Title
The Epistemic Weight of Silence

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Young, David J.
Madsen, Jens Koed
Yousefi, Sayeh

Publication Date
2023

Peer reviewed

The Epistemic Weight of Silence

David Jack Young (dy286@cam.ac.uk)
Department of Psychology,
University of Cambridge, Downing Place, CB2 3EB, Cambridge, UK

Jens Koed Madsen (j.madsen2@lse.ac.uk)
Department of Psychological and Behavioural Science,
London School of Economics and Political Science, Houghton Street, WC2A 2AE, London, UK

Sayeh Yousefi (s.yousefi@lse.ac.uk)
Department of Psychological and Behavioural Science,
London School of Economics and Political Science, Houghton Street, WC2A 2AE, London, UK

Abstract

When people "fail to deny" unflattering claims, it is commonly taken to imply they are true. Yet, the 'argument from ignorance' – arguing in favour of something due to a lack of evidence against it – is often deemed a fallacy. Why is abduction from missing evidence permissible in some cases, but not others? We present a framework of factors which disambiguate these cases, using a Bayesian Network model. We suggest that a source's silence often reflects a latent conflict between their motives for what they want their audience to believe, and complying with external constraints on their speech, like the need to be accurate. In these cases, silence implies that the source does not believe that what they would like to say is true, which licenses a probabilistic inference that it is false. We present data from two studies suggesting people infer from silence like this.

Keywords: Argument from ignorance, Bayesian network, argumentation, epistemology, silence.

Introduction

The argument from ignorance – arguing in favor of a claim by appealing to the lack of evidence against it – is often described as a fallacy. Famously, Bertrand Russell asked whether we should believe that there is a China teapot orbiting the Sun, too small to detect from Earth, on the basis that there is no evidence to disprove it (Russell, 1952). The answer is an obvious "no". Elsewhere, arguments from ignorance are intuitive: if a person suspects their partner of having an affair, and upon confronting them about it, the partner says nothing, the fact they *fail to deny* the accusation is damning, and it seems reasonable to believe they are having an affair. What, then, distinguishes cases where the argument from ignorance is reasonable and not?

In line with Harris, Corner, and Hahn (2013), we adopt a Bayesian perspective to the argument from ignorance. Bayes' theorem provides a normative model for how people should revise their beliefs in the face of new evidence, expressed here in odds form, where H is the hypothesis being considered and e the evidence which bears upon it:

$$Bayes'Rule\colon \frac{p(H|e)}{p(\neg H|e)} = \frac{p(H)}{p(\neg H)} \times \frac{p(e|H)}{p(e|\neg H)}$$

Here, $p(H|e)/p(\neg H/e)$ is the "posterior odds" – the odds we assign to H being true rather than false after observing the evidence – and $p(H)/p(\neg H)$ is the "prior odds" – the odds we had *before* observing the evidence. $p(e|H)/p(e|\neg H)$ is the "likelihood ratio" – how much more likely it is that we would observe this evidence if the hypothesis were true than if it were false. Likelihood ratios above 1 make us more confident H is true by increasing our posterior odds relative to the prior odds, whereas ratios below 1 make us more confident it is false, and ratios of 1 mean we do not 'update' our beliefs whatsoever.

Bayesian approaches have successfully been applied to reasoning (Hayes et al., 2019; Oaksford & Chater, 2007) and argumentation (Hahn & Oaksford, 2006, 2007). Of particular use, they have also described reasoning fallacies and errors from a rational perspective, including arguments from ignorance (Hahn, Oaksford, & Bayindir, 2005; Oaksford & Hahn, 2004), ad hominem (Harris et al., 2012; Oaksford & Hahn, 2012), slippery slope (Corner et al., 2011), and circular arguments (Hahn, Oaksford, & Corner, 2005).

Harris et al. (2013) argued that a Bayesian approach can be used to model arguments from silence. They show that a lack of evidence, n, can be thought of as having a likelihood ratio $p(n|H)/p(n|\neg H)$ with respect to a given hypothesis, H. Therefore, belief updating in response to a lack of evidence is warranted when $p(n|H) \neq p(n|\neg H)$, which occurs when the lack of evidence is systematically related to the truth of H.

