When agreement looks like copying: Testing a Bayesian model of source dependency inference

Sonja Belkin[1] s.belkin@lse.ac.uk
Jens Koed Madsen[1] j.madsen2@lse.ac.uk
Lee H de-Wit[2] lhd26@cam.ac.uk

[1]Department of Psychological and Behavioural Science, London School of Economics and Political Science, Houghton Street, WC2A 2AE, London, UK.
[2]Department of Psychology, University of Cambridge, Downing Pl, CB2 3EB, Cambridge, UK

Abstract

When multiple sources agree, their testimony provides stronger evidence if they are independent. But how do people infer whether sources are coordinated? We developed a Bayesian model predicting that dependency inferences should depend on three factors: verbal similarity between reports, a topic's knowability, and the diversity of possible expressions. In a pre-registered experiment (N = 156), participants viewed social media posts varying in these dimensions and judged the likelihood of coordination. Results revealed a striking dissociation: participants were highly sensitive to verbal similarity ($\eta^2 = .283$), inferring substantially more dependency when posts were near-identical versus substantively similar but differently worded. However, participants appeared insensitive to knowability and expressive similarity, contrary to normative predictions. These findings suggest people rely on a simple "copy detection" heuristic based on surface similarity rather than engaging in full Bayesian integration of domain-relevant probabilistic information.

Keywords: source dependency; testimony evaluation; Bayesian inference; verbal similarity; heuristic reasoning

Introduction

Imagine three people independently recommend the same new restaurant. Intuitively, this convergence feels like strong evidence, as three separate positive dining experiences are unlikely to all be wrong. But suppose all three describe the pasta as having "a revelatory depth of flavour that transports you straight to Naples". Now, the evidence appears different. While three people might independently enjoy the same dish, three people independently choosing the same distinctive phrasing when recommending it seems far less likely. The suspicion arises that your friends may have read the same review of the restaurant, or collectively discussed their experience, and that this may have informed their testimony, challenging the independence of their endorsements. If, for instance, their endorsement is due to reading a review, you should disregard your friends' testimonies, and care solely about whether the original source – the review – is reliable.

This example illustrates the inferential challenge at the heart of evaluating corroborating testimony. The evidential value of agreement depends not only on *whether* sources agree, but on features of the domain and the testimony itself. Three friends agreeing that a restaurant is "good" is unsurprising even under independence, since positive restaurants tend to be recognized as such. But three friends producing identical, distinctive descriptions suggests coordination, because independent diners would likely express their enjoyment differently. The more improbable the observed similarity under independence, the stronger the evidence for a common source.

More formally, agreement among sources is epistemically valuable only to the extent that sources are independent. According to Condorcet's Jury Theorem, the probability that a majority of independent witnesses are correct approaches certainty as the number of witnesses increases (Boland, 1989). However, when sources are dependent, which might include sharing information, copying from one another, or responding to common influences, agreement provides far less epistemic support. In the extreme case, where all sources simply copy their testimony from a single review, their collective agreement provides no more evidence than the original reviewer alone (Bovens & Hartmann, 2003).

The ability to detect source dependency has become critical in an information environment characterized by coordinated inauthentic behaviour (Giglietto et al., 2019), press release journalism (Lewis et al., 2008), and information cascades (Bikhchandani et al., 1992). Misjudging dependency in the information environment has practical consequences, such as making 'astroturfing' attempts by malicious state actors more persuasive (Goldstein et al., 2023). Further, in legal contexts, jurors may overestimate the strength of evidence, if they do not distinguish corroboration from contamination (Loftus, 2005). Similarly, replication crises may partly reflect hidden methodological dependencies (Ioannidis, 2005).

A Bayesian framework for inferring dependency

The formal epistemology literature provides a normative framework for reasoning about source dependency. Bovens and Hartmann (2003) developed Bayesian models showing how the epistemic value of converging testimony relies critically on assumptions about source independence (see Madsen et al., 2020 for an empirical test of some of these intuitions). When sources are known to be independent, agreement dramatically increases posterior probability in the reported proposition. When sources are known to be dependent, this boost is attenuated or eliminated (see also Pilditch et al., 2025). However, in most real-world situations, independence is not directly observable; it must be inferred from available cues. This raises the question: what

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

information do people use to judge whether sources are independent, and how do they integrate this information with observed agreement?

