WORKING PAPERS

N° 1678

October 2025

"Using Generative AI to Increase Skeptics' Engagement with Climate Science"

Bence Bago, Philippe Muller and Jean-François Bonnefon

Toulouse School of Economics

Using Generative AI to Increase Skeptics' Engagement with Climate Science

Look for the published version in Nature Climate Change, 2025

Bence Bago *[1], Philippe Muller[2], and Jean-François Bonnefon[3]

[1]Department of Social Psychology, Tilburg University, Tilburg NL
[2]Institut de Recherche en Informatique de Toulouse, Université de Toulouse, Toulouse, France
[3]Toulouse School of Economics, CNRS (TSM-R), Université Toulouse Capitole, Toulouse, France

*To whom correspondence should be addressed; E-mail: b.bago@tilburguniversity.edu

Abstract

Climate skepticism remains a significant barrier to public engagement with accurate climate information, because skeptics actively engage in information avoidance to escape exposure to climate facts. Here we show that generative AI can enhance engagement with climate science among skeptical audiences by subtly modifying headlines to align better with their existing perspectives, without compromising factual integrity. In a controlled experiment (N = 2000) using a stylized social media interface, headlines of climate science articles modified by an open-source large language model (Llama3 70B, version 3.0) led to more bookmarks and more upvotes, and these effects were strongest among the most skeptical participants. Participants who engaged with climate science as a result of this intervention showed a shift in beliefs towards alignment with the scientific consensus by the end of the study. These results show that generative AI can alter the information diet skeptics consume, with the promise that scalable, sustained engagement will promote better epistemic health. They highlight the potential of generative AI, showing that while it can be misused by bad actors, it also holds promise for advancing public understanding of science when responsibly deployed by well-intentioned actors.

1 Main

Tackling global challenges requires collective action, which is difficult when people lack a shared understanding of scientific facts [1, 2, 3]. Misconceptions about vaccines [4], migration [5], and climate science [6, 7, 8] have all undermined coordinated responses. Since ambitious climate policies require strong public support, many efforts have focused on messages that resonate with skeptics. The simplest approach here is perhaps the most effective: Communicating facts, particularly the scientific consensus, successfully decreases skepticism [9, 10, 11, 12]. In experiments, exposure is forced; in everyday life, skeptics often actively avoid these facts [13, 14, 15, 16, 17, 18]. The problem, then, is not delivering facts to skeptics, but bringing skeptics to the facts.

Information avoidance occurs when individuals actively steer clear of facts that could challenge their emotions, beliefs, or behaviors [19, 20, 21, 22, 23, 24]. First, climate change facts are emotionally charged and provoke feelings of fear, helplessness, or anxiety [25]. Skeptics may avoid them to preserve emotional well-being. Second, because climate change is a highly polarized issue, skeptics may avoid facts that clash with their identity or worldview [14, 26, 27, 28], and select out of news sources likely to present such information [29, 30, 16, 24]. Third, skeptics may avoid climate facts to sidestep costly lifestyle changes or moral duties, like changing one's diet [31] or cutting air conditioning use [32].

In sum, there are many reasons for skeptics to avoid news stories that contain climate science facts, especially when they anticipate that these stories will challenge their current views—in other words, when they anticipate that these stories are written for a non-skeptic audience, would not fit their views, would make them experience negative emotions, and would contain little to no useful information for them [33, 34]. This problem is exacerbated by changes in the news landscape over the last decades. Journalists working for outlets that report on climate science have been under pressure to create content fitted to an audience that is both shrinking and becoming more homogeneous in ideology [35, 36, 27, 24, 37, 38, 28]. This creates a feedback loop in which stories are tailored to a non-skeptic audience, pushing away skeptics, and further increasing the need to tailor stories to a non-skeptic audience [8, 39, 40].

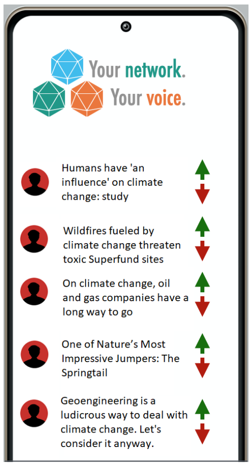
How can we bypass information avoidance and increase skeptics' engagement with climate news? Reducing the volume of inaccurate information is not a solution [41, 42], because we know from research on partisan news that reducing exposure to like-minded sources does not ensure that people opt into alternative, accurate sources [43, 44, 45]. Applying specific frames to climate headlines (e.g., environmental, public health, national security, economic, or moral angles) has shown little impact on skeptics' [13, 46, 47], and increasing the negativity may backfire [47, 48], despite its general efficacy elsewhere [49, 50]. One-shot interventions [51] also fall short, because their effects decay quickly [52, 53], and because climate science evolves too rapidly for static messages. We therefore need a scalable, repeatable intervention that boosts skeptics' engagement with climate headlines—without relying on negativity, fixed frames, or compromising factual integrity.

Here we show that an open-source generative AI model [54] (Llama3 70B, version 3.0; we used Llama3 because its open weights support reproducibility; unlike proprietary models, it is not subject to opaque updates that could alter future replications) can be used to increase skeptics' engagement with climate news by rewriting headlines to reduce anticipated disagreement, regret, and negative emotions—without increasing negativity or compromising factual integrity. This intervention has the largest impact on the most skeptical individuals, and shifts beliefs towards alignment with the scientific consensus. These outcomes contrast with the common view of generative AI as a driver of misinformation and miscalibrated beliefs [55, 56, 57, 58, 59, 60]. Recent work, however, shows that large language models can also be used constructively, e.g., to reduce conspiracy beliefs through personalized dialogues [61, 62], or to help people reach consensus on divisive issues [63]. Our approach follows this constructive line in the specific context of climate science communication. Algorithms have shaped journalism for over a decade [64, 65], and this influence is now accelerating with generative AI. While journalists have voiced ethical concerns about credibility, accuracy, and bias in AI-assisted news production, they tend to view its use in news distribution more positively [66, 67]—particularly when it helps reach news outsiders who lack the motivation or capacity to engage [68]. Our work is situated at this distribution stage, aiming to reach climate skeptics who typically avoid such coverage, while upholding journalistic integrity.