In many cases where the argument from ignorance is considered a fallacy, there is no systematicity – the lack of evidence is uncorrelated with the truth of the hypothesis at hand. For instance, in the case of Russell's celestial teapot, we on Earth lack *epistemic closure* – we have not been able to perform a complete search for the relevant information because the evidence for our claim is undetectable. This means that whether there is a teapot or not, we will always lack evidence. In this case $p(n|H) = p(n|\neg H)$, and the lack of evidence provides no justification for updating our beliefs (Oaksford & Hahn, 2004).

But in other cases, there is systematicity. In the case of a person accusing their partner of infidelity, there is a systematic relationship between the partner's silence and the

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

hypothesis – if we believe the partner has at least *some* reservations about lying, they should be more likely to deny the accusation if it is false than if it is true. Their failure to deny it implies they have reservations about doing so, which implies that denying it would be lying, and therefore it is likely true. In this case, $p(n|H) > p(n|\neg H)$, and their silence can be treated as evidence in favor of the claim.

Although work has been conducted showing that people do update their beliefs when confronted with a lack of testimonial evidence for claims, in proportion with predicted $p(n|H)/p(n|\neg H)$ ratios (Harris et al., 2013), and some of the circumstances under which doing so can be normative have been explored (Hahn & Oaksford, 2006; Goldberg, 2011), we lack a framework for how people can construct these ratios. What factors do people perceive to govern the silence of sources, and do they integrate them normatively?

We suggest that a source's silence is often a result of an internal conflict between testimonial motives. Sources are often *biased* towards making claims in support of the various goals they have regarding what they would like other people to believe, such as "I am a faithful partner". But sources also often want to maintain *honesty* in what they say, due to both internal factors like a felt moral obligation to be honest, and external factors like the penalties others may impose upon us for dishonesty such as revulsion, reprisal, loss of trust, and even fines or imprisonment. These motives clash when we want others to believe something untrue, meaning there is a higher probability the source will remain silent out of a reluctance to lie. Ergo, an observer can infer a latent conflict between honesty and bias motives when a source remains silent, and thereby make an inference as to what the source believes is true, which provides evidence for the hypothesis at hand.

The Model

We have developed a Bayesian Network model of how inferences from silence can arise from considering latent conflicts between testimonial motives. The idea of the model is that, when people observe a situation where a source remains silent when questioned about a particular hypothesis, they can infer a posterior degree of belief in that hypothesis by supplying the model with appropriate priors, then performing Bayesian updating. A diagram of the model as a directed acylic graph ("DAG") is shown in Figure 1.

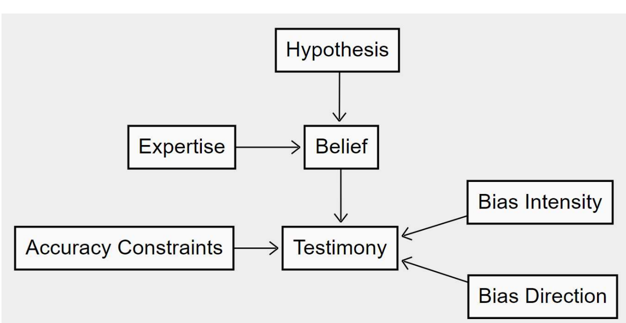

Figure 1. DAG of the Bayesian Network

The model assumes that people's biases are two-dimensional, possessing both a direction (whether they are biased towards or against the hypothesis) and an intensity (the extent to which they prioritise making claims that align with that direction over honesty). Direction is best thought of as a 'potential direction', as when the source's bias intensity is low, the direction has no effect on their speech. We separate these factors because it is common to know a person's bias direction but not their intensity – i.e., all Labour politicians are biased towards the Labour party, but some more so than others – and the two-dimensional structure allows for the common-sense prediction that when people have no bias (i.e., they are impartial), they just say what they believe is true, which is difficult to implement otherwise.