Psychological research has begun to address these questions. Yousif, Aboody, and Keil (2019) found that, in certain contexts, people are able to distinguish between independent and non-independent evidence, giving more weight to information from sources with distinct rather than shared origins. Harris and Hahn (2009) showed that coherence among testimonies increases belief more when sources are perceived as independent. Recently, Young et al. (2025) demonstrated that disagreements about source independence can produce issue polarisation: people who perceive sources as independent update more from their agreement than those who perceive them as dependent. Similarly, Alister et al. (2025) found that although most participants were insensitive to whether a consensus was formed independently or through shared sources, individual-level analysis revealed substantial heterogeneity: some participants reliably preferred independent evidence, whereas others showed the opposite pattern.

Despite this progress, several questions remain unresolved. First, prior work has typically manipulated independence perceptions through explicit instruction or contextual framing. Little is known about how people *spontaneously* infer dependency from observable features of testimony itself, such as the degree of verbal similarity between reports or characteristics of the topic being discussed.

Second, the normative framework suggests that dependency inferences should be sensitive to domain characteristics. Intuitively, if three people independently produce identical GDP forecasts to multiple decimal places, this is far more surprising, and thus more suggestive of coordination, than if three people independently report the same football score. The probability of coincidental agreement depends on both how 'knowable' the topic is and how many ways there are to express the same information. Yet whether lay reasoners are sensitive to such factors remains untested.

Third, dual-process theories of judgment (Chaiken, 1980; Kahneman, 2003) suggest that people may rely on simple heuristics rather than engaging in full Bayesian integration. Verbal similarity might function as a readily accessible cue that triggers coordination inferences regardless of domain-specific considerations. If so, people might be highly sensitive to surface features of agreement while being insensitive to deeper probabilistic structure.

We develop and test a formal Bayesian model of source dependency inference. The model specifies how observers should integrate three factors: (1) the degree of verbal similarity between reports ('noise'), (2) the inherent knowability of the topic, and (3) the expected similarity of independent expressions ('expressive similarity'). We derive specific predictions about how each factor should influence posterior dependency judgments.

In a pre-registered experiment, we test these predictions by presenting participants with social media posts about scenarios varying in knowability and expressive similarity and manipulating verbal similarity between conditions. We examine whether participants integrate these factors as the model predicts, or whether they rely on simpler heuristics that neglect domain-relevant information. By comparing observed behaviour to normative predictions, we aim to characterize both the sophistication and the limits of intuitive reasoning about source independence.

Theoretical model

The question of how to evaluate corroborating testimony has deep roots in social epistemology. Philosophers have long recognized that agreement among sources provides stronger evidence when those sources are independent (Coady, 1992; Goldman, 1999). Bovens and Hartmann (2003) formalized this intuition within a Bayesian framework, demonstrating that the evidential value of converging reports depends critically on the assumed dependency structure among informants.

The present model extends this work by specifying how observers might *infer* dependency from observable features of testimony. Rather than assuming independence is known, we model a situation in which the observer must determine whether sources have coordinated based on the similarity of their reports and characteristics of the domain.

Let D denote the event that sources are dependent (i.e., coordinating or copying from a common source). The observer begins with a prior probability $P(D)$ reflecting baseline expectations about coordination in the given context. Upon observing n sources producing similar reports, the observer updates to a posterior probability $P(D \mid \text{similar reports})$ via Bayes' rule. Two domain parameters are central to this inference:

Knowability (K). This parameter represents the probability that any single source, acting independently, would arrive at the correct or consensus answer. For 'easy' topics, such as reporting a well-publicized football score, K is high (approaching 1), meaning independent sources will naturally converge. For 'hard' topics, such as forecasting next quarter's GDP, K is low, meaning independent sources are unlikely to agree by chance. This aligns with Yousif et al.'s (2019) conceptualization of knowability as a variable which captures how easily one can establish the truth of a claim, without needing to make inferences from complex data. Formally, then, K captures what Goldman (1999) termed the "epistemic accessibility" of the domain. Higher K implies that agreement is more likely under independence and thus provides weaker evidence for dependency.

