CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Master's thesis

2021 **Mikayel Harutyunyan**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Mikayel Harutyunyan

Fighting Fake News with Accuracy: Dual Processing Perspective

Master's thesis

Prague 2021

Author: Mikayel Harutyunyan

Supervisor: doc. PhDr. Julie Chytilová, Ph.D.

Academic Year: 2020/2021

Bibliographic note

Harutyunyan, M. (2021) *Fighting Fake News with Accuracy: Dual Processing Perspective.* Master's Thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies.

Abstract

The phenomenon of "fake news", or misleading online content, is increasingly worrisome due to its large-scale socio-economic impact. Researchers and practitioners attempted to understand what drives the virality and believability of fake news and how to reduce its influence. This research aims to shed light on these questions. Building upon a theoretical account positing that people share fake news because they simply fail to engage in deliberate thinking, we designed an accuracy prompt intervention to encourage people to think effortfully. In a pre-registered study conducted via Prolific (N = 520), we find limited evidence supporting accuracy prompts stylized as warning labels, but only for increasing sharing discernment in true, not fake news. The veracity of news articles does not impact sharing intentions, despite having a sizeable effect on accuracy judgments. This and other findings support the dual processing theory of cognition in the context of fake news. Predispositions towards more intuitive thinking increased belief in fake news and higher distrust in true news. Conversely, a better ability to engage in effortful thinking increases truth discernment. In addition, confirmation bias decreases truth discernment and increases sharing intentions. Politically concordant true headlines are believed and shared more, but the effect differs on the individual (negative) and county (positive) level in case of fake news. Overall, our study provides a theoretical foundation for scalable dual processing-based interventions that social media could implement to fight online misinformation.

Abstrakt

Falešné zprávy neboli zavádějící či nepravdivý online obsah je fenomén, který je znepokojivý svým negativním potenciálním socio-ekonomickým dopadem. Výzkumníci i praktici se pokoušeli porozumět, co pohání jejich důvěryhodnost a šíření těchto falešných zpráv a jakými, způsoby lze omezit vliv falešných zpráv. Tato akademická práce si klade za cíl osvětlit některé teorie. Práce je postavena na předpokladu, že lidé sdílejí falešné zprávy, protože nezapojují dostatečně kritické myšlení. Navrhli jsme tzv. „accuracy prompt” intervenci, která podporuje kritické myšlení. Ve předregistrované studii realizované na platformě Prolific (N = 520) jsme našli omezené důkazy podporující „accuracy prompts” intervence, které jsou stylizované jako „výstražné nálepky”, ale pouze pro podporu šíření pravdivých zpráv, nikoliv však zpráv falešných. Pravdivost zpráv neovlivňuje záměr sdílení navzdory efektu na úsudky posuzující pravdivost. Tato a další zjištění podporují kognitivní teorii duálního zpracovávání v kontextu falešných zpráv. Predispozice k intuitivnímu přemýšlení zvyšují víru ve falešné zprávy a snižují věrohodnost pravdivých zpráv. Naopak zvýšená schopnost kritického myšlení zvyšuje rozlišení pravdivých zpráv. Konfirmační zkreslení pak snižuje rozlišení pravdivých zpráv a zvyšuje záměry sdílení zpráv. Politicky konkordantní pravdivé zprávy jsou více důvěryhodné a sdílené, ale efekt se různí u individuálních (negativních) a okresových (pozitivních) falešných zpráv. Tato práce poskytuje teoretický základ pro škálovatelné intervence vycházející z teorie duálního zpracovávání informací, které mohou být relevantní pro sociální média hledající způsoby, jak bojovat se šířením falešných zpráv.

Declaration of Authorship

1. The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

2. The author hereby declares that all the sources and literature used have been properly cited.

3. The author hereby declares that the thesis has not been used to obtain a different or the same degree.

Prague 27.07.2021

Acknowledgments

The last two years have been hard for everyone, and more so for some people. This thesis would not have been completed, had it not been for great mentors, family, and friends.

