Supporting User Information Processing Through Large Language Models Within the Political Sphere

Neeley Pate
Department of Computer Science
University of Rochester
Rochester, NY, USA
npate@ur.rochester.edu

Abstract

How do we support information within the political domain? By incorporating personalization and guardrails, large language model (LLM) systems can be leveraged to support navigation through the information ecosystem. In this work, I outline a proposal for designing LLM systems within two areas: mitigating misinformation belief and bolstering information processing. These tools aim to draw from theories within persuasion, information processing, and motivated reasoning to ultimately speak to the end user and nudge them to pursue accuracy when presented with information. These interventions will not only extend research within these relevant domains, but also support an individual's ability to interpret information provided.

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **HCI design and evaluation methods**.

Keywords

Large language models, misinformation, motivated reasoning, personalization, information processing

ACM Reference Format:
Neeley Pate. 2025. Supporting User Information Processing Through Large Language Models Within the Political Sphere. In *33rd ACM Conference on User Modeling, Adaptation and Personalization (UMAP '25), June 16–19, 2025, New York City, NY, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3699682.3727567

1 Introduction

Information processing, or how an individual receives and interprets a given piece of information, is extremely unique to the individual, their current state of being, and their environment [1, 20, 21]. The acceptance or rejection of certain information can also lead to belief in misinformation, or false information / information meant to mislead. Information processing and misinformation belief pose challenges within space of political science, where individuals may rely on partisan motivated reasoning for belief formation [2, 3, 19], introducing the potential for inaccurate perceptions and misinformation belief.

How do we support information within the political domain? Prior work has shown promising results at the intersection of large language model (LLM) interventions and the processing of political information [8]. Through this proposal, I aim to expand this research by focusing on interventions that mitigate misinformation belief and support accuracy in information processing. Utilizing LLMs, I outline the creation of specialized systems that accomplish these tasks. These systems will emphasize the use of guardrails to protect end users interacting with political information and incorporate personalization to meet the end user at their level. First, I aim to understand the current state of LLMs regarding misinformation, and how we can elevate LLM performance to combat misinformed beliefs. Second, I want to expand beyond misinformation to focus on supporting writers in reaching a broad audience, relying on theories of information processing and motivated reasoning. As LLMs become more accessible and prevalent in society, it is important to design positive interventions. These interventions will provide evidence of positive use cases for LLM systems and support individuals within an information ecosystem. While this proposal is based within the subtopic of political information (and misinformation), this work can easily be extended to other domains, such as public health or climate change, where misinformation may be encountered and navigating information may be difficult.

2 Related Work

2.1 Misinformation and LLMs

With the growth in pervasiveness of LLMs, concern of their ability to spread misinformation has also increased [10, 16]. While LLMs pose the ability to spread misinformation through hallucinations or limited real-time information, prior works have also shown their ability to combat misinformation through detection efforts [15, 16]. However, tackling misinformation through vanilla LLM fact-checking is not always effective [9]. One aspect of misinformation that makes it difficult to combat is familiarity associated with the misinformation [21]. Alternatively, authors have also expressed concern about the opposite result, where individuals are distrusting of truths [12]. Thus, both falsehood belief mitigation and truthhood belief support are crucial areas to explore within the misinformation intervention domain.

LLMs, while offering negative consequences within the misinformation sphere, also serve to help combat misinformation. Prior works have emphasized LLM persuasion [4] and have shown that simple conversations with LLM agents can help reduce belief in conspiracy theories that are maintained after initial exposure [8].

UMAP '25, New York City, NY, USA

ACM ISBN 979-8-4007-1313-2/25/06
https://doi.org/10.1145/3699682.3727567

Thus, it is important to explore LLM systems to fight misinformation, while also understanding that these systems need guardrails to protect end users.