The model contends that bias intensity and direction interact with the source's perception of the accuracy constraints they are operating under (how important it is for their speech to be accurate) in order to determine what testimony they give, if any (we do not want, at this stage, to differentiate too sharply between 'accuracy' and 'honesty' constraints – the demands on a source to be accurate are identical to those of being honest unless they believe the enforcers of the accuracy constraints have false beliefs, which is more complex than the scenarios used in the studies we discuss here). As Table 1 shows, the model assumes that people's testimony will always reflect their genuine belief in two cases: 1. When that belief aligns with their bias (i.e., they are "For" the hypothesis and believe it is true, or are "Against" the hypothesis and believe it is false), and 2. When they have no bias at all (Bias Int. = "Low"). When their bias *conflicts* with their belief, they lie if accuracy constraints are low, but say nothing if they are high.

Table 1: CPT of the Expected Testimonies

Bias Int.	Bias Dir.	Acc. Const.	Belief		
			True	False	None
High	For	High	*"True"*	*None*	*None*
High	For	Low	*"True"*	*"True"*	*"True"*
High	Against	High	*None*	*"False"*	*None*
High	Against	Low	*"False"*	*"False"*	*"False"*
Low	For/ Against	High/ Low	*"True"*	*"False*	*None*

Table 2: CPT of the Expected Beliefs

Hypothesis	Expertise	
	High	Low
True	*"True"*	*None*
False	*"False"*	*None*

Thinking about perceived accuracy constraints and the bias factors allows a listener to probabilistically infer a source's *belief* - but whether that belief reflects reality depends upon the source's *expertise*. We explicitly allow sources to have no belief about a hypothesis when they completely lack expertise, as this produces the common-sense result that one explanation for a source's silence on a particular issue is that they simply know nothing about it.

Each expected belief and testimony in the CPTs is modelled to occur with a likelihood of p=1. We conducted simulations to ascertain the predictions of the model across different conditions. The results are shown in Figure 2. To generate these predictions, we allowed the priors for accuracy constraints, bias direction, and expertise variables to vary from 0.2 to 0.8 in increments of 0.2, and generated posteriors in every permutation of conditions, holding bias intensity constant at 0.8, and the hypothesis at 0.5 – these constants reflect that in the studies that follow, the sources are always biased one way or another, and the hypotheses always uncertain.

As Figure 2 shows, people infer that hypotheses are likely true (> 0.5) if a silent source should be biased against them (Bias Direction < 0.5), but likely false (< 0.5) if they are biased towards them (Bias Direction > 0.5). These inferences are weakened if the source has lower expertise, since the silence is 'explained away' by their not knowing anything, rather than a latent motive conflict. The inferences are strengthened by higher accuracy constraints, as these make latent motive conflicts more of a possibility.

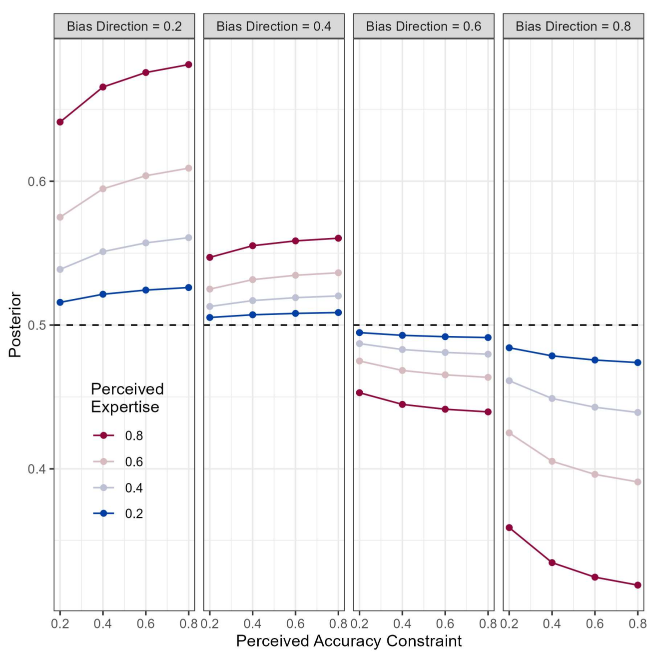

Figure 2. Predictions of the model across conditions.