2 Results

Measures We recruited 2,000 U.S. participants quota-matched for sex, age, and political partisanship. They reported their prior climate change beliefs before the experiments. Participants were categorized as believers, skeptics, or others based on a question from the Yale Climate Change Communication Center [69, 70]. To the question 'Assuming global warming is happening, do you think it is…?', participants who answered 'mostly caused by human activity' ($n = 1,414$) were categorized as believ-

(A)



(B)

Examples of original headlines	Llama3-modified headlines
The Foods that Reverse Climate Change	Precision Agriculture and Gene-Editing: The Future of Food Production?
Mysterious Origin of the 'Tree of Life' Revealed as Some of the Species Are Just Decades from Extinction	Ancient 'Tree of Life' Finally Has Its Origins Traced
Aircraft Turbulence is Worsening with Climate Change. Studying Birds Could Help	Birds May Hold the Key to Predicting Turbulent Skies
Inside the Giant 'Sky Rivers' Swelling with Climate Change	Chasing the Storm: Scientists Fly into Atmospheric Rivers to Improve Forecasts
Climate Science is Sound. Satellite Timelapse Doesn't Disprove Sea Level Rise	Experts Say Satellite Images Alone Can't Detect Sea Level Rise

Figure 1: Experimental methods. (A) The stylized social media platform used in the experiment. (B) Examples of original and Llama3-modified climate headlines.

ers. Participants who answered 'caused mostly by natural changes in the environment' ($n = 412$) or 'none of the above because climate change is not happening' ($n = 57$) were categorized as skeptics; and participants who answered 'other' ($n = 53$) and 'don't know' ($n = 63$) were categorized as others. Our primary preregistered analyses compare the 1,414 believers (coded as $+0.5$) to the 479 skeptics (coded as -0.5), and code 'others' as zero. Participants also rated three continuous 0–100 scales : belief that climate science is happening, belief it is caused by human activity, and belief it is a significant threat. Our secondary preregistered analyses use these continuous measures as an alternative to the categorical measure.

Participants engaged in a social media simulation featuring a feed of 20 news headlines: 11 on climate change and 9 on other science topics (Fig 1A). Their first task was to upvote or downvote each headline, providing us with our first measure of engagement: the probability of upvoting climate-related headlines, which would increase their visibility. Next, participants bookmark 10 headlines they would be interested in reading later, providing a second engagement measure: the probability of bookmarking climate headlines, reflecting willingness to be exposed to climate information. Third, participants read in full one of their bookmarked articles (always about climate, by design) and rated their experience: Did they regret this bookmark? Would they upvote or downvote after reading? How much did they trust the contents? Finally, participants reported their climate change beliefs post-reading. These measures assessed potential backfiring (skeptics feeling deceived into engaging with unwanted contents) and the intervention's impact on climate beliefs.

Treatments Our experiment compared the engagement and experience between participants shown either original headlines or Llama3-modified ones—true to the article's content, but phrased in a way that would not be inconsistent with the beliefs of someone who thinks climate change is not happening (see Fig. 1B for examples). Full details on article selection and headline modification are in the Methods section. We began with a large set of climate change articles (2022–2024) from trustworthy sources, and selected 58 articles containing scientific data, statements by scientists, or references to scientific studies. After using Llama3 to modify the headlines, we narrowed this set to 28 via two steps. First, we excluded 11 articles after a manipulation check with a separate sample of skeptics—these modified headlines did not significantly impact anticipated regret, agreement, or emotions. Next, a professional fact-checker excluded 19 more, rating the modified headlines as insufficiently relevant or accurate. This high exclusion rate reflects a deliberately high bar for relevance. Finally, a separate Llama3 instance screened the remaining 28 for undesirable clickbait features (e.g., negativity, sensationalism). We found no evidence of increased clickbait (and partial evidence of reduction) compared to the originals.

2.1 Engagement

Interaction of treatment and prior beliefs Figure 2 displays the effect of Llama3-modified headlines on upvotes and bookmarks (2A), as well as heterogeneity analyses at the participant (2B) and stimulus level (2C). Effects on downvotes are, by design, symmetrical to the effects on upvotes. On average (Figure 2A), modifying headlines led to a 11 percentage point increase in upvotes by skeptics, and a 7 percentage point increase in bookmarks. The preregistered outcome of interest was the interaction of treatment and prior beliefs, tested through 8 variants of the same general model. The general model was: $\texttt{Outcome} \sim \texttt{Belief} + \texttt{Treatment} + \texttt{Belief} \times \texttt{Treatment} + (1|\texttt{Headline}) + (1|\texttt{Participant})$. The outcome could be either an upvote or a bookmark, and the belief measure was either our categorical classification, or one of our three continuous measures. The interaction effect was statistically significant in all variants. Categorical classification (Figure 1A): upvotes: $b = 0.09$, $p < .001$; bookmarks: $b = 0.10$, $p < .001$. Continuous measure of belief that climate change is happening (Figure 1B top): upvotes: $b = 0.04$, $p < .001$; bookmarks: $b = 0.04$,