Expressive Similarity (E). This parameter represents the probability that two sources, given they agree in substance, would also agree in form, using identical or near-identical phrasing. For topics with constrained expression (such as numerical values or standardized formulations), E is high. For topics permitting diverse expression (such as personal opinions or creative descriptions), E is low. This parameter captures linguistic and social conventions that determine how surprising identical wording should be.

Under **independence** ($\neg D$), the probability of observing identical reports (R) from two sources is the joint probability of reaching the same conclusion *and* expressing it identically:

$$P(R|\neg D) = K\ x\ E$$

Under **dependency** (D), reports are dictated by a common source, so similar reports are expected with high probability:

$$P(R|D) \approx 1$$

Extending to n sources, the probability of all n sources producing identical reports under independence requires each of the $n - 1$ subsequent sources to match the first in both substance and form:

$$P(R|\neg D) = K\ x\ E^{n-1}$$

The likelihood ratio is therefore:

$$LR = \frac{P(R|D)}{P(R|\neg D)} = \frac{1}{K\ x\ E^{n-1}}$$

Applying Bayes' theorem, the posterior probability of dependency given the observed reports is:

$$P(D|R) = \frac{LR\ X\ P(D)}{LR\ X\ P(D) + \big(1 - P(D)\big)}$$

This formulation makes clear that the posterior probability of dependency depends on three factors: the prior probability of coordination $P(D)$, the knowability of the topic (K), and the expected expressive similarity (E). All three parameters can in principle be elicited directly from observers.

Model predictions

The model generates predictions about when agreement provides strong evidence for dependency (high LR) versus when agreement is uninformative ($LR \approx 1$). An agreement suggests dependency under three conditions:

1. **The topic is hard to know (low K):** When independent sources are unlikely to reach the same conclusion by chance, their agreement is surprising and suggests coordination.
2. **There are many ways to express an idea (low E):** When identical phrasing is improbable under independence, verbal similarity provides strong evidence for a common source.
3. **Many sources agree (high n):** As the number of agreeing sources increases, the probability that all have independently converged on identical wording decreases exponentially.

However, an agreement is uninformative under three conditions:

1. **The topic is easy to know (high K):** When everyone can independently verify the truth (e.g., a widely publicized sports score), convergence is expected, regardless of coordination.
2. **There are few ways to express an idea (high E):** When canonical phrasings exist (e.g., standard ways to report numerical values), identical wording is unsurprising, even under independence.
3. **Few sources are observed (low n)** With only two or three sources, coincidental agreement remains plausible.

To illustrate the model's predictions, consider three sources ($n = 3$) producing identical reports, with a prior probability of dependency $P(D) = .20$. Table 1 shows how the posterior probability varies with K and E. When both K and E are high (.90), agreement barely shifts the posterior from the prior (.26 vs. .20). When both are low (.15), even with a modest prior of .20, the posterior rises to .99.

*Table 1: posterior probability predictions**

K	E	LR	P(D | R)
.90	.90	1.37	.26
.50	.50	8.00	.67
.30	.30	37.04	.90
.15	.15	296.30	.99

** P(D | R) is calculated with n = 3, and P(D) = 0.2*

This illustrates how the same observation (identical reports from three sources) can provide anywhere from negligible to overwhelming evidence for coordination, depending on domain characteristics. A key advantage of this model is that all parameters can be measured directly from participants:

Prior probability P(D): "Before seeing their responses, how likely do you think it is that these sources are coordinated?"

Knowability (K): "If two experts independently analysed this topic, how likely do you think they would be to reach the same conclusion about it?"

Expressive similarity (E): "If two people held the same view on this topic, how likely is it that they would phrase their opinion identically?"

Responses to these questions (scaled 0 to 1) can be plugged directly into the model equations to generate predicted posteriors, which can then be compared to participants' actual dependency judgments. The following study tests these intuitions across three hypotheses:

H_1: Participants will infer greater dependency from low-noise reports, compared to high-noise reports.