2.2 Information Processing and Motivated Reasoning

Information processing models and theories aim at understanding how individuals perceive a message and what external conditions lead them to certain conclusions. Prior works emphasizes dual process models for understanding how individuals are persuaded (or not persuaded) to receive a given message [18]. While early persuasion begins with Aristotle [7], Petty and Cacioppo [22] propose the Elaboration Likelihood Model (ELM), which accounts for the peripheral route, or inference based on heuristics, and the central route, which emphasizes analyzing the strength of the argument. The authors suggest heuristics take over when an individual is unlikely to engage with a message. Motivated reasoning, proposed by Kunda [14], which can take on either directed or accurate dimension, ties in well to ELM and is easily extended to the political domain [2, 3, 11]. When individuals pursue directional motivated reasoning, they are looking to reach a specific conclusion and will utilize certain heuristics such as partisan alignment to reach that conclusion. When individuals pursue accurate reasoning, they are motivated to form accurate or easy to justify opinions without pursuit of a specific opinion. These two models both emphasize the dichotomy of engaging with the information provided and pursuing accurate opinions and disengaging with information or relying on heuristics to form pointed conclusions. Thus, the issue of prompting users to pursue the central route or accurate reasoning when engaging with a piece is a crucial one to solve, especially within the political landscape, where trust in media is highly partisan but generally on the decline across all parties [13].

3 Research Questions

My proposed research plan strives to support knowledge and information within the political sphere utilizing supported LLMs, and aims to answer the following questions:

- **Research Question 1**: How do current off-the-shelf LLM models perform in countering misinformation?
- **Research Question 2**: How can researchers design support for LLM systems to help combat political misinformation?
- **Research Question 3**: How do individuals respond to an LLM intervention designed to support the ground truth?
- **Research Question 4**: What natural trends arise to support accuracy reasoning and reader engagement?
- **Research Question 5**: How can LLM systems support information creators (journalists, academics, etc.) in engaging their audience and prompting critical thinking?
- **Research Question 6**: Do we see the intended effects of the LLM intervention downstream in reader engagement and accuracy reasoning?

4 Prior and Current Work

My current research efforts have focused on addressing the first two research questions – assessing off-the-shelf LLM performance in countering misinformation and the design of a system to combat political misinformation.

4.1 Evaluating Base LLM Performance in Countering Misinformation

In order to better support LLMs in countering misinformation, it is important first to understand the specific pitfalls of these models without any modifications. For this analysis, we utilized three models – ChatGPT 4o-mini, Claude 3.5-Haiku, and Gemini 1.5-flash – and had them generate statements that either proved or disproved a given prompt. To ensure robustness, the model was fed statements of varying ground truth labels from different broad political topics. The models first generate source material from a provided list of approved sources, then asked to summarize the information in a persuasive manner. Each model produced over 1750 final outputs for analysis.

This work resulted in several major findings, covered extensively in our preprint [24]. First, LLMs are unable to draw from real, credible source material. Utilizing GNews, we input the source and headline provided by the model after the source generation phase and compare the top 10 results of real articles returned by the API request. Figure 1 shows these results. We find that these models are providing headline and source pairs that do not exist or failing to provide both the headline and source in the majority of cases. Second, LLMs often impose their own bias on the source selection, either straying from the input list of sources (Gemini and Claude) or disproportionately selecting sources within the left-leaning category over other categories (ChatGPT and Claude). However, Gemini's relative fairness is attributed to the model generally regurgitating the list fed into the initial prompt; simply imposing no preference at all. Figure 2 shows the distribution of source appearance after source generation, and displays the 10 news sources selected for this experiment. This finding validates prior works that share similar findings of LLM biases [17, 25]. These findings suggest that LLMs alone should not be utilized for combating misinformation and individuals utilizing LLMs should be aware of the potential bias they are encountering. This work emphasizes the importance of the selection of the model and guardrail for specialized system design.

4.2 Designing an LLM System to Mitigate Misinformation Belief

The prior work highlights that when designing an LLM system to mitigate misinformation belief, two crucial areas to support are source selection and rhetorical styling. Because LLMs both fail to draw from existing sources directly and impose a bias on their selections, it is imperative to provide LLMs with sources directly and incorporate individual preference into source selection. Incorporating individual preference also extends well to rhetorical styling, since different attributes of a message may speak more to certain individuals than others. Thus, these two areas are emphasized in the creation of the LLM system.