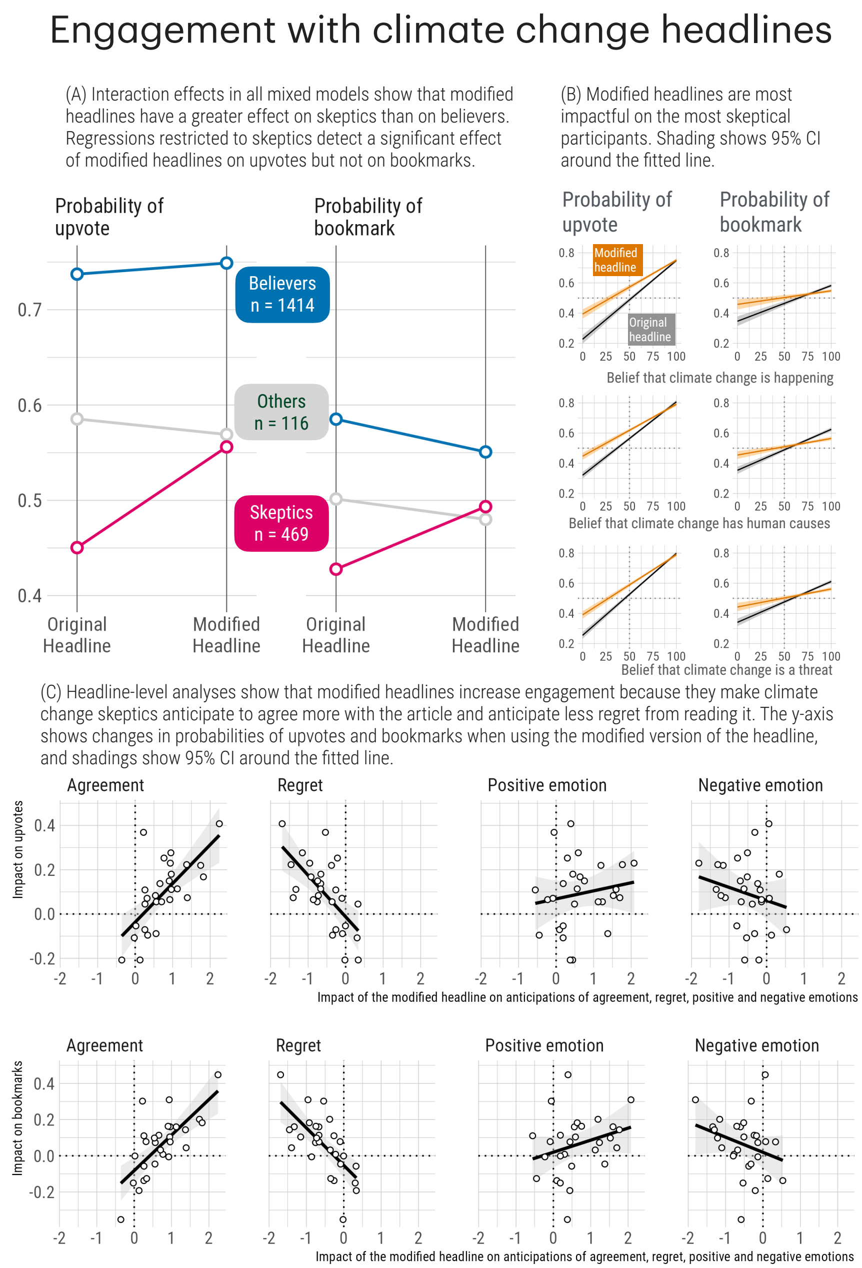


Figure 2: Main results on the engagement of skeptics with climate news. (A) Llama3-modified headlines increase the probability that skeptics upvote climate news, but this treatment effect is not detected as significant on bookmarks. For completeness, we show descriptive results for 'other' participants, who are neither believers nor skeptics. (B) Heterogeneity analyses at the participant level show that modified headlines are greater impact on both upvotes and bookmarks when people are more skeptic. (C) Heterogeneity analyses at the stimulus level show that the impact on both upvotes and bookmarks is larger for stronger manipulations of anticipated agreement and regret.

$p < .001$. Continuous measure of belief that climate change is caused by human activity (Figure 1B mid): upvotes: $b = 0.04, p < .001$; bookmarks: $b = 0.05, p < .001$). Continuous measure of belief that climate change is a significant threat (Figure 1B bottom): upvotes: $b = 0.04, p < .001$; bookmarks: $b = 0.05, p < .001$. Results were robust to the inclusion or exclusion of inattentive participants, see SI.

To explore the interaction further, we conducted separate analyses for skeptics and believers, defined as per the categorical variable. For each category, the model was: $\texttt{Outcome} \sim \texttt{Treatment} + (1|\texttt{Headline}) + (1|\texttt{Participant})$. For skeptics, we find a significant treatment effect on upvotes ($b = -0.1, p = 0.014$), but not on bookmarks ($b = -0.06, p = 0.157$). This analysis is exploratory since the preregistered outcome of interest was the interaction effect between belief and treatment. A simulation-based post-hoc power analysis suggested that we would have needed about 1,000 skeptics and 350 headlines to detect the $b = -0.06$ effect size on bookmarks with a 95% power—and that our sample of skeptics would have 95%-power to detect effect sizes $b > .16$ (see details in SI). For believers, both effects are non-significant (upvotes: $b = -0.01, p = 0.708$; bookmarks: $b = 0.04, p = 0.265$). We also conducted separate analyses based on a split of of each of the three continuous measures of beliefs, separating participants into four categories based on their scores (0–25, 26–50, 51–75, and 76–100). For each group, the model was: $\texttt{Outcome} \sim \texttt{Treatment} + (1|\texttt{Headline}) + (1|\texttt{Participant})$. Results (see SI) showed that the intervention had the strongest effects in the most skeptical group, and the weakest effect for the least skeptical group. These results confirm that modification was most effective for skeptical audiences: modifying headlines tailored them specifically for skeptics, while the original headlines were already appropriate for non-skeptics, requiring no further adjustment.

Heterogeneity across stimuli and individuals For each headline, we computed the effect of the Llama3 modification on skeptics' engagement as the difference in upvote (or bookmark) probability between modified and original versions. Additionally, our manipulation check with an independent sample of skeptics (see Methods) provided four measures of anticipation per original headline and its modified version: anticipated agreement, regret, positive emotions, and negative emotions. For each headline, we calculated the effect of Llama3 modification on these four anticipations. Figure 2C displays the correlation between the impact of Llama3 modification on anticipatory measures, and its impact on engagement measures. Consistent with information avoidance theory, we found strong correlations between impact on anticipated agreement and upvotes ($r = .70, p < .001$) and bookmarks ($r = .71, p < .001$); and between impact on anticipated regret and upvotes ($r = -.68$, $p < .001$) and bookmarks ($r = -.70$, $p < .001$). Correlations with anticipated generic emotions were in the expected direction but weaker and non-significant (all $|r| < .32$, all $p > .1$). Other heterogeneity analyses showed engagement results to be robust across age, sex, and education; globally robust across political ideology and partisanship; and stronger for participants who reported a lower interest in science news.

2.2 Experience

To test whether our approach may backfire, we asked skeptics to read one of the articles they had bookmarked and recorded three potential adverse outcomes. Because 10 of the 19 feed articles were climate-related and participants had to bookmark 10, we could ensure the article was about climate. After they read the article, we also measured the shift of their beliefs towards alignment with the scientific consensus.

Upvote reversals, bookmark regrets, and trust We tested whether participants reacted more negatively (more upvote reversals, more bookmark regret, less trust) when they read articles based on original vs. modified headline. For each outcome we ran four variants of the following model, one per prior belief measure: $\texttt{Outcome} \sim \texttt{Belief} + \texttt{Treatment} + \texttt{Belief} \times \texttt{Treatment} + (1|\texttt{Headline})$. Table 1 summarizes all models. Across all variants, prior beliefs consistently impacted reactions, which is unsurprising—the stronger the climate beliefs, the more positive the reactions to a climate science article. No credible main effect of treatment was found—Llama3-modified headlines did not backfire overall. However, we found some credible evidence for an belief-by-treatment interaction on bookmark regret, suggesting that skeptics may be more likely to regret bookmarking a climate article based on a modified headline. Post-hoc analyses (see SI) suggest this might be due to article negativity—modified headlines led skeptics to bookmark more negative or alarmist content than they otherwise would.

