudenberg and Tirole (1991).

---

*"A hearer is entitled to the belief  $b \in \Omega$  induced by a speaker's testimony alone if and only if it follows from an equilibrium communication process."*

It is important to note that Equilibrium Anti-Reductionism does not posit that the hearer is entitled by default to believe the *propositional* content of the speaker's claim, but only the induced belief acquired at equilibrium. Thus, Equilibrium Anti-Reductionism relies heavily on the three premises I defended in this section:

1. The definition of belief and its connection with credence I assumed in 3.1
2. The Nash equilibrium logic and the underlying behavioural beliefs I described in 3.2
3. Bayesian updating I described in the current sub-section

In the following I will explain why these assumptions lead to Equilibrium Anti-Reductionism.

## 4 Justification at equilibrium

A game with the same mathematical structure, but with a different interpretation from the one I described up to now has been studied in the game theory literature in Crawford and Sobel (1982). Their first result, Lemma 1, is of great interest for our discussion. I will rephrase it here:

**Lemma 1.** (*Crawford and Sobel (1982)*) *The game has a finite number of equilibria, all of which are partitional.*<sup>9</sup>

What is a partitional equilibrium? Figure 1 illustrates the structure of such an equilibrium. In this example, if the true state is in  $[0, \omega_1]$ , the speaker claims message  $p_{[0, \omega_1]}$  while if the state is in  $[\omega_1, 1]$ , she sends message  $p_{[\omega_1, 1]}$ .

Notice that what Lemma 1 implies is that there is partition of  $\Omega$  in a finite number of intervals and that at equilibrium, all elements of that interval would send the same message.

---

<sup>9</sup>The precise result is slightly more technical. There are in fact an infinite number of equilibria but all are outcome equivalent to a finite number of partitional ones. It is usually considered that focusing on the latter class does not result in a loss of generality.

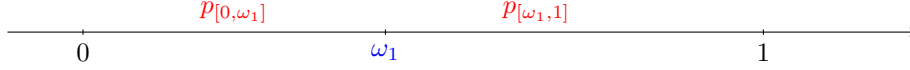


Figure 1: Example of a 2 cut-off equilibrium

It follows that at equilibrium the hearer knows exactly in what interval of the partition the hearer's belief  $\omega_0$  is. However, nothing imposes that the propositional content of equilibrium claim has to be truthful. In a two interval partition we are considering here, which is separated in  $\omega_1$ , the equilibrium claims can perfectly be "all rationals in  $[0, 1]$ " for all states below  $\omega_1$  and "all irrationals in  $[0, 1]$ " for states above. At equilibrium, the hearer will still be able to identify the interval of  $\Omega$  that made a given claim and rescale its credence on that interval.

Yet, in practice, it is without loss of generality to consider only claims whose propositional content is the equilibrium interval - here " $[0, \omega_1]$ " and " $[\omega_1, 1]$ "- as they will induce exactly the same beliefs as less natural claims such as those I presented above. Thus, if one assumes the speaker will use claims whose propositional content is the equilibrium interval, the hearer is justified in conditionalising his credence over the propositional content of the speaker's claim.

In principle, multiple equilibria can exist, each of which is characterised by a finite number of states called cut-off states— $\omega_1$ , in figure 1. For instance, there can also be a 4 cut-off equilibrium, where more propositions can be justified (see an illustration in figure 2).

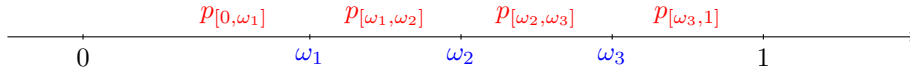


Figure 2: Example of a 4 cut-off equilibrium

In the words of game theorists, the main consequence of lemma 1 is that information transmission is always imprecise. In the words of epistemologists, this means that the hearer will never be justified through testimony alone to hold credence 1 that the state is  $\omega_0$ . For instance, in the examples above, if the speaker believes in state  $\frac{\omega_1}{3}$ , the hearer will only have the true<sup>10</sup>, Nash-justified credence that the state is in  $[0, \omega_1]$  and, if his prior credence is uniformly distributed, believes that the state is  $\frac{\omega_1}{2}$ .