The LLM system designed for this purpose relies on an individual's data in order to personalize responses. To generate responses, the LLM is given a statement and the factual ground truth of the statement. The model draws upon external sources to support its

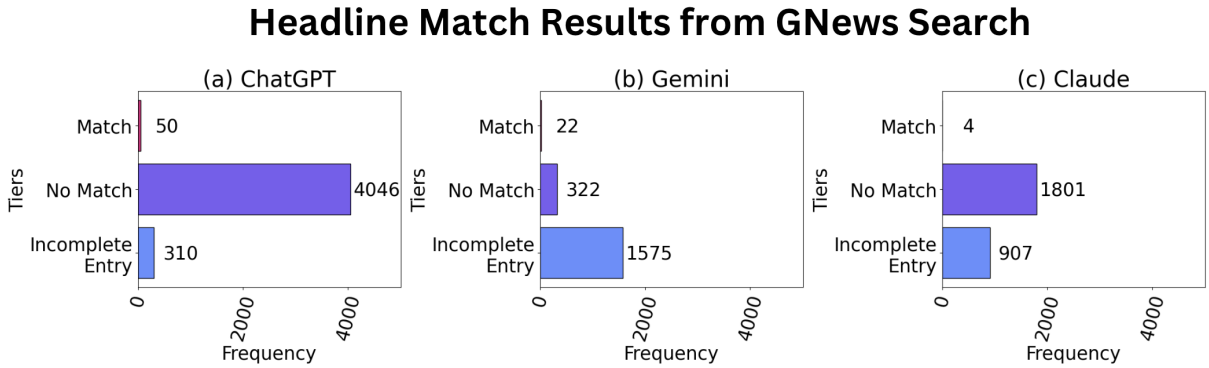

Figure 1: Frequency of Headline Matches Across Models, Based on GNews Search Results

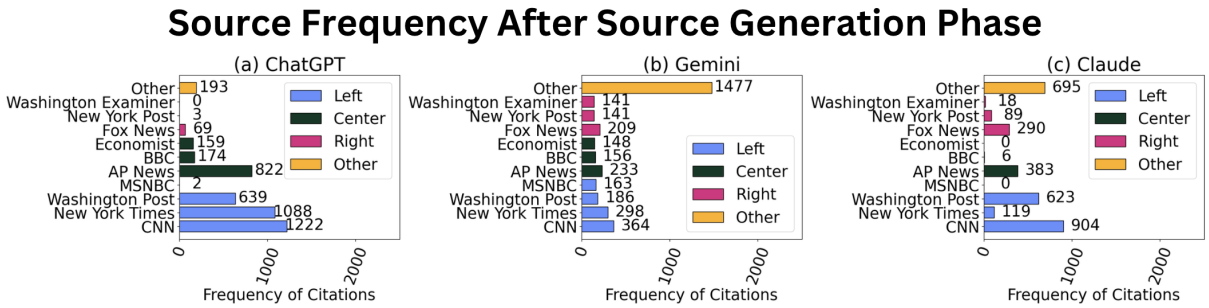

Figure 2: Frequency of Source Appearance After the Source Generation Phase Across Models

argument, and then reframes the summary of the source content in a particular rhetorical style. Both the sources and the style are determined by a predictive model based on data collected from other human participants. The backbone of the LLM model is ChatGPT 3.5-turbo.

This system was then evaluated by an LLM following the generation process previously described. The ChatGPT 4o-mini evaluator model scores the outputs and provides reasoning for the score. We also choose to look at source frequency and cosine similarity scores across outputs to verify the effect of personalization.

We find that the LLM system works well for its intended purpose, with the LLM evaluator rating most outputs as either a 4 or 5 out of 5. We also find a more balanced representation of news sources within the final output when compared to base ChatGPT 4o-mini, shown in Figure 3. It is important to note that while the pipeline has a significant amount of "Other", this often stems from the model citing sources within the original source material rather than source fabrication. Finally, we find that the pipeline output is more varied due to differences in source and rhetorical style (Table 1). These findings imply the pipeline is more grounded in real source material, exhibiting less of a bias than the base model, and incorporating personalization to have more diverse responses.

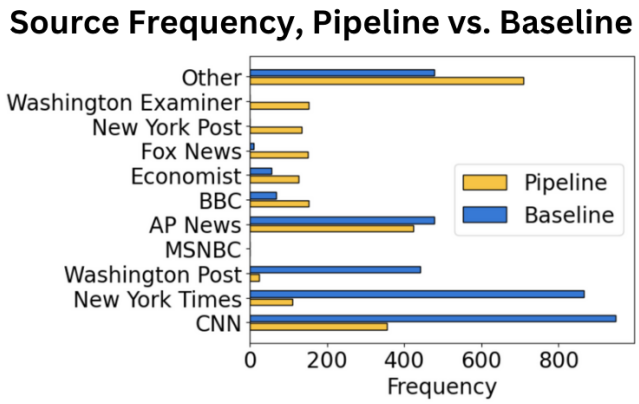

Figure 3: Frequency of Source Reference in Outputs

5 Future Work

Building off of the prior findings, I plan to host an empirical study evaluating the pipeline for combating misinformation and expanding into supporting reader reasoning.