Shift towards alignment with the scientific consensus After reading the climate article, participants were reminded of their initial responses to the three continuous belief measures and could revise them. Belief change was calculated as post minus prior. We detected a significant shift towards alignment with the scientific consensus (i.e., a shift toward stronger agreement) for all three measures (climate change is

Predicted Outcome	Belief Measure	Belief Effect	Treatment Effect	Interaction Effect
Reversal of Upvote	(1)	$\mathbf{b = -0.18}$ $p =< .001$	$b = 0.01$ $p = .608$	$b = -0.03$ $p = .497$
	(2)	$\mathbf{b = -0.08}$ $p =< .001$	$b = 0.003$ $p = .908$	$b = 0.01$ $p = .544$
	(3)	$\mathbf{b = -0.08}$ $p =< .001$	$b = 0.002$ $p = .953$	$b = -0.01$ $p = .606$
	(4)	$\mathbf{b = -0.08}$ $p < .001$	$b = 0.004$ $p = .884$	$b = 0.01$ $p = 0.562$
Regret about Bookmark	(1)	$\mathbf{b = 0.25}$ $p < .001$	$b = 0.07$ $p = .288$	$\mathbf{b = -0.24}$ $p = .025$
	(2)	$\mathbf{b = 0.14}$ $p < .001$	$b = 0.02$ $p = .726$	$b = -0.06$ $p = .183$
	(3)	$\mathbf{b = 0.14}$ $p < .001$	$b = 0.03$ $p = .670$	$\mathbf{b = -0.12}$ $p = .006$
	(4)	$\mathbf{b = 0.16}$ $p < .001$	$b = 0.03$ $p = .670$	$\mathbf{b = -0.08}$ $p = .070$
Trust in Article	(1)	$\mathbf{b = 0.69}$ $p < .001$	$b = -0.04$ $p = .580$	$b = -0.08$ $p = .437$
	(2)	$\mathbf{b = 0.35}$ $p < .001$	$b = -0.04$ $p = .529$	$b = 0.006$ $p = .889$
	(3)	$\mathbf{b = 0.37}$ $p < .001$	$b = -0.03$ $p = .635$	$b = -0.06$ $p = .143$
	(4)	$\mathbf{b = 0.40}$ $p < .001$	$b = -0.03$ $p = .615$	$b = -0.04$ $p = .276$

Table 1: Main results about participants' self-rated experience after reading one bookmarked climate article. Belief measures: (1) Climate change is happening, binary, yes/no (2) Climate change is happening, continuous, 0–100 (3) Climate change is caused by human activity, continuous, 0–100 (4) Climate change is a significant threat, continuous, 0–100.

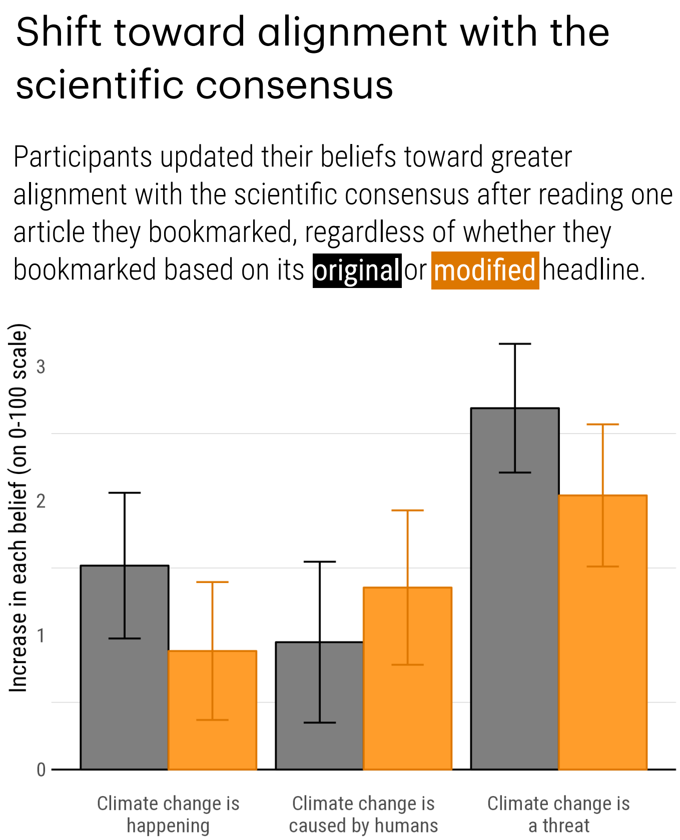


Figure 3: Participants' shift towards alignment with the scientific consensus at the end of the study, for three measures of belief. Error bars show the 95% confidence interval around the mean.

happening; $t(1998) = -6.3, p < .001$; climate change is caused by human activity; $t(1998) = -5.44, p < .001$; climate change is a threat; $t(1998) = -13.0, p < .001$. The average shift was small (from 0.8 to 2.7 points on a 0–100 scale, see Figure 3). We find no credible evidence that the persuasion effect was affected by the treatment in any of the three measures (see SI for detailed analysis, including moderation by partisanship, as well as secondary analyses using degrees of prior skepticism as predictor). These results are in line with previous work showing a positive effect of exposure to climate science, but we need to be careful to not over-interpret its observed size or over-estimate its duration, as it based on exposure to a single article, within a limited observation time frame, and could partly reflect experimenter demand.

3 Discussion

Generative AI appears to provide a viable pathway for increasing the reach and consumption of accurate climate information in communities that would otherwise resist it: directly, by increasing the likelihood of skeptics to read climate articles; and indirectly, by amplifying the spread of science-related content within skeptics' networks, which likely include other skeptics. Results were most pronounced among the most skeptical participants, which is essential since it may allow generative AI to address entrenched resistance, increasing engagement where it is the lowest yet the most needed. Moreover, the mechanisms behind these effects align well with theoretical expectations: headline modification influenced skeptics' anticipations, reducing their inclination to avoid information that might otherwise challenge their beliefs, and thus lowering the psychological barriers leading to information avoidance.

We acknowledge several limitations to the ecological validity of our study. First, while our social media simulation more closely approximates real-world engagement than typical survey experiments, it remains a simplified environment. In actual platforms, engagement is shaped by competing content, social dynamics, and opaque recommendation algorithms, all of which could influence intervention effectiveness. Second, our Prolific sample is self-selected and digitally literate, which may limit generalizability. Third, our findings are specific to the United States, where climate change is unusually polarized, and media habits as well as media coverage may be different from other countries. Taken together, these limitations caution against interpreting absolute engagement or belief shifts too strongly and underscore the need for future research in more ecologically valid and diverse settings. In parallel, the specific tailoring strategies used by the language model should be interpreted with caution, as they may not generalize across future models given the rapid pace of change in generative AI behavior. While generative AI will likely remain a viable tool for modifying headlines to achieve similar communicative goals, future models may rely on different strategies than those observed in this study.