First and foremost, I would like to express my utmost gratitude to my supervisor, doc. PhDr. Julie Chytilová, Ph.D., for her strategic guidance, support, and unparalleled patience. Next, I would like to thank professor Wayne Landsman and Jiří Novák M.Sc., Ph.D., for giving me the opportunity of a lifetime. Next, professors Michael Christian, William Maddux, and Jan-Benedict E.M. Steenkamp from the University of North Carolina at Chapel Hill have been integral to coaching me as an aspiring researcher and helping develop the ideas incorporated in this manuscript.

Lastly, I would like to thank all my friends who assisted me with perfecting the thesis: Nicolas Say, for challenging my thoughts from the standpoint of cognitive psychology; Léon Kamenický, Yulia Rebrova, Ayla Sattarzada, and Meng Ting Shi for their heartwarming support and constructive feedback. Furthermore, I am grateful to Eszter Takacs for her help with setting up the Qualtrics questionnaire.

Ultimately, nothing would be possible to achieve without the unconditional love and support of my sister, parents, and my family in Armenia and beyond.

Master's Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University

Author: Bc. Mikayel Harutyunyan
Supervisor: doc. PhDr. Julie Chytilová Ph.D.
Specialization: FFM&B

Proposed Topic:

The impact of personality on human behavior in digital age.

Background and Motivation:

With the onset of the digital era came the dusk of the Gutenbergian era of printed media, drastically decreasing the cost of the mass distribution of information. As the Web spread worldwide, it became easier than ever to publish incorrect information, ready to be shared and taken to be true. In today's interconnected world, the role of veracity became crucial than ever, since it costs nothing to publish a lie, but the lies do have real costs. Nevertheless, what factors contribute to the information being perceived as true?

The answer to this question is important for several reasons. First of all, seminal theories of decision-making (e.g. Savage 1951; Becker, 1968; Mazar et al, 2008), cooperation (Fehr & Fischbacher, 2003), communication (Shannon, 1948) all rely on the concept of truth, accuracy or honesty in their goal of explaining human behavior (Vosoughi et al., 2017). In addition, humans are cognitive misers living in fast-paced environments overloaded with information. We thus use "heuristics", or mental shortcuts, to make choices - and some of them are based upon whether or not to believe a fact we are told. We are inclined to assign more value to the information that confirms previous beliefs, recall more easily, or believe to be popular opinion. Furthermore, while heuristics simplify our decision-making, they may lead to mistakes. The cost of these mistakes, however, is higher than ever.

The second half of the decade gave rise to fake news, or "fabricated information that mimics news media content in form but not in organizational process or intent" (Lazer et al., 2018). Empirically, it has been demonstrated that fake news can reinforce climate change denial (Van der Linden et al., 2017), cause misallocation of resources during terror attacks and natural disasters (Vosoughi et al., 2017), adverse and profound reactions of the stock markets (Rapoza, 2017), as well as misinformed elections (Mustafaraj & Metaxas, 2017; Guess et al., 2019).

A good illustration of the latter were the 2016 U.S. presidential elections, when false information amounted to around 6% of all news consumption (Grinberg et al., 2019), the average U.S. citizen saw at least one false news story and the vast majority of exposed individuals believed it to be true (Silverman & Singer-Vine, 2016). The trust in media dropped to a historic low, with an estimated 51% of Democrats and 14% of Republicans expressing at least a "fair amount" of trust in mass media as a news source (Gallup, 2016).

In the light of this all, there is a unified, interdisciplinary call to find avenues to mitigate the impact of fake news and make truth louder (Lazer et al., 2017). On the user side of the issue, fact-checking, debiasing, and media literacy activities have been suggested as potential remedies (Lazer et al, 2018). This line of research is mainly based on the seminal inoculation theory (McGuire, 1964), the notion of developing individual resistance by exposing the person to a weakened version of a (counter)-argument, and refuting it. One of the most cost effective solutions in line with the theory is showing a warning label (also called "disputed" flag) next to the article, cautioning users to content with questionable origin. It was, however, demonstrated that this may be undermined by politically motivated reasoning (Flynn, Nyhan, & Reifler, 2017), contribute to a more favourable view of fake news that have not yet been flagged (Pennycook & Rand, 2017) or backfire altogether and increase belief (Berinsky, 2015). Facebook abruptly cancelled their own "disputed" flags, since they were perceived as a way to silence free speech by users on both extremes of the political spectrum, and attracted more attention to the fake news publication (Lazer et al., 2017).