H_2: Participants will infer greater dependency for scenarios that are lower knowability.

H_3: Where reports are low-noise, participants will infer greater dependency for scenarios that are lower in expressive similarity than scenarios with high expressive similarity.

Method

The target sample size for this study was determined through an a priori power analysis using *G*Power* (Faul et al., 2009), which aimed to achieve 80% power at an alpha level of .05. The analysis was based on the hypothesis with the largest sample size requirement, H3, which required 156 participants to detect a small effect size of .20.

166 participants (M_{age} = 46, SD_{age} = 12.3, N_{Female} = 77) were recruited via Prolific Academic. Participants were pre-screened for living in the UK, being fluent in English, having a platform approval rate between 95-100%, and for completing at least 100 previous submissions – to reduce the risk of bot accounts and maximize the recruitment of genuine participants. Nine participants did not fully complete the study, and one participant failed the attention checks. These submissions were excluded from analyses, leaving N = 156.

Participants in the final sample were reimbursed at an hourly rate of £9 for participating in the 5-minute experiment. Each participant provided informed consent and was debriefed at the end of the study. The study was preregistered on AsPredicted and received ethical approval from LSE (Submission ID: 67445).

Design & Procedure

The study employed a 2 (condition [low noise vs high noise]: between subjects) x 2 (knowability [high vs low]: within subjects) x 2 (expressive similarity [high vs low]: within subjects) mixed design. The experiment was designed using *Qualtrics* software.

Participants were randomly assigned to either the high-noise or low-noise condition. Following exclusions, there were 75 participants allocated to the high-noise condition (48.1%), and 81 allocated to the low-noise condition (51.9%).

Participants then viewed four scenarios presented sequentially, in a randomized order. Each scenario corresponded to one combination of the within-subjects factors of knowability and expressive similarity, such that all participants were exposed to each level of both factors once. The four scenarios were selected on the basis of a pilot study, where three alternative scenarios were developed for each combination of expressive similarity and knowability. The final scenarios selected were (1) reports on which country had won gold in the Olympics synchronized diving competition (high knowability, high expressive similarity), (2) reports expressing suspicion that a company's recent ad campaign involved the use of AI-generated content (high knowability, low expressive similarity), (3) posts predicting a country's GDP in the upcoming quarter (low knowability, high expressive similarity), and (4) posts speculating about a government minister's likely actions following implication in a corruption scandal.

For each scenario, participants provided their judgements, on a scale of 0-100, of the scenario's knowability, expressive similarity, and prior ratings of dependency. After providing their ratings, participants viewed five posts formatted to resemble X (formerly Twitter), that had been generated using *TweetGen*. The presentation of five posts was motivated by prior findings that sensitivity to informational dependency emerges only after exposure to multiple sources, while prior studies with three or four sources often fail to reveal these effects (Desai et al., 2025).

The five posts presented were manipulated to vary in noise. In the low-noise condition, posts were highly similar in phrasing and included overlapping, unusually specific details. Conversely, in the high-noise condition, posts conveyed agreement for the core claim, but differed in the wording used and details reported. This manipulation follows Pillai and Fazio's (2024) use of naturalistic paraphrasing, which more closely reflects real-world variation in reports than paraphrasing which has been varied according to strict linguistic rules (e.g. Silva et al., 2017). Figure 1 shows an example of the low-noise manipulation for the scenario high in knowability and low in expressive similarity.