Our approach is holistic in that it bypasses information avoidance by flexibly adapting to any framing that may be effective for a given article, rather than applying a uniform frame across all contents, and it has potential for automation and, thus, scalability. We emphasize 'potential' here, as our study still required substantial human oversight: nearly half of the modified headlines had to be discarded for failing our (admittedly high) quality standards. Finally, this approach aligns well with the incentives of social media and news organizations, as it drives greater engagement among audiences they do not typically reach, without reducing engagement within their usual audiences. This alignment may improve the chances of large-scale deployment and cooperation across the media industry, in the context of a growing role of AI in journalism, with recent work highlighting journalists' increasing willingness to integrate AI tools into news creation and distribution [66, 71, 72]. News outlets may prepare different versions of their headlines, to be circulated in parallel on social media platforms, without the need to identify climate skeptics—or they may be routed specifically to their intended audience when social media platforms are able to identify users who are more likely to be climate skeptics. We recognize that practical implementation of such approaches raises important ethical considerations. Identifying climate skeptics at scale could involve sensitive data and pose privacy risks, and even well-intentioned efforts to increase engagement may be perceived as manipulative. These concerns are particularly salient in light of growing public distrust and could backfire if skeptics interpret such interventions as evidence of ideological targeting.

4 Online Methods

4.1 Materials

Collection of climate science articles We used the Bing API service to search for and collect headlines of climate change related articles. Using the terms 'global warming' and 'climate change', we targeted trusted, mainstream news sites where the full text of each article was openly available. After scraping the article text, we applied several initial filters: we selected articles published no earlier than 2022, that mentioned 'climate change' or 'global warming' at least twice, contained numeric data, and were under 4,000 tokens (the input limit for Llama3). Following this, we used the Llama3 API to apply two additional filters. Llama3 was prompted to answer the following questions for each article: ‘Is this article primarily about some aspect of climate change/global warming? Return only Yes or No. Article: [article text].’ Then: ‘Does this article contain any scientific data, references to scientific studies, or feature scientists? Please analyze the following text and return only “Yes” or “No,” but nothing else: [article text].’ Finally, we manually reviewed every article for which Llama returned Yes and Yes, to ensure that these answers were correct.

Headline modification We used the following protocol to create modified headlines that would be less aversive to climate skeptics :

1. We provided Llama3 with the full text of the article embedded in the following prompt: ‘Create 5 headlines that must be true to the contents of the article, and are not inconsistent with beliefs of somebody who thinks that climate change is not happening - they need not be fully consistent, they can also take a neutral stance. Return the 5 titles, but nothing else. Article: [article text]’

2. We provided Llama3 with the five headlines it generated in the previous stage, plus the original headline, embedded in the following prompt: ‘Select the headline that is the least inconsistent with the beliefs of somebody who believes that climate change is not happening. Return only the selected headline. Headlines: [headline variants]’.

3. We provided Llama3 with the headlines it generated in the previous stage and the text of the article, embedded in the following prompt: ‘Is this headline true to the contents of the article, or is it misleading in any way? Return either Misleading or Not Misleading. Headline: [selected headline variant] Article: [article text]’.

4. We repeated this whole loop until Llama3 selected a headline that was not the original headline, and judged that headline to be not misleading.

Steps 2 and 3 were included in the process in light of results showing that self-evaluation can sometimes improve LLM outputs [73, 74], with the caveat that we cannot be sure these results apply to our particular use case. We followed a similar procedure to create headlines variants aimed at people who believed that climate change is happening but is not caused by human activity.

Manipulation check After generating modified headlines for 58 climate-related articles in the previous step, we conducted a manipulation check to insure that modified headlines did change the expectations of skeptics, in order to eliminate the headlines for which the manipulation was unsuccessful, prior to conducting our main experiment. This also allowed us to collect headline-level data for the analysis about heterogeneity across materials, reported in the results section. We recruited 302 participants from the US (158 identified as women, mean age = 45.9, sd= 14.3), using filters to target climate skeptics (all participants answered 'No' or Don't know' to the question 'Do you believe in climate change?'). Each participant saw a random subset of 30 headlines (10 unmodified, 10 modified for people who believe climate change is not happening, 10 modified for people who believe climate change is happening but not caused by human activity). They were instructed to 'Imagine you had to read the article with the following headline: [Headline]. When reading the article, how much do you expect to…' (1) feel positive emotions like enthusiastic, happy, excited, or cheerful; (2) feel negative emotions like angry, annoyed, afraid, or resentful; (3) agree with the contents of the article; (4) regret engaging with the article. All four ratings used a scale from 1 to 7. Overall, we found that headlines which were modified for people who believe climate change is not happening had the intended effect on the expectations of skeptics. They increased anticipated positive feelings ($b = 0.25, p < .001$) , increased anticipated agreement ($b = 0.41, p < .001$), decreased anticipated regret ($b = -0.26, p < .001$) and decreased anticipated negative emotions ($b = -0.12, p < .001$). However, we discarded eleven headlines for failing an individual manipulation check, since for this eleven headlines, the average effect of modification went in the wrong direction. As a result, after this manipulation check, we obtained a set of 47 articles. Finally, we observed that the headlines modified for people who believe climate change is not happening always outperformed in the manipulation check the headlines modified for people who believe climate change is happening but not caused by human activity. As a result, we decided to focus on the former in our main experiment.

Fact check In order to make sure that the modified headlines used in the experiment did not compromise factual integrity, we recruited a professional fact-checker who read all articles and their modified headlines. We asked the fact checker whether the headline was accurate and did not contain any untrue information (yes/no), whether the headline accurately represented the contents of the article (yes/no), and to further rate this accuracy on a scale from 0 to 5. We decided to adopt a conservatively high bar for factual integrity by using headlines for which the responses were yes, yes, and at least 4. This eliminated 19 articles which had passed the manipulation check, resulting in our final set of 28 articles.

Neutral headlines For the experiment, we also needed foil headlines unrelated to climate change. For this, we collected 62 science news headlines from nationalgeographic.com that did not contain references to climate change or global warming in their headline. These articles came from the 'animals', 'history and culture', and 'science' categories.

4.2 Participants

We collected data from 1999 participants (1033 identified as women, M = 45.9 years, SD = 15.8 years) using Prolific, an online survey platform commonly used in academic research to obtain access to a diverse pool of pre-screened participants, who are compensated for their time. We used a quota-sampling procedure so the sample was representative of age, sex, and political affiliation in the US population. In total, 997 participants took part in the control (original headlines) and 1002 in the experimental (modified headlines) condition. While 2083 people started the experiment, 3 did not consent to participate and 81 did not finish the experiment; these participants either produced no data or were excluded from the analysis.