5.1 Empirical Evaluation of the LLM System for Mitigating Misinformation Belief

In order to measure the effects of the intervention, human participants will enter a setting in which they are exposed to both political

Statement Theme	Pipeline - Baseline Difference
U.S. Oil Production	**-0.073**
Electric Vehicles	**-0.124**
Gas Price	0.022
Population Increase	**-0.164**
Population Trends	**-0.132**
U.S. Birth Rate	**-0.06**
Inflation Effects	**-0.102**
Price of Used Cars	**-0.018**
Airline Ticket Price	0.028
Illegal Immigrant Population	**-0.136**
Humanitarian Parole	**-0.12**
NYC Illegal Immigrants	**-0.068**

Table 1: Average Cosine Similarity Based on Statement ID Between Pipeline and Baseline

statements and the intervention (in the case condition). Within the study, the effect of the bot will be captured through belief updates and network updates compared between the case, where only humans enter the study, and the case, where each individual would have a personalized bot response. When participants enter the study, they will provide their demographic and personality information. Then, they will go through a series of three rounds, each with three stages. Each round contains one of the three statements within that topic. In stage 1, participants will be shown the statement and asked on a Likert scale whether they agree or disagree with the statement. In stage 2, participants will be shown the responses of a few others within the experiment and asked if they would like to update their opinion. In the control, all other responses shown will be humans, while in the case, one of the responses will be the personalized bot. In stage 3, participants will be shown more peers and asked to update their network by either "following" or "unfollowing" them. In the control, all peers will be other participants while in the case one of the peers will be the personalized bot for that individual. After completing three rounds, users will fill out an exit survey and will be debriefed regarding which statements were true or false within the experiment. Ideally, we would find the bot has an effect on the case population, pushing them towards the truth between stage 1 and stage 2. We would also hope to find that individuals in the case choose to follow the bot, highlighting its perceived credibility.

5.2 Measuring Trends in Inducing Accuracy Reasoning and Reader Engagement

Extending message framing beyond simple rhetorical styling, it has been shown that the mindset of an individual when receiving a message can greatly affect their means to process and retain the message, as well as the conclusion they reach after interacting with the message [6, 11, 26]. While most prior works within information processing and motivated reasoning induce the reader into a current state prior to their receiving of the message [2, 3, 11], this is not the case for real-world scenarios where individuals may receive a passage based on social media algorithms or internet browsing. As a result, we must understand how factors of information retention and accuracy reasoning emerge naturally, under no prior priming.

In order to explore this phenomenon, I propose enlisting both experts and non-experts to write passages regarding political topics and priming *the writer* rather than the readers. Priming methods would follow examples in prior works, such as emphasizing reader self-reflection, non-partisan aspects of issues (cross-partisan agreement for example [3]), prompting the reader to pursue accuracy reasoning, etc. These samples would then be ingested by readers who receive no priming cues at all, and then will be asked questions regarding their perception of the issue to gauge how much information they retained and whether they pursued the accuracy reasoning dimension. One common example of this would be having an individual read a piece related to the economy and then ask their perception of the current economy, as well as write a summary of the article they read several minutes later [6]. From the large body of written samples and annotations, we can unearth trends in which aspects of writing work well to support accuracy and message retainment. Based on prior works, I would expect that writing that invokes personal relevance, strong presentation of arguments, personal responsibility, and inaction vs. action framing will naturally lead to accuracy pursuit in information processing, for example, though other trends may appear as well [23].

5.3 Designing an LLM System to Support Written Work

From the phenomena that naturally emerge from these annotations, I will design an LLM system that will emphasize the trends that lead to accuracy reading and message retainment within the reader. Ideally, the tool will initially prime the writer of the piece with a given goal relevant to the piece they are writing (for example, personal relevance appeal, cross-partisan agreement, etc.), then evaluate the work as the journalist writes the piece. The LLM system will include a database of recent political information to support current LLM knowledge and writing samples that draw out the desired themes that inspire accuracy reasoning. The system will point out arguments that lack support, suggest places for calls of self-reflection and accountability, and other thematic improvements based on the trends from the prior study described. This work would differ from other writing tools (for example, Grammarly [1]) because of the support of current political information and the specific design for invoking accuracy reasoning. Because the proposed work is an architecture for outfitting an LLM, different LLM models could be placed within the system. Testing LLM performance within the system across models would influence the final model used in the empirical study. This work should ultimately help writers create more compelling pieces by keeping the reader at the forefront of the writing process.