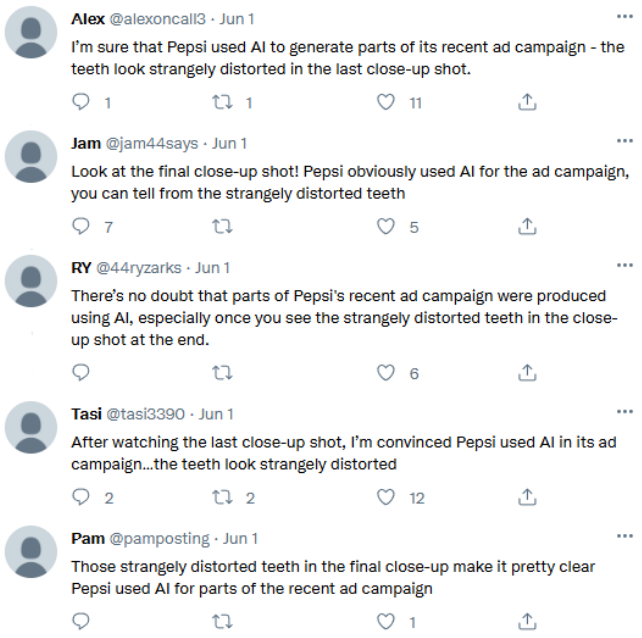

Figure 1: Low-noise condition for the high knowability, low expressive similarity scenario

After viewing the posts, participants were asked for their posterior ratings of dependency: "*Having now seen the posts from people discussing [topic], how likely do you think it is that people who made these posts may have coordinated their answers (whether accidentally, e.g. by getting the same information from a shared source, OR purposefully e.g. by discussing in advance)?*". Responses were recorded on a scale from 0 ("extremely unlikely") to 100 ("extremely likely").

Results

All analyses were conducted using *R* version *4.5.2* and are available at the OSF repository. To test H1 and H2, posterior dependency judgements were analyzed using a 2 (Noise: low vs. high; between-subjects) × 2 (Knowability: low vs. high; within-subjects) × 2 (Expressive Similarity: low vs. high; within-subjects) mixed ANOVA, computed with the *afex* package (Singmann et al., 2021).

Results showed a significant main effect of Noise, $F(1, 154) = 91.53$, $p < .001$, generalized $\eta^2 = .283$. This indicated that participants in the low-noise condition inferred substantially more dependency across the posts they viewed, compared to participants in the high-noise condition. Estimated marginal means, conducted using the *emmeans* package (Lenth, 2021), confirmed this effect: scenarios in the low-noise condition elicited higher dependency judgements ($M = 73.7$, $SE = 2.46$, 95% CI [68.8, 78.6]) than in the high-noise condition ($M = 39.7$, $SE = 2.56$, 95% CI [34.6, 44.8]), mean difference = 34.0, $SE = 3.55$, $t(154) = 9.57$, $p < .001$, supporting H1. Results are visualised in Figure 2.

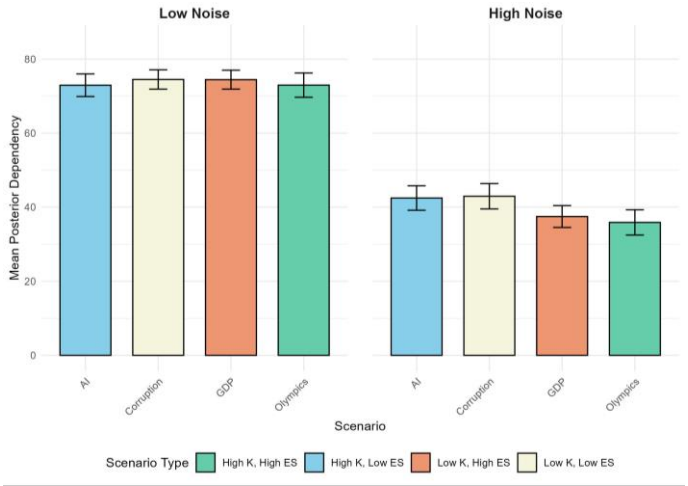

Figure 2: The Effect of Noise on Dependency Judgements

However, there was no significant main effect of Knowability, $F(1, 154) = 0.75$, $p = .387$, generalized $\eta^2 < .001$, indicating that posterior dependency judgements did not differ significantly between low- ($M = 57.3$, $SE = 1.80$, 95% CI [53.8, 60.9]) and high-knowability scenarios ($M = 56.1$, $SE = 2.04$, 95% CI [52.0, 60.1]). Further, the Noise × Knowability interaction was not significant, $F(1, 154) = 0.03$, $p = .869$, suggesting that the increase in inferred dependency in the low-noise condition was not moderated by knowability. Consequently, H2 was not supported.