There is always at least a 1-cut-off equilibrium where whatever the true state, the speaker sends the same message (and the cut-off then is  $\omega_1 = 1$ ):  $p_\Omega$ . We mentioned this proposition

<sup>10</sup>The posterior credence of the hearer is true in the sense that the true state is consider as possible. This is because under the equilibrium assumption, the speaker cannot be deceiving, as the speaker will not claim  $p_S$  if  $\omega_0$  is not in  $S$ .

---

earlier, and noticed it was fully uninformative. Why is it an equilibrium message? Because clearly the speaker cannot be lying when saying something the hearer already knew. So she could always say  $p_\Omega$  and be believed by the hearer. The more cut-offs an equilibrium has, the more precise the propositions the hearer is justified to believe.

Why are these communication strategies part of an equilibrium? Let us take the example of the 2 cut-off equilibrium of Figure 1, where  $\omega_1$  is the only cut-off in  $(0, 1)$ . Here, the speaker has the choice between stating two propositions which will be believed by the hearer:  $p_{[0, \omega_1]}$  and  $p_{[\omega_1, 1]}$ . Any other proposition will be ignored. For this to be an equilibrium, it must be that all states below  $\omega_1$  prefer to separate themselves from the states above it. That is: if the state the speaker believes in is below  $\omega_1$ , the speaker prefers to claim  $p_{[0, \omega_1]}$ , while if that state is above  $\omega_1$ , proposition  $p_{[\omega_1, 1]}$  will be the preferred claim.

If the speaker claims  $p_{[0, \omega_1]}$ , Bayes' rule will cause the credence of the hearer to re-scale over  $[0, \omega_1]$ , leading to a belief  $b_{[0, \omega_1]}$ <sup>11</sup> which is necessarily in this interval. Similarly, if  $p_{[\omega_1, 1]}$  is claimed, the belief of the hearer,  $b_{[\omega_1, 1]}$ <sup>12</sup>, will be in  $[\omega_1, 1]$ . Yet, because we assumed that epistemic utility functions are scoring rules, they possess a unique maximum and strictly increase for beliefs before that maximum and strictly decrease for beliefs after it. Therefore, it can be that if the true state is  $\omega_1$ , the speaker is exactly indifferent between the hearer believing  $b_{[0, \omega_1]}$  and  $b_{[\omega_1, 1]}$ . Figure 3 illustrates how this can be possible. A direct consequence is that if the true state is below  $\omega_1$ , the speaker strictly prefers claiming  $p_{[0, \omega_1]}$  than  $p_{[\omega_1, 1]}$  and vice versa if the true state is above  $\omega_1$ . Thus, both messages conform to the Nash principle: the speaker can do no better than using them as we prescribed.

One important thing to notice is that the existence of cut-off states is dependent on the misalignment. Assume that, in our climate scientist example, the misalignment is very important, for instance 50%. This means that whatever the effect of a 2°C increase in temperature on the current GDP, the climate scientist wants the hearer to believe it is 50 percentage points below what she really believes it to be. Recall it is assumed that this number is between 0 and 100. For there to be a cut-off state, there must be a percentage such that the speaker is indifferent between the belief induced by saying the effect on GDP is below that level and above it. Clearly, given the misalignment and as the minimal level is 0%, that level has to be above 50%. In addition, given the misalignment is so strong, whatever

---

<sup>11</sup>Formally,  $b_{[0, \omega_1]} = \operatorname{argmax}_{b \in \Omega} \int_{\omega \in \Omega} c(\omega | [0, \omega_1]) u_H(b, \omega) d\omega$

<sup>12</sup>Formally,  $b_{[\omega_1, 1]} = \operatorname{argmax}_{b \in \Omega} \int_{\omega \in \Omega} c(\omega | [\omega_1, 1]) u_H(b, \omega) d\omega$