Research Question and Expected Contribution:

In general, the overarching topic of the thesis is "What is the mechanism behind the influence of fake news and what can be done to mitigate their impact?".

From the standpoint of behavioural economics, the mechanism of influence of fake news can be evaluated from the standpoint of confirmation bias or rational inattention. Indeed, individuals online are likely to share and believe in the information that confirms their beliefs, and face limited time to evaluate the information. This, however, does not explain why individuals may show increased support in case of warning labels. This manuscript aims to scrutinize further the following phenomenon applying the Elaboration Likelihood Model – a theoretical perspective not yet widely applied in the context of fake news.

Applying this theoretical framework, which is widely used in psychology and marketing literature, predominantly in the field of influence, would help evaluate the appropriateness of certain strategies of combating fake news used today.

For instance, one cause of the malfunction of the warning label interventions could be that they do not motivate individuals to process the information deeply. Based on the dual-route theory of Elaboration Likelihood Model, I argue that (partisan) individuals do not engage in higher-order critical evaluation of the falsehood that is congenial to their political views. In addition, I predict that warning labels can have an unintended negative effect, triggering biased elaboration. In its turn, that results in increased support and further dissemination of the article.

Hypotheses:

H1. Individuals are more likely to engage in high level of elaboration for the articles that have the same partisan affiliation, and less likely to engage in high level of elaboration for non-partisan articles.

H2.a Partisan individuals are more likely to support partisan article in case it is labeled as disputed.

H2.b Partisan individuals are more likely to be willing to disseminate the partisan article in case it is labeled as disputed.

Methodology:

Experimental Design Participants were presented with six ”real” news stories, as well as six actual “fake news” articles. Fake news stories were adapted from Silverman et al (2016) and concerned the most widely circulated stories during the 2016 US presidential election cycle. The stories were balanced across the political spectrum (3 left-leaning, 3 right-leaning). Headlines were presented in a random order. For each headline, participants answered three questions, adapted from Pennycook (2019): “Have you seen or heard about this story before?”, “To the best of your knowledge, how accurate is the claim in the above headline?” (rated on a Likert scale from “1” to “5”), “Would you consider sharing this story online (for example, through Facebook or Twitter)?” (rated: I would never share something political online (data removed), no, maybe, yes). Two types of disputed labels were introduced. The first type, “Disputed by 3rd party checkers” corresponds to the label Facebook used. The second one was formulated as follows: “Rated false by X users in your area. Information shared by a source with below-average trustworthiness ranking by Politifact. Research shows that users who fail to verify the story’s correctness have an increased risk of being misled by fake news” (partially adapted from Ross et al., 2017). Participants were randomized into three conditions:

(a) The warning condition, where all of the fake-news headlines (but none of the real-news headlines) contained the short disputed label (b) The control condition, where none of the headlines (but none of the real-news headlines) contained the disputed label.

The articles snapshots were shown in a simulated environment reflecting Facebook’s UI. The participants were asked to scroll through the newsfeed that contained other multimedia posts (such as videos or photos) and evaluate those that were highlighted with prompts. As a proxies of elaboration, the dependent variable, time spent at each article and whether or not the subject expanded it will be recorded.

The participants then proceeded to the second stage, where they were asked to complete a set of filler demographic questions (age, sex, income, education, political affiliation), as well as the Cognitive Reflection Test. Following Pennybook and Rand (2019), we used a seven-item CRT: the original three-item CRT by Frederick (2005) and less math-focused version from Thomson and Oppenheimer (2016). Next, in the final stage of the experiment, participants will be shown all the news article titles and will be asked to categorize them (1 = fake, 0= real news), as well as rate the familiarity with the headline (controls for recall and its impact).

Data Analysis

The data will then be analyzed via econometric methods as follows.