4.3 Procedure

The median time for completion was 12 minutes. Participants did not have to complete the study in a single session, but Prolific rules required them to complete within 67 minutes of staring the study, or else be timed out. Participants were randomly assigned to the original or modified headlines treatments. Regardless of treatment, they went through the same experimental stages, detailed below.

Demographic questions and belief elicitation Full details of all questions are provided in the SI. We asked participants about their education level, age, sex, and partisan affiliation. We also asked whether they leaned democratic or conservative on economic issues and social issues, separately. We also recorded their preferences for reading news on the following topics: science, technology, US politics, international politics, culture, sports and entertainment. Then, as reported in the main text, we elicited their beliefs about climate change through five questions. Three of these questions used continuous 0–100 scales to measure belief that climate science is happening, belief that climate change is caused by human activity, and belief that climate change is a significant threat. Two other questions, taken from the Yale Climate Change Communication Center [69, 70] asked whether they believed in climate change; and whether, assuming that climate change is real, it is caused by human activity.

Upvotes Participants were shown a stylised social media interface displaying a feed of 20 posts. All these posts were headlines of news articles: 11 were randomly selected from the pool of climate science headlines, and 9 were foils, randomly selected from the neutral headline pool Participants were asked to either upvote or downvote each post. Here is how we described this task to the participants:

> *Welcome to the experiment! You will be participating in a social media simulation where you will see news articles as posts. You will have an upvote and downvote button next to each post, which will determine the ranking of the post. The upvote and downvote buttons function similarly to the voting system on a website called Reddit, where users can vote on content to determine its popularity and visibility. The higher the vote, the more people will see the post. Just like on Reddit, upvoting means you think the post is positively contributing to the community and downvoting means the opposite. This is how posts look like: On the right, you can upvote by clicking on the green and downvote by clicking on the red arrow. You will see 20 posts and must vote on each one. Once you voted on each post, the next button will appear and you can advance to the next page. Click on 'Next' to start the Simulation!*

After participants finished upvoting or downvoting all posts, they moved on to an attention check. They were presented with four headlines and had to identify the one which had not *not* appeared in their feed. This was then repeated a second time, with another set of four headlines.

Bookmarks Participants were presented again with the same feed of 20 headlines as in the Upvote phase, and were now asked to bookmark 10 of these articles for later reading, knowing that one of these decisions would be implemented in the next phase of the experiment. Here is how we described this to participants:

> *In the next section, you will have to read one of the articles. Now, this is your chance to say which ones you are, and which ones you are not interested in*

reading. You will see the same titles as you have seen before. This time, you can bookmark the ones you are the most interested in reading by clicking on the bookmark button: You will have to bookmark at least 10 posts, but please bookmark all that you would be interested in reading. We will select one article out of the bookmarked list that you will have to read after this stage.

Experience We randomly selected one of the articles participants bookmarked, with the constraint that this article had to be about climate science (there was always at least one such article because participants had to bookmark 10 articles out of 19, and only 9 articles were not about climate). Participants read the full version of this article which was selected from their bookmarks. After finishing it at their own pace, they are asked three questions:

1. *You read this article because you bookmarked it. How much do you regret bookmarking it?* [0-100 scale anchored at 'I regret it very much' and 'I do NOT regret it', with 'Neutral' written over the middle]

2. *Now that you know the contents of the article, would you upvote or downvote it on social media?* [Upvote?Downvote]

3. *How much do you trust that the information in this article is reliable?* [0-100 scale anchored at 'Not at all' and 'Completely', with 'Neutral' written over the middle]

Posterior beliefs The experiment ended by asking people the three continuous belief questions about climate change. Participants were shown the responses they gave at the start of the experiment, and were offered the opportunity to change these answers if they wished to.

4.4 Statistical analysis

We used linear mixed-effect regression models to estimate the effect of the treatment. We used linear models even when the outcome was binary (votes/bookmarks) as it is a preferred method to gain unbiased interpretable estimates of treatment effects in experimental settings [75]. We included random intercepts for both headlines and participants in the analyses of bookmarks and votes. For regret, credibility judgments, and belief change we included random intercepts for headlines only, as adding participant-level random effects would be redundant, since these measures were collected only once per participant. We z-scored continuous priors and all the continuous dependent variables (bookmark regret, credibility, belief update). Vote and bookmark were coded as 0: downvote/not bookmarked or 1: upvote/bookmarked. The treatment variable was coded as preregistered (0.5: original headline, -0.5: modified headline). Categorical prior belief variable was coded as per participants' response to the question: 0.5 = Believers (selected that climate change is caused by human activity); −0.5 = Skeptics (either selected that climate change is not happening or that it is happening but not caused by human activity); 0 = Selected 'Other' or 'Don't know'. Note that the preregistration did not specify how these two latter categories should be coded. To include them in the analysis without biasing the preregistered contrast, we assigned them a neutral value of 0.

4.5 Ethics

Ethical approval for this study was obtained by Tilburg University, under the reference: TSB_RP1173.

4.6 Data, code and material availability

Code for the analysis and all materials are available on the GitHub page of the project, along with all stimuli used in the experiment: `https://github.com/bencebago/news_personalization`. The preregistration, available at `https://aspredicted.org/wfvn-c2tg.pdf`, unfortunately contains a double typo stemming from a late terminology change, when we decided to write of climate 'skeptics' rather than climate 'deniers'. One key sentence reads:

> We will categorize participants based on their response to this question: 'Assuming global warming is happening, do you think it is…?' people who respond by clicking on the option that is caused mostly by human activity will be categorized as **skeptics**, people clicking on the other options that it does not happen or that it is not caused by human activity will be categorized as **deniers.**

But it should have been:

> We will categorize participants based on their response to this question: 'Assuming global warming is happening, do you think it is…?' people who respond by clicking on the option that is caused mostly by human activity will be categorized as **believers**, people clicking on the other options that it does not happen or that it is not caused by human activity will be categorized as **skeptics.**

Acknowledgments

JFB acknowledges support from grant ANR-19-PI3A-0004, grant ANR-17-EURE-0010, Grant ANR-22-CE26-0014-01, and the research foundation TSE-Partnership. PM acknowledges support from grant ANR-19-PI3A-0004 and the DesCartes project: the National Research Foundation, Prime Minister's Office, Singapore under its CREATE program. The authors gratefully acknowledge the help of Iyad Rahwan at the Center of Humans and Machines, Max Planck Institute for Human Development, for obtaining funding for data collection.