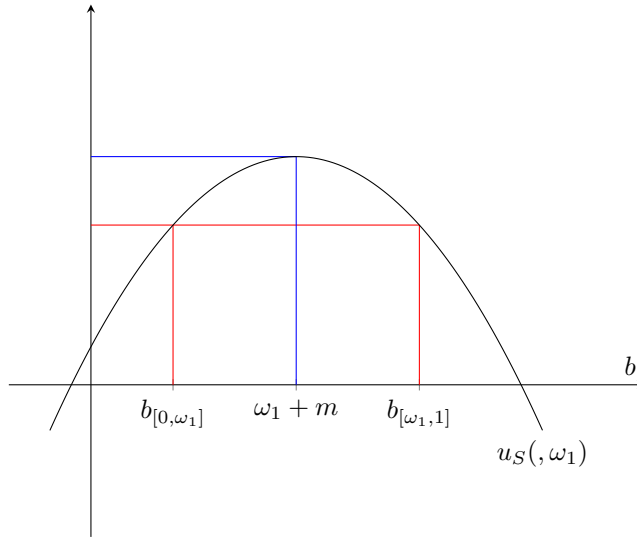


Figure 3: Identifying cut-off  $\omega_1$

the state  $\omega$  between 50 and 99, the speaker will always strictly prefer to induce a belief as low as possible. So  $p_{[\omega, 1]}$  is not an equilibrium message and there is no 2 cut-off equilibrium. The smaller the misalignment, the more cut-offs can exist and at the limit, when  $m$  is close to 0, the speaker can almost be perfectly precise.

Because the hearer is only interested in the truth, he would always like to know more about the state. Interestingly, despite misalignment, the speaker would also like to be perfectly precise about his belief<sup>13</sup>. Imagine, as a thought experiment, that before becoming an expert, the speaker could commit herself to always speak the truth about what she will learn. Then, if you average her epistemic utility over all the possible states she could learn<sup>14</sup>, she would be better off than in any other partitional equilibrium agents would end up in without commitment. In other words, the weakening of justification due to the misalignment is a strategic effect which happens despite both agents' epistemic interest. It is entirely a consequence of the environment they are in. This result goes further on, even without commitment. If the speaker could choose a communication strategy before learning the state, she would always choose the one corresponding to the equilibrium with most cut-offs. That is, she would want to be as precise as she could be, in a certain sense.

<sup>13</sup>See Crawford and Sobel (1982) section 4.

<sup>14</sup>according to credence  $c$ , assuming this is the state generating probability



---

## 5 Credence Imprecision and Belief Accuracy

The theory I proposed offers a game theoretical foundation for testimony-based credence justification. Given the equilibrium claim  $p_S$  made by the speaker, the hearer is justified in adopting the credence  $c(|p_S)$ . In addition, given what I assumed regarding the connection between credence and beliefs, it also follows that the hearer is justified in believing  $b_S$ , where

$$b_S = \operatorname{argmax}_{b \in \Omega} \int_{\omega \in \Omega} c(\omega|p_S) u_H(b, \omega) d\omega$$

Note that this specific connection between credence and beliefs it is not necessary to support my claim. As my theory provides a justification for testimony-based credences, it is sufficient to assume that beliefs can be reduced to credences, as done for instance in Sturgeon (2008), Foley (2009) or Leitgeb (2013), to further obtain a justification for testimony-based beliefs.

However, the precision of the credence obtained through equilibrium communication plays an important part in the accuracy of the resulting belief of the hearer. Belief in my settings are elements of  $\Omega \subset \mathbb{R}$ . Therefore I will measure the accuracy of belief  $b \in \Omega$  through the natural euclidean distance to the true state:  $|b - \omega_0|^{15}$ . By credence precision I mean the size of the interval of states making the claim  $p_S$  at equilibrium. Formally, call  $S = [l, u]$  an interval of  $\Omega$  such that there is an equilibrium strategy  $\sigma$  where  $\sigma(\omega) = p_S$  for any  $\omega \in [l, u]$ . The size of  $S$  is  $|u - l|$ .

First, note that the speaker's posterior credence can never be perfectly accurate. This is because claims of the type  $q_\omega$ , where  $\omega \in \Omega$  are always out of equilibrium claims in our game. In other words, they are not justified to be believed by the hearer, even if the speaker is telling the truth<sup>16</sup>.

Equilibrium claims are always of the form  $q_I$ , where  $I$  is an interval of  $\Omega$  containing the true  $\omega$ . The hearer is thus always justified to hold a less precise posterior credence than the speaker. But this has consequences on the accuracy of the hearer's resulting belief.

**Lemma 2.** *On average, the smaller the epistemic conflict of interest, the more accurate the*

---

<sup>15</sup>An important literature studies the existence of proper scoring rules and the notion of accuracy for credences on infinite domain (Pruss (2022), Nielsen (2023)). This problem does not apply to me though, as my beliefs are point estimates.