Finally, there was a significant, though very small, main effect of Expressive Similarity, $F(1, 154) = 4.84$, $p = .029$, generalized $\eta^2 = .003$. This suggested – contrary to the direction of expectations of H3 – that the scenarios higher in Expressive Similarity elicited slightly higher inferences of dependency than scenarios low in Expressive Similarity. This effect was qualified by a significant Noise × Expressive Similarity interaction, $F(1, 154) = 4.80$, $p = .030$, generalized $\eta^2 = .003$. Follow-up simple effects analyses showed that, in the low-noise condition, there was no significant difference in dependency judgements between scenarios with low- and high- Expressive Similarity, mean difference = 0.01, $SE = 1.91$, $t(154) = 0.01$, $p = .995$. In the high-noise condition, however, scenarios high in Expressive Similarity elicited significantly higher dependency judgements, than scenarios low in Expressive Similarity, mean difference = 6.03, $SE = 1.98$, $t(154) = 3.05$, $p = .003$.

To test H3, a paired samples t-test was conducted to determine whether, within the low-noise condition, judgements of dependency differed between scenarios that were high or low in Expressive Similarity. When noise was low, there was no significant difference in inferences of dependency between the low ($M = 73.72$, $SD = 22.06$) and high Expressive Similarity condition ($M = 73.70$, $SD = 22.21$), $t(80) = 0.01$, $p = .50$, 95% CI $[-2.55, \infty]$.

To confirm that scenarios intended to be high in knowability and expressive similarity were rated as significantly more knowable and expressively similar than scenarios intended to be low in knowability and expressive similarity, two paired-samples t tests were conducted. Results indicated that there was a significant difference between the low and high knowability scenarios, $t(155) = 11.98$, $p < .001$, 95% CI [14.73, 20.54], and between the low and high expressive similarity scenarios, $t(155) = 2.39$, $p = .018$, 95% CI [0.64, 6.74]. Results from the manipulation checks are visualized in Figure 3.

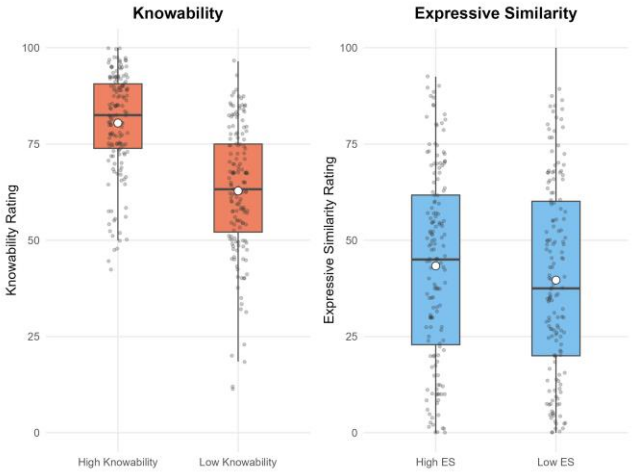

Figure 3: Manipulation Checks for Knowability and Expressive Similarity

Discussion

The present study tested a Bayesian model of source dependency inference, which predicts that observers should infer greater coordination among sources when they agree on topics that are difficult to know or when identical phrasing is improbable by chance. Our results revealed a striking dissociation: while participants were highly sensitive to verbal similarity between reports (H1 strongly supported), they appeared to show no sensitivity to the normatively relevant factors of knowability and expressive similarity (H2 and H3 not supported). These findings contribute to our understanding of how people evaluate corroborating testimony, a question with implications for epistemology, misinformation research, and the psychology of social learning.

Sensitivity to verbal similarity: The effect of noise (the degree of verbal similarity between posts) was large and robust ($\eta^2 = .283$). Participants who saw near-identical posts with overlapping specific details inferred substantially more dependency between sources ($M = 73.7$) than those who saw posts that agreed in substance but differed in phrasing ($M = 39.7$). This 34-point difference confirms that people are highly attuned to verbal similarity as a cue to coordination.