References

[1] Meng, Y., Broom, M. & Li, A. Impact of misinformation in the evolution of collective cooperation on networks. *Journal of the Royal Society Interface* **20**, 20230295 (2023).

[2] Kopp, C., Korb, K. B. & Mills, B. I. Information-theoretic models of deception: Modelling cooperation and diffusion in populations exposed to" fake news". *PloS one* **13**, e0207383 (2018).

[3] Levin, S. A. & Weber, E. U. Polarization and the psychology of collectives. *Perspectives on Psychological Science* **19**, 335–343 (2024).

[4] Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K. & Larson, H. J. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA **5**, 337–348.

[5] Abascal, M., Huang, T. J. & Tran, V. C. Intervening in anti-immigrant sentiments: The causal effects of factual information on attitudes toward immigration **697**, 174–191.

[6] Hornsey, M. J., Harris, E. A., Bain, P. G. & Fielding, K. S. Meta-analyses of the determinants and outcomes of belief in climate change **6**, 622–626.

[7] Van der Linden, S., Leiserowitz, A. & Maibach, E. The gateway belief model: A large-scale replication. *Journal of Environmental Psychology* **62**, 49–58 (2019).

[8] Lewandowsky, S. Climate change disinformation and how to combat it **42**, 1–21.

[9] Van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D. & Maibach, E. W. The scientific consensus on climate change as a gateway belief: Experimental evidence. *PloS one* **10**, e0118489 (2015).

[10] Goldberg, M. H., van der Linden, S., Maibach, E. & Leiserowitz, A. Discussing global warming leads to greater acceptance of climate science. *Proceedings of the National Academy of Sciences* **116**, 14804–14805 (2019).

[11] Većkalov, B. *et al.* A 27-country test of communicating the scientific consensus on climate change. *Nature Human Behaviour* 1 (2024).

[12] Bago, B., Rand, D. G. & Pennycook, G. Reasoning about climate change **2**, pgad100.

[13] Feldman, L. & Hart, P. S. Broadening exposure to climate change news? how framing and political orientation interact to influence selective exposure. *Journal of Communication* **68**, 503–524 (2018).

[14] Peterson, E. & Iyengar, S. Partisan gaps in political information and information-seeking behavior: Motivated reasoning or cheerleading? **65**, 133–147.

[15] Areni, C. S. Motivated reasoning and climate change: Comparing news sources, politicization, intensification, and qualification in denier versus believer subreddit comments. *Applied Cognitive Psychology* **38**, e4167 (2024).

[16] Newman, T. P., Nisbet, E. C. & Nisbet, M. C. Climate change, cultural cognition, and media effects: Worldviews drive news selectivity, biased processing, and polarized attitudes. *Public Understanding of Science* **27**, 985–1002 (2018).

[17] Bolin, J. L. & Hamilton, L. C. The news you choose: News media preferences amplify views on climate change. *Environmental Politics* **27**, 455–476 (2018).

[18] Wang, Y. & Jaidka, K. Confirmation bias in seeking climate information: Employing relative search volume to predict partisan climate opinions. *Social Science Computer Review* **42**, 4–24 (2024).

[19] Sweeny, K., Melnyk, D., Miller, W. & Shepperd, J. A. Information avoidance: Who, what, when, and why. *Review of general psychology* **14**, 340–353 (2010).

[20] Bénabou, R. & Tirole, J. Identity, morals, and taboos: Beliefs as assets. *The quarterly journal of economics* **126**, 805–855 (2011).

[21] Golman, R., Hagmann, D. & Loewenstein, G. Information avoidance. *Journal of economic literature* **55**, 96–135 (2017).

[22] Andersen, K., Toff, B. & Ytre-Arne, B. Introduction: What we (don't) know about news avoidance. *Journalism Studies* **25**, 1367–1384 (2024).

[23] Skovsgaard, M. & Andersen, K. Conceptualizing news avoidance: Towards a shared understanding of different causes and potential solutions. *Journalism studies* **21**, 459–476 (2020).

[24] Mangold, F., Schoch, D. & Stier, S. Ideological self-selection in online news exposure: Evidence from europe and the us. *Science Advances* **10**, eadg9287 (2024).

[25] Hickman, C. *et al.* Climate anxiety in children and young people and their beliefs about government responses to climate change: a global survey. *The Lancet Planetary Health* **5**, e863–e873 (2021).

[26] Bayes, R. & Druckman, J. N. Motivated reasoning and climate change. *Current Opinion in Behavioral Sciences* **42**, 27–35 (2021).

[27] Newman, N., Fletcher, R., Eddy, K., Robertson, C. T. & Nielsen, R. K. Reuters institute digital news report 2023. Tech. Rep., Reuters Institute for the Study of Journalism, Oxford, UK (2023). URL `https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023`.

[28] Chinn, S., Hart, P. S. & Soroka, S. Politicization and polarization in climate change news content, 1985-2017. *Science Communication* **42**, 112–129 (2020).

[29] Malka, A., Krosnick, J. A. & Langer, G. The association of knowledge with concern about global warming: Trusted information sources shape public thinking. *Risk Analysis* **29**, 633–647 (2009).

[30] Feldman, L., Myers, T. A., Hmielowski, J. D. & Leiserowitz, A. The mutual reinforcement of media selectivity and effects: Testing the reinforcing spirals framework in the context of global warming. *Journal of Communication* **64**, 590–611 (2014).

[31] Edenbrandt, A. K., Lagerkvist, C. J. & Nordström, J. Interested, indifferent or active information avoiders of carbon labels: Cognitive dissonance and ascription of responsibility as motivating factors. *Food Policy* **101**, 102036 (2021).

[32] d'Adda, G., Gao, Y., Golman, R. & Tavoni, M. Strategic information avoidance, belief manipulation and the effectiveness of green nudges. *Ecological Economics* **222**, 108191 (2024).

[33] Sharot, T. & Sunstein, C. R. How people decide what they want to know. *Nature Human Behaviour* **4**, 14–19 (2020).

[34] Dorison, C. A., Minson, J. A. & Rogers, T. Selective exposure partly relies on faulty affective forecasts. *Cognition* **188**, 98–107 (2019).

[35] Newman, N., Fletcher, R., Schulz, A., Simge, A. & Nielsen, R. K. Reuters institute digital news report 2020. Tech. Rep., Reuters Institute for the Study of Journalism, Oxford, UK (2020). URL `https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf`.