Study 1

In Study 1 we set out to test whether people's inferences from silence reflected the predictions of the Bayesian Network model. In the main part of Study 1, we obtained estimates of people's posterior degree of belief in an hypothesis after learning of a source's silence in response to a question about it across different scenarios. In a follow-up, we obtained people's priors for the variables in the Bayesian Network across these scenarios, then used these to generate predictions about what the main participants' posteriors should have been. We then compared the predicted and observed posteriors.

Participants

For the main study, 351 completed the study, but 26 were excluded – one for completing twice, and 25 for failing at least one attention check (see below), leaving N=325. Participants were recruited via Prolific and paid £9/hour for a median study time of 5 minutes. All participants were British with a median age of 41 (*Range*: 20 – 80); 11 provided no gender, but of those that did, 160 were female, 153 male, and 1 non-binary.

For the follow-up, 50 completed the study, none of whom failed the attention check or completed twice, but one was excluded for also having completed Study 1A, leaving N=49. Participants were recruited via Prolific and paid £7.50/hour for a median study time of 5 minutes. All participants were British with a median age of 33 (*Range*: 21 – 69); 25 were female and 24 male.

Design and Procedure

For the main study, participants read eight dialogues describing a scenario in which a biased source was silent when asked a question. There was a 2x2x2 within-subjects manipulation of the source's bias direction, expertise, and accuracy constraints. After each dialogue, participants provided posterior probability estimates of the degree of belief someone who heard about the scenario would have had about the hypothesis entailed by the question being true.

The dialogues concerned two fictional characters – Alice and Bob – discussing politics. We told participants their government – which was not the US's or UK's, and about whose "ideology or competence" they should make no assumptions – had passed a series of tax reforms last year. They were told Alice and Bob had seen a news item explaining that an expert report on the impact of those reforms had just been delivered to the government and was being "shared privately among the government's most senior economic officials". In each of the eight dialogues, Alice initially said she had "no idea" whether the tax reforms had been successful or detrimental to the nation's wealth, then Bob told Alice about an additional news clip in which a politician ignored a question about the report's findings, which Alice said she hadn't seen. We varied the context of the politician's silence to effectuate treatment.

In the high expertise condition, the politician was described as "the minister in charge of the Government's finances"; in the low expertise condition, they were described as "a really junior Government MP who doesn't work in the Government's finance department". In the high accuracy constraints condition, the question was put to them "in Parliament" (before seeing the first dialogue, we told participants it was illegal to lie in Parliament in this country, as it is in the UK); in the low accuracy constraints condition, it was "by a journalist as they were walking down the street".

To vary the source's *bias*, we varied the question that was asked: in the biased-for condition, they were asked "Is it true that the Government's new tax policy has improved the nation's wealth"; in the biased-against condition, they were asked whether it had "…worsened the nation's debt?". In each dialogue, Bob said the politician "said nothing, just ignored the question", then asked Alice whether she had seen it, to which she replied she had not.

Participants were then asked "Now please rate what Anne's belief should be about whether the tax reforms have improved the nation's wealth or worsened the nation's debt, and how confident she should be" on a 0-100 scale, where 100 was labelled "Completely certain the tax reforms have **improved the nation's wealth**", 0 was labelled "Completely certain the tax reforms have worsened the nation's debt", and 50 labelled "Completely unsure". Participants answered using a slider which began at 50 and displayed the number they had chosen.

Since each participant saw all eight dialogues, we asked "Please treat each dialogue as self-contained and unrelated to the others, do not use anything you learn in one dialogue to affect your interpretation of the other dialogues". Dialogues were presented in a random order. The dialogues were intermixed with two attention check questions which asked them to give a particular numerical response on the slider.

Before taking part, participants provided informed consent, and after the blocks and attention checks, provided demographic data, were debriefed, and thanked. We did not ask participants to estimate any priors or likelihoods.

For the follow-up, we gave participants the same task instructions and information as Study 1A, prior to the dialogues, but were asked to estimate Anne's *priors* rather than posteriors. First, we asked for Anne's prior for the tax reforms having been successful (M = 46.3, SD = 9.2), then, what Anne's prior for the source's bias intensity, bias direction (where high scores = biased towards wanting people to believe the policy had been successful), expertise, and accuracy constraints would have been across the different dialogues. Rather than showing them the dialogues, we told them that Bob was about to tell Alice about what a politician had said when asked about the policy, and to imagine what Anne's priors would be if the source possessed different characteristics.