<sup>16</sup>A similar point has been made by Bright (2017), section 4, regarding the lack of credibility of "truth seeking scholars".

---

*belief of the hearer.*

**Proof of Lemma 2:**

Assume the equilibrium with the most cut-off in the game has  $n$  cut-offs. I will show that in that equilibrium, on average, the smaller the epistemic conflict of interest, the more accurate the belief of the hearer. The same logic applies for any other equilibrium.

Notice that whatever  $\omega_0 \in \Omega$ ,  $p_S(\omega_0)$  the corresponding equilibrium claim of the speaker which will induce a belief  $b_S$ , is a function of  $m$ . A formal rephrasing of Lemma 2 is to say that the expected distance between the optimal action for the speaker  $\omega_0$  and the equilibrium action of the hearer  $b_S$  is decreasing in  $m$ . Yet, given that the expected utility of the hearer is a proper scoring rule, the expected distance between  $\omega_0$  and  $b_S$  is decreasing in  $m$  if and only if the expected utility of the hearer is decreasing in  $m$ . It also follows from Crawford and Sobel (1982) equation (25) that the expected utility of the hearer is:

$$-\frac{1}{12} - \frac{m^2(n^2 - 1)}{3}$$

which is decreasing in  $m$ .

□

What Lemma 2 shows is that the smaller the epistemic conflict of interest, the more accurate, on average, the credence of the hearer. This is because the smaller the epistemic conflict of interest, the more precise, on average, the credence of the hearer. Thus, credence imprecision gives a weaker evidential support to the hearer's belief, leading, on average, to a less accurate belief than under reduced epistemic conflict of interest.

Notice that this result is only valid *on average*. It can be that for a given state, a less biased speaker induces a less accurate belief than a more biased speaker. It can even be that for some state, a very biased speaker induces the hearer to the true belief. But given the continuous nature of our setting, this happens with 0 probability. On average, epistemic conflict of interest leads to more imprecise (but credible) claims which themselves lead to less accurate beliefs.

---

Dang and Bright (2021) show that in practice, 20<sup>th</sup> century scientific research has often sacrificed the norm of accuracy for the sake of credibility—as I argue here. This result also has a flavour to those obtained by O’Connor (2014) regarding vagueness in predicates and in particular in natural languages. While O’Connor shows that they arise as a result of quick and successful signalling conventions, I show that they are the result of a quest for strategic compatibility for the speaker and justification for the hearer. A similar point has been made by Mayo-Wilson (2014) who builds on a formal approach to show that the degree of insularity of scientific communities has a major impact on their credibility regarding non-expert communities under epistemic conflict of interest<sup>17</sup>. Although formally different as Mayo-Wilson builds on networks and not on communication models, his result conceptually parallels the one I derived regarding the impact of misalignment on the accuracy of equilibrium belief.

## 6 Discussion

Even under epistemic conflict of interest, I offered reasons to think that the hearer can still be justified to trust the speaker and to build belief upon her words. But this can happen only in very specific situations and only for a certain class of beliefs.

**Equilibrium in practice**—Justification is only granted *at equilibrium*. Two things are to be noted regarding equilibrium behaviours.

First, acting according to the Nash equilibrium logic is the result of beliefs about other players behaviour: all players’ must believe that other players will act as such<sup>18</sup>. Then, these *behavioural* beliefs will act as a self-fulfilling prophecy, leading the speaker’s claim to conform to the Nashian logic, and thus justifying the hearer to hold *propositional* beliefs such as  $q_{[\omega_1, \omega_2]}$ . As argued by Voss (2001) or Bicchieri and Sontuoso (2020) the adoption of a social norm can be seen as resulting from that kind of *behavioural* belief. The idea has notably been applied by Lewis (2008) in a sender-receiver game similar to the one studied here to suggest that the emergence of languages results from an equilibrium behaviour<sup>19</sup>. The same idea has

---

<sup>17</sup>Mayo-Wilson designates this as situations of potential miscommunication.

<sup>18</sup>More specifically, players must believe others are maximising their utility function, know the pay-offs and the strategies of all parties in the game. See Aumann and Brandenburger (1995) for a detailed account of the behavioural assumptions underlying the Nash equilibrium.