This finding aligns with research on plagiarism detection and the psychology of copying (Marsh & Bower, 1993; Roediger & McDermott, 1995). Verbatim similarity is a powerful diagnostic cue because it is unlikely to arise by chance when multiple people independently describe the same event. Indeed, the plagiarism detection literature confirms this intuition: systems designed to identify copying rely heavily on exact string matches and statistically

improbable phrases precisely because such overlap is rare under independence (Weber-Wulff, 2014). Our participants appear to apply similar reasoning, treating near-identical wording as a strong signal of coordination. The strength of this effect is consistent with dual-process models of persuasion and judgment (Chaiken, 1980; Petty & Cacioppo, 1986). According to the heuristic-systematic model (Chaiken et al., 1989), people often rely on simple cues rather than engaging in effortful probabilistic reasoning. Verbatim similarity may function as a highly accessible heuristic cue, one that triggers coordination inferences without requiring integration of more abstract probabilistic information about topic characteristics.

Insensitivity to normatively assumed factors: Despite successfully manipulating knowability (19-point difference in K ratings between high and low scenarios), participants did not adjust their dependency inferences accordingly. When three sources produced identical GDP forecasts (a domain where independent agreement is normatively surprising), participants were no more suspicious of coordination than when three sources identically reported an Olympic result. This null effect, consistent with Alister et al.'s (2025) finding that that consensus persuasiveness does not vary with knowability regardless of source independence, is striking: it suggests a fundamental disconnect between lay reasoning and the Bayesian model. The failure to integrate knowability information echoes findings from the base-rate neglect literature (Kahneman & Tversky, 1973; Bar-Hillel, 1980): just as people often ignore prior probabilities when evaluating diagnostic evidence, our participants ignored the prior probability of independent agreement (determined by K) when evaluating the diagnostic value of similar reports. This pattern is consistent with Kahneman's (2003) account of attribute substitution; rather than computing the normatively relevant probability P (dependency | identical reports, topic), participants may substitute a simpler judgment ("how similar do these posts look?"), independent of topic characteristics.

Theoretical implications

Our findings speak to broader debates about the nature of human reasoning under uncertainty. Two complementary explanations may account for the observed pattern:

A Bayesian-consistent interpretation: The noise manipulation may have created such extreme likelihood ratios that K and E effects were swamped. When five sources produce near-identical text with specific overlapping details, the evidence for coordination is overwhelming, regardless of domain characteristics. Under the proposed model, if P(identical wording | independence) approaches zero in all scenarios, the likelihood ratio approaches infinity and domain priors become negligible. This interpretation suggests our design was underpowered to detect K and E effects because the low-noise stimuli were *too* similar (a design limitation rather than a failure of Bayesian reasoning).

Heuristic alternative: People may rely on a simpler 'copy detection' heuristic (a threshold-based rule where sufficiently similar text triggers coordination inferences without integration of probabilistic cues about the domain). This would be consistent with Simon's (1955, 1990) concept of bounded rationality: rather than optimising, people satisfice by using computationally cheap rules that work well enough in most contexts. The ceiling effect observed in the low-noise condition (all scenarios ≈ 74) is consistent with a threshold model in which near-identical wording crosses a fixed detection threshold regardless of other factors. Moreover, this explanation aligns with findings by Yousif et al. (2019) that individuals over-rely on superficial cues, like source consensus, while failing to account for source dependence. These interpretations are not mutually exclusive. People may be *capable* of Bayesian integration when evidence is moderate but may resort to heuristics when evidence is extreme. This account parallels findings in the decision-making literature which show that people integrate information more normatively when cues are probabilistic rather than deterministic (Juslin et al., 2009).

Limitations Several limitations qualify these conclusions. First, the expressive similarity manipulation achieved a substantially weaker point difference than the knowability manipulation (3.7 vs 19). Future research should develop scenarios with clearer ES variation, perhaps by comparing topics where canonical phrasing exists (e.g., scientific abstracts) with topics permitting diverse expression (e.g., creative description), to provide a stronger test of whether people are sensitive to this factor. Second, the strong noise manipulation may have produced ceiling and floor effects that obscured moderation. A design with moderate similarity levels (i.e., posts share some phrasing) might reveal K and E effects masked here. This would test whether the absence of effects reflects genuine insensitivity or merely the overwhelming strength of near-identical wording as evidence. Third, our stimuli were social media posts on relatively mundane topics. The stakes of coordination detection may matter, as people might reason more carefully about source independence when evaluating expert testimony or consequential decisions. Future research could examine whether the observed pattern generalizes to high-stakes contexts or whether motivated reasoning modulates sensitivity to dependency cues (Kunda, 1990).