We asked participants to provide bias intensity, bias direction, and expertise priors for both the high-expertise and low-expertise source. We then asked them for the perceived accuracy constraint priors for each source under both the high-accuracy and low-accuracy conditions. Table 3 shows the mean and standard deviations of these priors.

The questions were blocked by source expertise, and blocks were presented in a random order and intermixed with an attention check that asked for a particular numerical response on the sliders. After providing consent, reading task instructions and completing the experimental blocks, participants provided demographic information, were debriefed, and thanked.

Table 3: Mean (and SD) of Follow-Up Priors for High-Expertise and Low-Expertise Source Characteristics

Source	Bias Int.	Bias Dir.	Exp.	Acc. (High)	Acc. (Low)
High	70.6	68.3	72.3	67.1	42.4
Exp.	(21.6)	(28.1)	(21.1)	(23.9)	(26.3)
Low	54.8	58.7	44.7	62.3	40.9
Exp.	(19.0)	(23.8)	(21.1)	(24.0)	(25.9)

Results

We coded the Bayesian Network in *R* using functions from the package {gRain}. We then used the network to simulate what each follow-up study participant's posterior would have been for each main study dialogue, and compared these to main study participants' corresponding posteriors. We changed priors of 100 to 99 and 0 to 1, as Bayesian Networks can produce undefined or infinite posteriors when given such priors, then divided them by 100 to produce probabilities on a 0.01-0.99 scale. For dialogues in the biased-against condition, we flipped people's priors for the hypothesis and the source direction, as these were both defined relative to the policy having been successful in the task instructions, whereas in this condition the source was asked whether the policy had *failed*. We also divided the main study participants' posteriors by 100.

Figure 3 shows the mean observed and predicted posteriors, with 95% CIs, for each dialogue condition. The dotted line shows the mean prior in each condition (53.7 or 46.3).

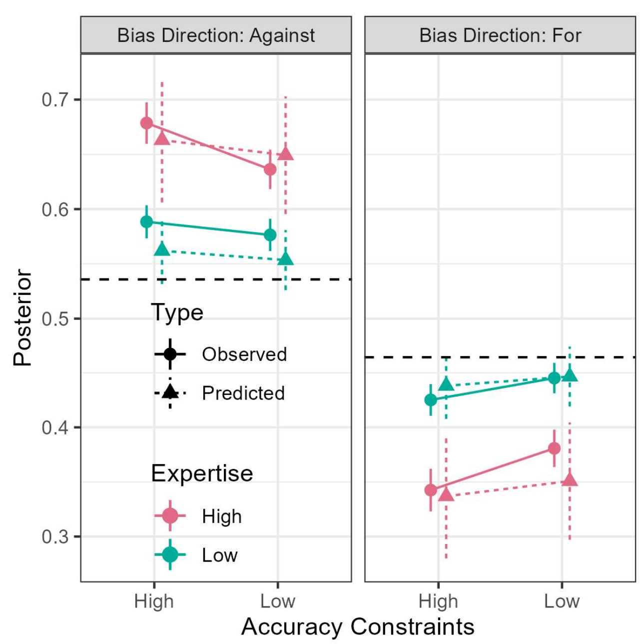

Figure 3. The Observed and Predicted Posteriors (Mean and 95% CIs) and Priors (dotted lines).

People's updating conformed to the model's predictions from Figure 1: they increased their confidence in the hypothesis being true when the silent source was biased against it, and decreased when they were biased towards it; the higher the source's expertise and accuracy constraints, the more they updated in that direction. These trends were present in both the predicted and observed posteriors.