<sup>19</sup>In Lewis’ game there is no misalignment and the sender/receiver role is randomly assigned at beginning of the game. Lewis

---

been widely defended by proponents of evolutionary game-theory (Smith and Price, 1973), for whom collective norms appear in the long run as a result of an equilibrium process.

Second, it is important to note that social norms, just like equilibrium behaviour, take time to establish themselves. In a social context, agents try different speaking and hearing strategies until their joint behaviour stabilises, because no epistemic benefit can be gained by a further change. As a result, in situations which have not reached their equilibrium state, belief justification fails. This would be the case for new or unexpected events, such as the early stages of the COVID-19 pandemic, where the settings of the game are still not well understood by all players. Claims made by experts by that time still had not reached equilibrium level. This observation shows how important scientific norms and practices are for enforcing the credibility of their claims. Because scientific norms are the result of equilibrium behaviour, they provide justification to an audience of laypeople to rely on expert advice. In the same line as Zollman (2019), this result gives another reason to understand why deviation from scientific good practices, such as scientific fraud, is extremely detrimental to Science's credibility. By drawing a link between equilibrium behaviours and collective norms, this paper also aligns with Gerken (2015) and Gerken (2022) regarding the importance of scientific norms in the context of intra-scientific testimony and in expert/laypeople communication.

**Similar approaches**—To my knowledge, two similar formal approaches to the Source Problem have been proposed in the epistemological literature: Faulkner (2011) and Duijf (2021).

While, similarly to the analysis defended in the present paper, Faulkner agrees that the SP is essentially a problem of cooperation (or a game-theoretical problem), he concludes against AR:

*[S]peakers and audiences have different interests in communication[...] Our interest, qua audience, is learning the truth. Engaging in conversations as to the facts is to our advantage as speakers because it is a means of influencing others: through an audience's acceptance of what we say, we can get an audience to think, feel, and act in specific ways. So our interest, qua speaker, is being believed...because we have a more basic interest in influencing others...[T]he commitment to telling the truth*

---

also shows that pooling equilibria—of which partitional equilibria are a special case—can emerge.

---

*would not be best for the speaker. The best outcome for a speaker would be to receive an audience's trust and yet have the liberty to tell the truth or not. (2011, 5–6)*

The main difference between Faulkner's approach and mine is that Faulkner does not assume any equilibrium concept. Therefore, his reasoning essentially takes place out of equilibrium. It is thus natural that we agree on the fact that the hearer is not justified in believing  $q_\omega$  type claims. But assuming behavioural beliefs and thus looking at the Nash equilibrium of the game enables me to grant justification for certain classes of beliefs.

A second reason for these differences is that Faulkner's premises are slightly different from mine. Although we agree that the default position for the speaker is to promote her own interest, Faulkner also considers that her interest is to be believed. This latter assumption is at odds with the equilibrium logic. In the long run, if being believed was the interest of the speaker regardless of anything else, she could do no better than simply saying the truth. In addition it seems to me that, in situations such as those I've listed above, the speaker's interest goes beyond the mere aim of being believed.

Another paper close to mine is Duijf (2021). Duijf considers an expert/layperson communication model where the former informs the latter regarding the consequences of a given decision. The expert has varying degrees of expertise (captured by the probability that he knows the state) and there are varying degrees of misalignment between parties (captured by a commonly known probability). Duijf examines the delegation question: when is it rational for the layperson to delegate an action to the expert and when is it not? Although an important question, it is distinct from the one I'm interested in in the current paper: when is it rational for the hearer to hold a proposition conveyed by a speaker for knowledge? Interestingly though, our results have a similar flavour: it must be that the misalignment between speaker and hearer is not too wide<sup>20</sup>.

An alternative approach to the problem of expert reliability builds upon the concept of reputation. The speaker is justified to believe the hearer in virtue of the former's reputation, or, in other words, in virtue of the repeated observation that the speaker is saying the truth. In the words of game theorists, one should consider these situations as repeated games and not *one-shot* ones—as I did in the present paper. While I agree that, in many situations, this

---

<sup>20</sup>Duijf's allowance for imperfectly informed speakers would, in my model, have no qualitative impact on the results: equilibria would still be partitional and the hearer would still know less than the speaker.