5.4 Empirical Evaluation of the LLM System for Reader Engagement and Accuracy Reasoning

After creating the specialized LLM system, the system will be evaluated by both experts and non-experts. In the control, the writers

[1] https://www.grammarly.com/

will not be exposed to the system, while in the case the writers will work with the system to generate a piece. Written pieces will be about certain political topics, and writers will be given a certain amount of time to create the piece. After writing these pieces, they will be given to a general audience to be read. Readers will rate the piece on various metrics such as strength of the argument, personal relevance, etc. Readers in both the case and control will also be given a "fluff" piece, for example about a movie review (following prior work [5]), and provide the same ratings and feedback. Readers will then select which piece they would like to summarize and write a short summary. This will give some indication of reader engagement and retention of the political pieces depending on how often they were selected over the "fluff". These summaries could in turn be provided *again* to another set of readers who go through the same rating and summarization process to further measure the downstream effects of the original written piece, similar to a Telephone Game style experiment [5, 6]. Ideally, the outcome of this would find that the pieces written with the intervention involved will be rated more highly and engage the users through prompting them to write the summary about the piece, and that the summary would contain relevant information (verifying information retention). Even more interestingly, the results could show that individuals who received the summary rather than the original piece also experience this effect to some extent.

6 Dissertation Specifics

I am currently a second year PhD student at the University of Rochester, advised by Dr. Ehsan Hoque. I aim to finish my PhD in the spring of 2028. While my dissertation is currently underway, I have not yet proposed my topic for my thesis, though the current idea is outlined here as future work. While I plan to conduct the proposed future work which includes human evaluation, the lack of human evaluation within the context of this proposal can be considered a limitation at this time.

In the long term, I hope upon graduation to receive a teaching position within a computer science department at a university. Along with holding teaching roles (and hopefully teaching courses at the intersection of ethics and computer science) I hope to continue to conduct research within the human-computer interaction space.

References

[1] Saifuddin Ahmed and Han Wei Tan. 2022. Personality and perspicacity: Role of personality traits and cognitive ability in political misinformation discernment and sharing behavior. *Personality and Individual Differences* 196 (2022), 111747. doi:10.1016/j.paid.2022.111747

[2] Robin Bayes, James N. Druckman, Avery Goods, and Daniel C. Molden. 2020. When and How Different Motives Can Drive Motivated Political Reasoning. *Political Psychology* 41, 5 (Oct. 2020), 1031–1052. doi:10.1111/pops.12663

[3] Toby Bolsen, James N. Druckman, and Fay Lomax Cook. 2014. The Influence of Partisan Motivated Reasoning on Public Opinion. *Political Behavior* 36, 2 (June 2014), 235–262. doi:10.1007/s11109-013-9238-0

[4] Simon Martin Breum, Daniel Vædele Egdal, Victor Gram Mortensen, Anders Giovanni Møller, and Luca Maria Aiello. 2024. The Persuasive Power of Large Language Models. *Proceedings of the International AAAI Conference on Web and Social Media* 18 (May 2024), 152–163. doi:10.1609/icwsm.v18i1.31304

[5] Troels Bøggild, Lene Aarøe, and Michael Bang Petersen. 2021. Citizens as Complicits: Distrust in Politicians and Biased Social Dissemination of Political Information. *American Political Science Review* 115, 1 (2021), 269–285. doi:10.1017/S0003055420000805

[6] Taylor N. Carlson. 2019. Through the Grapevine: Informational Consequences of Interpersonal Political Communication. *American Political Science Review* 113, 2 (May 2019), 325–339. doi:10.1017/S000305541900008X

[7] Edward Meredith Cope et al. 1877. *The rhetoric of Aristotle*. Vol. 2. University Press.