H1 will be tested in two ways. Firstly, and ordinary least squares regression will be used to test the hypothesis with time spent on reading the article as a dependent variable. Secondly, a logistic regression will be run, with the interaction with the post (1 = expanding the article/comment section, 0 = inaction) as a dependent variable. The various degrees of treatment will be coded as follows: 0 = no warning label, 1 = basic warning, 2 = expanded warning. H2 a,b will be tested via multivariate regression with self-reported level of likelihood to share the article based upon the headline and self-reported level of support of the article as dependent variables. Finally, mediation analyses proposed by Baron and Kenny (1986) will be conducted to validate the applicability of Elaboration Likelihood Model to the context of online media and fake news.

Key Literature:

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, *31*(2), 211-36. An, J., Quercia, D., & Crowcroft, J. (2014, October). Partisan sharing: Facebook evidence and societal consequences. In *Proceedings of the second ACM conference on Online social networks* (pp. 13-24). ACM.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. Journal of Human Resources, 43(4), 972- 1059

Bullock, J. G., Gerber, A. S., Hill, S. J., & Huber, G. A. (2013). *Partisan bias in factual beliefs about politics* (No. w19080). National Bureau of Economic Research.

Clayton, K., Blair, S., Busam, J. A., Forstner, S., Glance, J., Green, G., ... & Sandhu, M. (2019). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior*, 1-23.

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports*, *6*, 37825.

Ebnali, M., & Kian, C. (2019, July). Nudge Users to Healthier Decisions: A Design Approach to Encounter Misinformation in Health Forums. In *International Conference on Applied Human Factors and Ergonomics* (pp. 3-12). Springer, Cham.

Ecker, U. K., Lewandowsky, S., & Tang, D. T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & cognition*, *38*(8), 1087-1100.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, *80*(S1), 298-320.

Freelon, D. (2017). Personalized information environments and their potential consequences for disinformation. *Understanding and addressing the disinformation ecosystem. Annenberg School for Communication*, 38-44. Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social clicks: What and who gets read on Twitter?. *ACM SIGMETRICS Performance Evaluation Review*, *44*(1), 179-192.

Glaeser, E. L., & Ujhelyi, G. (2010). Regulating misinformation. *Journal of Public Economics*, *94*(3-4), 247-257. Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011, October). Predicting personality from twitter. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing* (pp. 149-156). IEEE. Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374-378.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances*, *5*(1), eaau4586.

Harrison, B., Svetieva, E., & Vishwanath, A. (2016). Individual processing of phishing emails: How attention and elaboration protect against phishing. *Online Information Review*, *40*(2), 265-281.

Hirsh, J. B., Kang, S. K., & Bodenhausen, G. V. (2012). Personalized persuasion: Tailoring persuasive appeals to recipients' personality traits. *Psychological science*, *23*(6), 578-581.

Horne, B. D., & Adali, S. (2017, May). This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In *Eleventh International AAAI Conference on Web and Social Media*.

Jang, S. M., & Kim, J. K. (2018). Third person effects of fake news: Fake news regulation and media literacy interventions. *Computers in Human Behavior*, *80*, 295-302.

Johnson, J. A. (2014). Measuring thirty facets of the Five Factor Model with a 120- item public domain inventory: Development of the IPIP NEO-120. Journal of Research in Personality, 51, 78-89.

Kriel, C. (2017). Fake news, fake wars, fake worlds. *Defence Strategic Communications*, *3*(3), 171-190.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Schudson, M. (2018). The science of fake news. *Science*, *359*(6380), 1094-1096.

Li, H., & Sakamoto, Y. (2014). Social impacts in social media: An examination of perceived truthfulness and sharing of information. *Computers in Human Behavior*, *41*, 278-287. Levi, L. (2017). Real fake news and fake fake news. *First Amend. L. Rev.*, *16*, 232.

Lyons, T. (2017). News feed FYI: Replacing disputed flags with related articles. *Facebook newsroom*, *20*. Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*(3), 106-131.