---

argument is compelling to insure the justification for knowledge, it deals with situations which are beyond the anti-reductionist realm. Repeated observations of the conjunction between the speaker's claims and reality cannot be made without perception or logic. I believe that anti-reductionist is precisely relevant for cases where reputation cannot play a part and where the decision to trust has to be done in *one-shot*. These cases can be of prime importance: how can one assess the reliability of climate scientists regarding (b) when it comes to describing a +5°C planet? Cases of misalignment can exist, such as differences in validation standards (Lloyd et al., 2021), moral disagreements (Gundersen, 2020) or publication incentives. How can one assess the reliability of epidemiologists regarding (b), when assessing the long-term effects of mRNA-based vaccines? Similar misalignment incentives can easily be imagined.

## 7 Conclusion

Can a hearer be rationally justified to have beliefs based on testimony alone when there is a known epistemic conflict of interest with the speaker? While this question has been mainly discarded by the literature, I have argued that it is essential for many practical situations. In particular, being able to provide a theory that supports a positive answer to it is essential to support modern production of scientific knowledge. On the basis of a game-theoretical approach, I have shown that epistemic dependence under epistemic conflict of interest is possible, but only *at equilibrium*. Equilibrium behaviours appear in the long run, and are embodied by scientific norms. This result contributes to highlighting the—already stressed—importance of scientific norms for Science's credibility. Moreover, even at equilibrium, the belief the hearer will be justified to hold will be on average *less accurate*, than the speaker's original belief.

## References

- Aumann, R. and Brandenburger, A. (1995). Epistemic conditions for nash equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1161–1180.
- Battaglini, M. and Makarov, U. (2014). Cheap talk with multiple audiences: An experimental analysis. *Games and Economic Behavior*, 83:147–164.

- 
- Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.
- Bicchieri, C. and Sontuoso, A. (2020). Game-theoretic accounts of social norms: The role of normative expectations. In *Handbook of experimental game theory*, pages 241–255. Edward Elgar Publishing.
- Bradley, R. (2017). *Decision theory with a human face*. Cambridge University Press.
- Bright, L. K. (2017). On fraud. *Philosophical Studies*, 174(2):291–310.
- Burge, T. (1993). Content preservation. *The Philosophical Review*, 102(4):457–488.
- Burge, T. (1997). Interlocution, perception, and memory. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 86(1):21–47.
- Burge, T. (2013). *Cognition Through Understanding: Self-Knowledge, Interlocution, Reasoning, Reflection: Philosophical Essays*, volume 3. Oxford University Press.
- Chisholm, R. M. (1989). *Theory of knowledge*, volume 3. Prentice-Hall.
- Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, pages 1431–1451.
- Dang, H. and Bright, L. K. (2021). Scientific conclusions need not be accurate, justified, or believed by their authors. *Synthese*, 199(3):8187–8203.
- Duijf, H. (2021). Should one trust experts? *Synthese*, 199(3):9289–9312.
- Faulkner, P. (2007). On telling and trusting. *Mind*, 116(464):875–902.
- Faulkner, P. (2011). *Knowledge on trust*. Oxford University Press.
- Fehr, E., Fischbacher, U., and Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature*, 13(1):1–25.
- Foley, R. (2009). Beliefs, degrees of belief, and the lockean thesis. In *Degrees of belief*, pages 37–47. Springer.
- Fourny, G. (2020). Perfect prediction in normal form: Superrational thinking extended to non-symmetric games. *Journal of Mathematical Psychology*, 96:102332.
- Fricke, M. (2012). Group testimony? the making of a collective good informant. *Philosophy and Phenomenological Research*, 84(2):249–276.