[8] Thomas H. Costello, Gordon Pennycook, and David G. Rand. 2024. Durably reducing conspiracy beliefs through dialogues with AI. *Science* 385, 6714 (Sept. 2024), eadq1814. doi:10.1126/science.adq1814

[9] Matthew R. DeVerna, Harry Yaojun Yan, Kai-Cheng Yang, and Filippo Menczer. 2024. Fact-checking information from large language models can decrease headline discernment. *Proceedings of the National Academy of Sciences* 121, 50 (Dec. 2024), e2322823121. doi:10.1073/pnas.2322823121

[10] Maryanne Garry, Way Ming Chan, Jeffrey Foster, and Linda A. Henkel. 2024. Large language models (LLMs) and the institutionalization of misinformation. *Trends in Cognitive Sciences* 28, 12 (Dec. 2024), 1078–1088. doi:10.1016/j.tics.2024.08.007

[11] Eric Groenendyk and Yanna Krupnikov. 2021. What Motivates Reasoning? A Theory of Goal-Dependent Political Evaluation. *American Journal of Political Science* 65, 1 (Jan. 2021), 180–196. doi:10.1111/ajps.12562

[12] Brian Guay, Adam J. Berinsky, Gordon Pennycook, and David Rand. 2023. How to think about whether misinformation interventions work. *Nature Human Behaviour* 7, 8 (Aug. 2023), 1231–1233. doi:10.1038/s41562-023-01667-w

[13] Gallup Inc. 2023. Media Confidence in U.S. Matches 2016 Record Low. https://news.gallup.com/poll/512861/media-confidence-matches-2016-record-low.aspx

[14] Ziva Kunda. 1990. The case for motivated reasoning. *Psychological Bulletin* 108, 3 (1990), 480–498. doi:10.1037/0033-2909.108.3.480

[15] Markus Leippold, Saeid Ashraf Vaghefi, Dominik Stammbach, Veruska Muccione, Julia Bingler, Jingwei Ni, Chiara Colesanti Senni, Tobias Wekhof, Tobias Schimanski, Glen Gostlow, Tingyu Yu, Juerg Luterbacher, and Christian Huggel. 2025. Automated fact-checking of climate claims with large language models. *npj Climate Action* 4, 1 (Feb. 2025), 17. doi:10.1038/s44168-025-00215-8

[16] Aiwei Liu, Qiang Sheng, and Xuming Hu. 2024. Preventing and Detecting Misinformation Generated by Large Language Models. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Washington DC, USA) *(SIGIR '24)*. Association for Computing Machinery, New York, NY, USA, 3001–3004. doi:10.1145/3626772.3661377

[17] Ruibo Liu, Chenyan Jia, Jason Wei, Guangxuan Xu, and Soroush Vosoughi. 2022. Quantifying and alleviating political bias in language models. *Artificial Intelligence* 304 (March 2022), 103654. doi:10.1016/j.artint.2021.103654

[18] Andrew Luttrell. 2018. *Dual Process Models of Persuasion*. Oxford University Press. doi:10.1093/acrefore/9780190236557.013.319

[19] Kevin J. Mullinix. 2016. Partisanship and Preference Formation: Competing Motivations, Elite Polarization, and Issue Importance. *Political Behavior* 38, 2 (June 2016), 383–411. doi:10.1007/s11109-015-9318-4

[20] Gordon Pennycook and David G. Rand. 2019. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 188 (2019), 39–50. doi:10.1016/j.cognition.2018.06.011 The Cognitive Science of Political Thought.

[21] Gordon Pennycook and David G. Rand. 2021. The Psychology of Fake News. *Trends in Cognitive Sciences* 25, 5 (2021), 388–402. doi:10.1016/j.tics.2021.02.007

[22] Richard E. Petty. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change* (1st ed ed.). Springer New York, New York.

[23] Richard E. Petty and Jon A. Krosnick. 2014. *Attitude strength: antecedents and consequences*. Psychology Press, Taylor & Francis Group, New York London.

[24] Adiba M. Proma, Neeley Pate, James Druckman, Gourab Ghoshal, Hangfeng He, and Ehsan Hoque. 2025. An Empirical Analysis of LLMs for Countering Misinformation. https://arxiv.org/pdf/2503.01902

[25] Luca Rettenberger, Markus Reischl, and Mark Schutera. 2025. Assessing political bias in large language models. *Journal of Computational Social Science* 8, 2 (May 2025), 42. doi:10.1007/s42001-025-00376-w

[26] Norbert Schwarz and Gerald L. Clore. 1983. Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology* 45, 3 (Sept. 1983), 513–523. doi:10.1037/0022-3514.45.3.513