Markovikj, D., Gievska, S., Kosinski, M., & Stillwell, D. J. (2013, June). Mining facebook data for predictive personality modeling. In *Seventh International AAAI Conference on Weblogs and Social Media.*

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research*, *45*(6), 633-644.

Mele, N., Lazer, D., Baum, M., Grinberg, N., Friedland, L., Joseph, K., ... & Mattsson, C. (2017). Combating fake news: An agenda for research and action. *Di https://www. hks. harvard. edu/publications/combating-fake-news-agenda-research-and-action (Retrieved October 17, 2018).*

Mustafaraj, E., & Metaxas, P. T. (2017, June). The fake news spreading plague: was it preventable?. In *Proceedings of the 2017 ACM on web science conference* (pp. 235-239). ACM.

Na, E. Y. (1999). Is biased processing of strong attitudes peripheral? An extension of the dual process models of attitude change. *Psychological Reports*, *85*(2), 589-605.

O'Keefe, D. J. (2008). Elaboration likelihood model. *The international encyclopedia of communication*.

Osatuyi, B., & Hughes, J. (2018, January). A tale of two internet news platforms-real vs. fake: an elaboration likelihood model perspective. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion* (pp. 1-24). Springer, New York, NY. Pennycook, G., & Rand, D. G. (2017). The implied truth effect: Attaching warnings to a subset of fake news stories increases perceived accuracy of stories without warnings.

Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, *116*(7), 2521-2526.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of experimental psychology: general*.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39-50.

Pennycook, G., & Rand, D. G. (2019). Cognitive reflection and the 2016 US Presidential election. *Personality and Social Psychology Bulletin*, *45*(2), 224-239.

Rapoza, K. (2017). Can 'fake news' impact the stock market?.

Glossary

Bots	Automated accounts on social media impersonating humans
Fake news	Fabricated information that mimics news media content in form but not in organizational process or intent
Misinformation	False or misleading information
Newsfeed	A sequence of social media content that is algorithmically optimized to show posts users or their connections may like or engage with
Disinformation	False information that is purposely spread to deceive people

List of Acronyms

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
COVID-19	Coronavirus Disease 2019
CRT	Cognitive Reflection Test
ELM	Elaboration Likelihood Model
OLS	Ordinary Least Squares
MMR	Measles, Mumps, and Rubella
MS2R	Motivated System 2 Reasoning (Kahan, 2012), a theoretical concept positing that deliberation leads to increased belief in identity-concordant information
PET	Preference for Effortful Thinking
PIT	Preference for Intuitive Thinking
UN	United Nations
WHO	World Health Organisation

Table of Contents

List of Figures

List of Tables

Section 1 - Introduction

False news is nothing new. On the evening of October 30, 1938, radio listeners across the US could hear an astonishing account of a meteor strike in the small town of Grover's Mill, New Jersey. Sirens blaring in the background, the anchormen reported sights of otherworldly creatures, accompanied by futuristic war machines moving toward the Golden Apple under the veil of poison gas. That, of course, was Orson Welles' radio adaptation of "The War of the Worlds". Even though the broadcast contained a few brief warnings that it was fiction, its masterful execution fooled many people. Newspapers reported people rushing to police stations or having heart attacks due to the shock (Heller, 2018). The panic that ensued was the biggest news story for weeks, propelling Welles into stardom. Nonetheless, academics argue that the ultimate hoax was not the alien invasion (Schwartz, 2015). The real misinformation culprits were the newspapers, perpetuating the myth of widespread panic for weeks to discredit their novel competitor - radio. Overstating isolated clusters of impressionable people, they crafted a mass hysteria narrative that lingers on.