Concluding remarks

This study demonstrates that people are highly sensitive to verbal similarity when inferring source coordination, but do not appear to modulate this inference based on the normatively relevant factors of knowability or expressive diversity. These results contribute to our understanding of how people evaluate corroborating testimony in an information environment where coordination, copying, and cascading are increasingly common. Findings add to a growing literature documenting the sophistication and limits of intuitive statistical reasoning (Griffiths & Tenenbaum, 2006), suggesting that while people possess heuristics for detecting blatant coordination, they may lack the capacity (or motivation) to integrate more subtle probabilistic cues about the domains in which testimony is offered.

References

Alister, M., Ransom, K., Connor Desai, S., Soh, E. V., Hayes, B., & Perfors, A. (2025). How convincing is a crowd? Quantifying the persuasiveness of a consensus for different individuals and types of claims. *Psychological Science, 36*(7), 483–498

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

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy, 100*(5), 992–1026.

Boland, P. J. (1989). Majority systems and the Condorcet jury theorem. *Journal of the Royal Statistical Society: Series D, 38*(3), 181–189.

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford University Press.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*(5), 752–766.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). Guilford Press.

Coady, C. A. J. (1992). *Testimony: A philosophical study*. Oxford University Press.

Desai, S. C., Ransom, K. J., Angles, L., & Hayes, B. (2025). Limits of repetition in the illusion of consensus. *Proceedings of the Annual Meeting of the Cognitive Science Society, 47*(0)

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160

Giglietto, F., Righetti, N., Rossi, L., & Marino, G. (2019). Coordinated inauthentic behavior on Facebook. *Information, Communication & Society, 23*(5), 672–693.

Goldman, A. I. (1999). *Knowledge in a social world*. Oxford University Press

Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., & Katerina Sedova. (2023). Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations. *ArXiv* [*Preprint*]

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science, 17*(9), 767–773.

Harris, A. J. L., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1366–1373.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124.

Juslin, P., Nilsson, H., & Winman, A. (2009). Probability theory, not the very guide of life. *Psychological Review, 116*(4), 856–874.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697–720.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237–251.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480–498.

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* (R package version 1.6.1)

Lewis, J., Williams, A., & Franklin, B. (2008). A compromised fourth estate? UK news journalism, public relations and news sources. *Journalism Studies, 9*(1), 1–20.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361–366.

Madsen, J. K., Hahn, U., & Pilditch, T. D. (2020). The impact of partial source dependence on belief and reliability revision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(9), 1795–1805

Marsh, R. L., & Bower, G. H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 673–688.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205.

Pilditch, T. D., Hahn, U., & Lagnado, D. (2025). The problem of dependency. *Synthese, 205*, 143

Pillai, R. M., & Fazio, L. K. (2024). Repeated by many versus repeated by one: Examining the role of social consensus in the relationship between repetition and belief. *Journal of Applied Research in Memory and Cognition, 14*(2), 154–166

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814.

Silva, R. R., Garcia-Marques, T., & Reber, R. (2017). The informative value of type of repetition: Perceptual and conceptual fluency influences on judgments of truth. *Consciousness and Cognition, 51*, 53–67

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*(1), 99–118.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*(1), 1–20.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2021). *afex: Analysis of factorial experiments* (R package version 1.0-1)

Weber-Wulff, D. (2014). *False feathers: A perspective on academic plagiarism*. Springer.

Young, D. J., Mirabile, P., Mallpress, D. E. W., & Lagnado, D. A. (2025). Belief polarization can be caused by disagreements over source independence. *Cognition, 256*, 106022.

Yousif, S. R., Aboody, R., & Keil, F. C. (2019). The illusion of consensus: A failure to distinguish between true and false consensus. *Psychological Science, 30*(8), 1195–1204.