The predicted posteriors showed a very good fit to the observed posteriors. Looking at the mean posteriors across the eight dialogue conditions, the predictions explained 98.14% of the variance in observations (adjusted R-squared), with a mean-squared error (Brier score) of 0.0003. We also merged the individual-level predicted and observed posteriors into one dataset and ran a 2x2x2x2 mixed ANOVA (Type [between]: Predicted vs Observed x Expertise x Accuracy Constraints x Bias [all within]) with Type III Sum of Squares, finding a main effect of Bias ($F(1, 372) = 146.49$, $p < .001$), an Expertise x Bias interaction ($F(1, 372) = 77.60$, $p < .001$), and an Accuracy x Bias interaction ($F(1, 372) = 12.30$, $p < .001$) and no other significant main effects or interactions. The lack of any interactions with 'Type' shows the observed trends did not significantly diverge from those predicted given people's priors.

Discussion

Study 1 provides strong evidence that our Bayesian Network model makes accurate predictions about people's inferences from silence. One potential objection is that we vary the direction of the source's bias by changing the wording of the question put to them, a subtle change which may not have been particularly salient to participants. In Study 2, we keep the question posed the same, and vary the characteristics of the source to manipulate bias direction instead.

Study 2

The data for Study 2 were collected before Study 1, but we present them second as they are more complicated. Study 2 was initially pre-registered (https://osf.io/vzm7a/) with a slightly different version of the Bayesian Network, which had one 'Bias' variable with three levels – For, Against, or Neutral – and did not explicitly model the source's belief. We now prefer the newer model on conceptual grounds, but acknowledge this means Study 2 provides weaker evidence for it than if it had been pre-registered.

Study 2 contained a programming error which meant some participants were ejected at random intervals, meaning some experimental and demographic data is missing. We still obtained 939 observations from 185 participants (without the error, we would have had 1,696 from 212), and the errors occurred at random, so we decided the data was still appropriate to analyse.

Participants

Two-hundred and twelve people participated in the study. The programming error meant that 27 participants were excluded before supplying any data (they were still paid the advertised amount), leaving a sample of N=185. They were British nationals recruited via Prolific, paid £9/hour for a median completion time of ~3 minutes. The programming error only 123 participants provided demographic data – of these, 77 were female, 55 male, and 1 'Other' gender, and the median age was 34 (*Range*: 18-70).

Design and Procedure

We showed participants eight dialogues in which two people – James and Anne – discussed the fact that their colleague Keith "did not say anything" when "directly asked" whether another colleague – Carl – had committed a workplace misdemeanor. Each dialogue began with James asking Anne whether she believed Carl had committed the misdemeanor, to which Anne replied "I really have no idea". James then told Anne about Keith's avoidance of the question. We varied contextual aspects of James' description of Keith's silence, and asked participants how confident they were that Carl had committed the misdemeanor.

We employed a 2x2x2 within-subjects factorial manipulation of Keith's bias, expertise and the accuracy constraints. With bias, in the 'biased-towards' condition, Keith was described as "an enemy of Carl's". This makes Keith biased *towards* the hypothesis which was put to participants, which was that Carl *had committed* the misdemeanor. In the 'biased-against' condition, Keith was described as "a close personal friend of Carl's", which makes him biased *against* the hypothesis. To manipulate expertise, we varied whether Keith could be expected to know about Carl's workplace conduct – in the 'high expertise' condition, James said to Anne that Keith "works with him [Carl] every day", whereas in the 'low expertise' condition James said that Keith "rarely sees him [Carl] at work". We varied the accuracy constraints by changing the situation in which Keith was asked about Carl's alleged misdemeanor – either at "an official enquiry" in the high accuracy condition, or at "a company party" in the low accuracy condition.

Participants completed each trial in a random order. We fully randomized the alleged misdemeanor across trials – and ensure they were all serious enough for a workplace enquiry to be plausible: "embezzled money from the company", "lost their temper with a co-worker", "leaked sensitive documents to a competitor", "faked being sick to go on holiday", "sent the wrong files to the director", "failed to file the quarterly report on time", "has been having an affair with a colleague", and "was late for a meeting with an important client".

Participants provided their confidence ratings in Carl having committed the misdemeanor on 0-100 sliders, with 0 labelled "I am 100% confident that he DID NOT" and 100 labelled "I am 100% confident that he DID". Two attention check questions were intermixed with the eight trials, which simply asked participants to give a particular numerical response. Before completing the eight trials, participants read task information and provided consent. After the trials, they provided age and gender information, then were debriefed and thanked.