[36] Newman, N., Fletcher, R., Eddy, K., Robertson, C. T. & Nielsen, R. K. Reuters institute digital news report 2022. Tech. Rep., Reuters Institute for the Study of Journalism, Oxford, UK (2022). URL `https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022`.

[37] González-Bailón, S. *et al.* Asymmetric ideological segregation in exposure to political news on facebook. *Science* **381**, 392–398 (2023).

[38] Iyengar, S. & Hahn, K. S. Red media, blue media: Evidence of ideological selectivity in media use. *Journal of communication* **59**, 19–39 (2009).

[39] Petersen, A. M., Vincent, E. M. & Westerling, A. L. Discrepancy in scientific authority and media visibility of climate change scientists and contrarians. *Nature communications* **10**, 3502 (2019).

[40] Koehler, D. J. Can journalistic "false balance" distort public perception of consensus in expert opinion? *Journal of Experimental Psychology: Applied* **22**, 24 (2016).

[41] Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting covid-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological science* **31**, 770–780 (2020).

[42] Pennycook, G. *et al.* Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021).

[43] Frimer, J. A., Skitka, L. J. & Motyl, M. Liberals and conservatives are similarly motivated to avoid exposure to one another's opinions. *Journal of Experimental Social Psychology* **72**, 1–12 (2017).

[44] Nyhan, B. *et al.* Like-minded sources on facebook are prevalent but not polarizing. *Nature* 1–8 (2023).

[45] Robertson, R. E. *et al.* Users choose to engage with more partisan news than they are exposed to on google search. *Nature* 1–7 (2023).

[46] Janét, K., Richards, O. & Landrum, A. R. Headline format influences evaluation of, but not engagement with, environmental news. *Journalism Practice* **16**, 35–55 (2022).

[47] Feldman, L. & Hart, P. S. Upping the ante? the effects of "emergency" and "crisis" framing in climate change news. *Climatic Change* **169**, 10 (2021).

[48] Chapman, D. A., Lickel, B. & Markowitz, E. M. Reassessing emotion in climate change communication. *Nature Climate Change* **7**, 850–852 (2017).

[49] Robertson, C. E. *et al.* Negativity drives online news consumption. *Nature Human Behaviour* **7**, 812–822 (2023).

[50] Zhang, M. *et al.* Negative news headlines are more attractive: negativity bias in online news reading and sharing. *Current Psychology* 1–14 (2024).

[51] Vlasceanu, M. *et al.* Addressing climate change with behavioral science: A global intervention tournament in 63 countries. *Science advances* **10**, eadj5778 (2024).

[52] Maertens, R., Anseel, F. & van der Linden, S. Combatting climate change misinformation: Evidence for longevity of inoculation and consensus messaging effects. *Journal of Environmental Psychology* **70**, 101455 (2020).

[53] Nyhan, B., Porter, E. & Wood, T. J. Time and skeptical opinion content erode the effects of science coverage on climate beliefs and attitudes. *Proceedings of the National Academy of Sciences* **119**, e2122069119 (2022).

[54] Dubey, A. *et al.* The llama 3 herd of models. *CoRR* **abs/2407.21783** (2024). URL `https://doi.org/10.48550/arXiv.2407.21783`. 2407.21783.

[55] Capraro, V. *et al.* The impact of generative artificial intelligence on socioeconomic inequalities and policy making. *PNAS Nexus* (2024).

[56] Spitale, G., Biller-Andorno, N. & Germani, F. Ai model gpt-3 (dis) informs us better than humans. *Science Advances* **9**, eadh1850 (2023).

[57] Kidd, C. & Birhane, A. How ai can distort human beliefs. *Science* **380**, 1222–1223 (2023).

[58] Shin, S. Y. & Lee, J. The effect of deepfake video on news credibility and corrective influence of cost-based knowledge about deepfakes. *Digital Journalism* **10**, 412–432 (2022).

[59] Simchon, A., Edwards, M. & Lewandowsky, S. The persuasive effects of political microtargeting in the age of generative artificial intelligence. *PNAS nexus* **3**, pgae035 (2024).

[60] Augenstein, I. *et al.* Factuality challenges in the era of large language models and opportunities for fact-checking. *Nature Machine Intelligence* **6**, 852–863 (2024).

[61] Costello, T. H., Pennycook, G. & Rand, D. G. Durably reducing conspiracy beliefs through dialogues with ai. *Science* **385**, eadq1814 (2024).

[62] Bago, B. & Bonnefon, J.-F. Generative ai as a tool for truth. *Science* **385**, 1164–1165 (2024).

[63] Tessler, M. H. *et al.* Ai can help humans find common ground in democratic deliberation. *Science* **386**, eadq2852 (2024).

[64] Dörr, K. N. Mapping the field of algorithmic journalism. *Digital journalism* **4**, 700–722 (2016).

[65] Diakopoulos, N. *Automating the news: How algorithms are rewriting the media* (Harvard University Press, 2019).

[66] Cools, H. & Diakopoulos, N. Uses of generative ai in the newsroom: mapping journalists' perceptions of perils and possibilities. *Journalism Practice* 1–19 (2024).

[67] Guenther, L., Kunert, J. & Goodwin, B. "away from this duty of chronicler and towards the unicorn": How german science journalists assess their future with (generative) artificial intelligence. *Journal of Science Communication* **24**, A06 (2025).

[68] Opdahl, A. L. *et al.* Trustworthy journalism through ai. *Data & Knowledge Engineering* **146**, 102182 (2023).

[69] Howe, P. D., Mildenberger, M., Marlon, J. R. & Leiserowitz, A. Geographic variation in opinions on climate change at state and local scales in the usa. *Nature climate change* **5**, 596–603 (2015).

[70] Marlon, J. R. *et al.* Change in us state-level public opinion about climate change: 2008–2020. *Environmental research letters* **17**, 124046 (2022).

[71] Nishal, S. & Diakopoulos, N. Envisioning the applications and implications of generative ai for news media. *arXiv preprint arXiv:2402.18835* (2024).

[72] Nishal, S., Sinchai, J. & Diakopoulos, N. Understanding practices around computational news discovery tools in the domain of science journalism. *Proceedings of the ACM on Human-Computer Interaction* **8**, 1–36 (2024).

[73] Ren, J., Zhao, Y., Vu, T., Liu, P. J. & Lakshminarayanan, B. Self-evaluation improves selective generation in large language models. In *Proceedings on Machine Learning Research*, 49–64 (PMLR, 2023).

[74] Kadavath, S. *et al.* Language models (mostly) know what they know. *arXiv preprint arXiv:2207.05221* (2022).

[75] Gomila, R. Logistic or linear? estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General* **150**, 700 (2021).