- 
- Fudenberg, D. and Tirole, J. (1991). Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53(2):236–260.
- Gerken, M. (2015). The epistemic norms of intra-scientific testimony. *Philosophy of the Social Sciences*, 45(6):568–595.
- Gerken, M. (2022). *Scientific Testimony: Its roles in science and society*. Oxford University Press.
- Goldberg, S. C. (2010). *Relying on others: An essay in epistemology*. Oxford University Press.
- Goldberg, S. C. (2014). Interpersonal epistemic entitlements. *Philosophical Issues*, 24(1):159–183.
- Graham, P. J. (2010). Testimonial entitlement and the function of comprehension. *Social Epistemology*.
- Graham, P. J. (2015). Epistemic normativity and social norms. *Epistemic Evolution: Purposeful Epistemology*.
- Gundersen, T. (2020). Value-free yet policy-relevant? the normative views of climate scientists and their bearing on philosophy. *Perspectives on Science*, 28(1):89–118.
- Hardwig, J. (1985). Epistemic dependence. *The Journal of Philosophy*, 82(7):335–349.
- Henderson, D. (2020). Are epistemic norms fundamentally social norms? *Episteme*, 17(3):281–300.
- Henrich, J. and Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, 72:207–240.
- Hinchman, E. S. (2005). Telling as inviting to trust. *Philosophy and Phenomenological Research*, 70(3):562–587.
- Hinchman, E. S. (2014). Assurance and warrant. *Philosopher’s Imprint*, 14(17).
- Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science*, 65(4):575–603.
- Kreps, D. M. (1990). *Game theory and economic modelling*. Oxford University Press.
- Leitgeb, H. (2013). Reducing belief simpliciter to degrees of belief. *Annals of Pure and Applied Logic*, 164(12):1338–1389.



- 
- Lewis, D. (2008). *Convention: A philosophical study*. John Wiley & Sons.
- Lloyd, E. A., Oreskes, N., Seneviratne, S. I., and Larson, E. J. (2021). Climate scientists set the bar of proof too high. *Climatic Change*, 165(3):1–10.
- Luce, R. D. and Raiffa, H. (1989). *Games and decisions: Introduction and critical survey*. Courier Corporation.
- Mayo-Wilson, C. (2014). Reliability of testimonial norms in scientific communities. *Synthese*, 191(1):55–78.
- McMyler, B. (2011). *Testimony, trust, and authority*. Oxford University Press.
- Meinshausen, M., Meinshausen, N., Hare, W., Raper, S. C., Frieler, K., Knutti, R., Frame, D. J., and Allen, M. R. (2009). Greenhouse-gas emission targets for limiting global warming to 2 degrees. *Nature*, 458(7242):1158–1162.
- Moran, R. (2006). Getting told and being believed. *The epistemology of testimony*, pages 272–306.
- Moran, R. (2018). *The exchange of words: Speech, testimony, and intersubjectivity*. Oxford University Press.
- Müller, B. (2022). Epistemic dependence, cognitive irrationality, and epistemic conflicts of interests: Why there is a need for social epistemic norms. *Logos & Episteme*, 13(3):287–313.
- Nielsen, M. (2023). Accuracy and probabilism in infinite domains. *Mind*, 132(526):402–427.
- O’Connor, C. (2014). The evolution of vagueness. *Erkenntnis*, 79(4):707–727.
- Pettigrew, R. (2016). *Accuracy and the Laws of Credence*. Oxford University Press.
- Poundstone, W. (1993). *Prisoner’s Dilemma/John Von Neumann, game theory and the puzzle of the bomb*. Anchor.
- Pruss, A. R. (2022). Accuracy, probabilism and bayesian update in infinite domains. *Synthese*, 200(6):444.
- Ross, A. (1986). Why do we believe what we are told? *Ratio*, (1).
- Savage, L. J. (1972). *The Foundations of Statistics*. Courier Corporation.
- Simion, M. (2021). Testimonial contractarianism: A knowledge-first social epistemology. *Noûs*, 55(4):891–916.

- 
- Smith, J. and Price, G. R. (1973). The logic of animal conflict. *Nature*, 246(5427):15–18.
- Sturgeon, S. (2008). Reason and the grain of belief. *Noûs*, 42(1):139–165.
- Tanveer, S., Rowhani-Farid, A., Hong, K., Jefferson, T., and Doshi, P. (2022). Transparency of covid-19 vaccine trials: decisions without data. *BMJ Evidence-Based Medicine*, 27(4):199–205.
- Voss, T. (2001). Game-theoretical perspectives on the emergence of social norms. *American Political Science Review*, 91(2):245–63.
- Weatherson, B. (2016). Games, beliefs and credences. *Philosophy and Phenomenological Research*, 92(2):209–236.
- Zagzebski, L. T. (2012). *Epistemic authority: A theory of trust, authority, and autonomy in belief*. Oxford University Press.
- Zollman, K. J. (2019). The scientific ponzi scheme.
- Zollman, K. J. (2021). The theory of games as a tool for the social epistemologist. *Philosophical Studies*, 178(4):1381–1401.