Results

We transformed all the experimental estimates by dividing by 100 before analysis. Figure 4 shows the mean posteriors per condition. The pattern should be qualitatively identical to that of Figure 3 – but it is not. Whereas the trends for the 'Bias Direction: Against' panel are as expected, those in the 'Bias Direction: For' are partially discrepant. As expected, these posteriors are lower than when the source is biased against the hypothesis, but there are two discrepancies: people should have updated more, leading to lower posteriors, when the source had *high* expertise, and under *high* accuracy contstraints – but instead they have updated more with low accuracy constraints and low expertise.

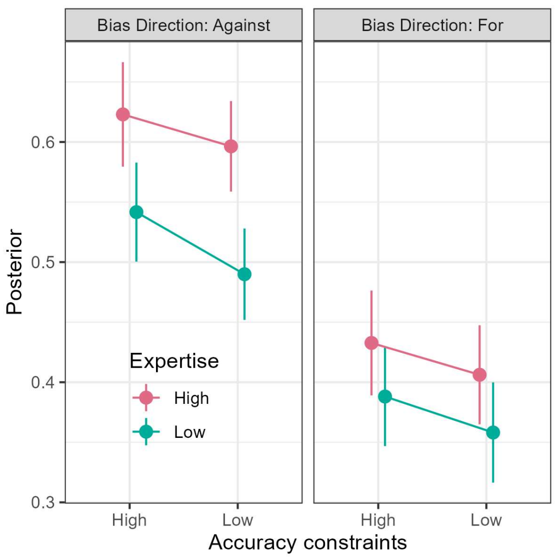

Figure 4. Observed posteriors in Study 2.

Discussion

There is still further work to do with Study 2, but the results so far provide a partial vindication of the model as well as implying it is missing something, as it cannot fully explain the trends when the source is biased towards the hypothesis. The design of Study 2 may have inadvertently introduced the kinds of additional complexities we were hoping to avoid during the early stages of testing this model, and these might help explain our results. Primarily, participants may have anticipated that as well as being constrained by accuracy, Keith was also constrained by loyalty – Keith may not have accused Carl of committing the misdemeanor, even when he was biased towards doing so, because he feared the repercussions he would receive from his colleagues for getting a fellow colleague into trouble. If this loyalty is expected more at the party than during the enquiry, and more from someone who didn't know Carl very well than from someone who did (as the person with higher expertise may actually be correct, which warrants lesser reprisal), this would make loyalty constraint a negative function of expertise and accuracy constraint, potentially explaining the stronger updating in the conditions when they are lower. However, we are yet to find a variation on the model which integrates a loyalty node of this kind that actually produces the observed results.

General Discussion

These results provide preliminary evidence that the epistemic weight of silence may derive from people's ability to infer latent conflicts between sources' testimonial motives. Sources have particular things they would like their audience to believe, but are subject to constraints on what speech is permissible. Silence implies a conflict between these factors, which allows witnesses to abduct that the source can't say what they would like to. When the constraint is accuracy, this implies the source believes that what they want to say is false. In corroboration, we have seen that when sources are silent in the face of hypotheses they are biased towards, people infer the hypothesis is likely false, but when the source is biased against the hypothesis, they infer the hypothesis is likely true. We have also seen that people's inferences are sensitive to the expertise of the source and the intensity of the accuracy constraints they are operating under. In most cases, people's inferences are consistent with our Bayesian Network model, implying that people infer from silence in a rational way.

However, we observed an unexpected pattern of results in Study 2, implying that inferences from silence are more complex than our model currently anticipates. The model is necessarily a simplification, and is intended to be specific instantiation of one of the many ways in which institutional constraints and biases can interact with beliefs to produce testimony – other constraints, like loyalty constraints, honesty constraints (which differ from accuracy *per se*), or total bans on speech, may apply in other contexts. Our hope is that the model can serve as a base to which additional nodes can be added to explain inferences from silence in more complex situations.

Our model could also integrate biases which affect people's *beliefs* rather than just their testimonial intentions. Belief updating is sensitive to biases that affect sources' beliefs (Wallace et al., 2020). Belief biases are distinct from intentional biases (e.g., partisan media sources may withhold facts that are uncongenial to their narrative, but know perfectly well which are true or false), and in these studies we ensured that sources' beliefs are unlikely to be influenced by biases. A node for belief-biases can be easily added into the model.

It is also worth noting that in many real-world cases, sources 'dodge' questions without being silent *per se*, instead using evasions such as "No comment", "I don't recall", "I plead the Fifth", or "I don't know", changing the subject, or appealing to some reason why they cannot speak. In many cases, these occur in the same situations as silence as we have modelled (see Table 1), and so should lead to the same inferences. One exception is when the source has a genuine reason they cannot speak on a matter, such as ongoing legal proceedings. Such an overriding reason for silence can modelled by adding a node to the model, with a prior for how constraining it should be.

These studies suggest the model is a good starting point, but further research will be required to determine it can be adjusted to account for additional constraints and other complexities.

References

Corner, A., Hahn, U. & Oaksford, M. (2011). The psychological mechanisms of the slippery slope argument. *Journal of Memory and Language*, *64*, 133-152.

Goldberg, S. (2011). The Division of Epistemic Labor. *Episteme*, *8*(1), 112-125.

Hahn, U., & Oaksford, M. (2006). A normative theory of argument strength. *Informal Logic, 26*, 1-24.

Hahn, U. & Oaksford, M. (2007). The burden of proof and its role in argumentation, *Argumentation, 21*, 39-61.

Harris, A. J. L., Corner, A., & Hahn, U. (2013). James is polite and punctual (and useless): A Bayesian formalisation of faint praise. *Thinking & Reasoning, 19*(34), 414–429.

Hahn, U., Oaksford, M. & Bayindir, H. (2005). How convinced should we be by negative evidence? In Bara, B., Barsalou, L. & Bucciarelli, M. (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, Mahwah, NJ: Erlbaum, 887-892.

Hahn, U., Oaksford, M., & Corner, A. (2005). Circular arguments, begging the question and the formalization of argument strength, in A. Russell, T. Honkela, K. Lagus & M. Polla (Eds.), *Proceedings of AMKLC '05 International Symposium on Adaptive Models of Knowledge, Language and Cognition*, Helsinki, 34-40.

Harris, A., Hsu, A. & Madsen, J. K. (2012). Because Hitler did it! Quantitative tests of Bayesian argumentation using Ad Hominem. *Thinking & Reasoning, 18*(3), 311-343.

Hayes, B. K., Banner, S., Forrester, S., & Navarro, D. J. (2019). Selective sampling and inductive inference: Drawing inferences based on observed and missing evidence. *Cognitive Psychology*, *113*, 101221.

Oaksford, M. & Chater, N. (2007). *Bayesian Rationality: the probabilistic approach to human reasoning*, Oxford University Press, Oxford: UK.

Oaksford, M., & Hahn, U. (2004). A Bayesian approach to the argument from ignorance. *Canadian journal of experimental psychology = Revue canadienne de psychologie experimentale*, *58*(2), 75–85.

Oaksford, M., & Hahn, U. (2012). Why are we convinced by the ad hominem argument? Source reliability or pragmadialectictics. *Bayesian Argumentation*, 39-58.

Russell, B. (1952). Is There a God? In Slater, J. G. (ed.). *The Collected Papers of Bertrand Russell*, *Vol. 11: Last Philosophical Testament*, *1943–68*. Routledge.

Wallace, L. E., Wegener, D. T., & Petty, R. E. (2020). When Sources Honestly Provide Their Biased Opinion: Bias as a Distinct Source Perception With Independent Effects on Credibility and Persuasion. *Personality and Social Psychology Bulletin*, *46*(3), 439–453.

{gRain}: Højsgaard, S. (2012). Graphical Independence Networks with the gRain Package for R. *Journal of Statistical Software*, 46(10), 1-26. URL https://www.jstatsoft.org/v46/i10/

R: R Core Team (2022). R: A language and environment for statistical computing. R. *Foundation for Statistical Computing*, Vienna, Austria. www.R-project.org/.