Radio's distant successors - social media, have significantly lowered the cost of the mass distribution of information. Today, it costs nothing to publish a lie, but the lies themselves have real costs. Fake news can reinforce climate change denial (Van der Linden et al., 2017), cause misallocation of resources during terror attacks and natural disasters (Vosoughi et al., 2018), misinformed elections (Guess et al., 2019; Mustafaraj & Metaxas, 2017) as well as strong adverse reactions in stock markets amounting to $39 billion annually (Cavazos, 2019; Rapoza, 2017). As such, online fake news is estimated to cost us $78 billion a year (Cavazos, 2019). In spring 2020, World Health Organization (WHO) and United Nations (UN) officials stated that the “infodemic” of misinformation was as dangerous as the novel coronavirus itself (UN, 2020). The misinformation on COVID-19 went viral online, amassing 3.8 billion views on Facebook alone in 2020. Alarmingly, the top ten fake news outlets had almost four times as many estimated views as the top ten leading health institutions, such as WHO (Avaaz, 2020). The falsehoods reduced mask-wearing and social distancing (Ioannidis, 2020a; Ioannidis, 2020b; Kouzy et al., 2020; Mattiuzzi & Lippi, 2020) and increased vaccine hesitancy (e.g., Carrieri et al., 2019), notably contributing to an almost 21% drop in immunization intent in September 2020 (Pew Research Center, 2020).

In the light of this all, there is a unified, interdisciplinary call to find avenues to mitigate the impact of fake news and amplify the truth (Lazer et al., 2018). What specifically influences human behavior that provides the rationale for the inability to discern between accurate and fake content? Answering that question could help design effective digital interventions to fight fake news. From behavioral economics and social psychology standpoint, dual-processing theories may shed more light on the subject (Pennycook & Rand, 2021b). Indeed, although people are more likely to trust politically concordant news (Pennycook et al., 2021b), it is not the only explanation of the phenomenon. In contrast to the politically concordant conclusion causally affecting the reasoning (or politically motivated reasoning; Tappin et al., 2020), the latest findings seem to indicate that people fall for fake news because they simply fail to deliberate; not because they deliberate in a motivated or identity-protective manner (Pennycook & Rand, 2019b).

With that in mind, it is crucial to evaluate the landscape of the interventions against fake news. The tools researchers and companies suggest vary: from those aimed at preventing individuals from encountering fake news in the first place to strategies targeted at priming individuals for a more critical assessment of the fake news they are exposed to (Lazer et al., 2018). The former strategy includes algorithmic discernment of fake news and labeling potentially false information. Although promising, it balances a thin line between the fight against false news and censorship and is proven to be largely inefficient at combating qualitatively novel topics - COVID-19, for instance (Pennycook et al., 2020). Fact-checking, in turn, is hard to scale. To this date, Facebook was forced to remove over 18 million pieces of coronavirus-related misinformation and flagged over 167 million pieces as "false" without removing them (Yaffe-Bellany, 2021). Alarmingly, however, fake news spreads faster and farther than true news or the fact-checked versions of the false articles (Vosoughi et al., 2018). As social media has become the top source of information for internet users, scalable intervention strategies that encourage analytic thinking are crucial. One promising approach is shifting the attention to accuracy - so-called "accuracy prompts" (Pennycook & Rand, 2021). Early meta-analytic evidence suggests that the method successfully encourages deliberate thinking (Pennycook & Rand, in press).

Therefore, combining the scalability of automated warning labels with the encouraging results of accuracy prompts in the context of social media seems a natural next step and is supported by specialists in the field (e.g., Pennycook & Rand, 2021a; Pennycook & Rand, 2021b). To the best of the author's knowledge, there are currently no studies scrutinizing the effect of accuracy prompts in the form of warning labels on fake news discernment. The current study attempts to investigate this further. An additional contribution of this thesis is the review of the available intervention strategies and recent findings through the lenses of the dual-processing theory. Specifically, the framework of the Elaboration Likelihood Model of Persuasion (ELM; Petty & Cacioppo, 1986), something only a few researchers have focused on as of today. Lastly, the study furthers the analysis of what factors contribute to susceptibility to fake news.

In an online pre-registered study (n = 520) of US residents recruited on Prolific, we find that accuracy prompts in the form of warning labels work, but only for increasing sharing intentions for true news. Although people are generally good at discerning fake news, they seem to fail to consider the veracity of the news when they want to share it. Problems with truth discernment are exacerbated by confirmation bias. People favoring higher-order, analytical thinking seem to be more confident in true news and more skeptical about fake news. This is the case for cognitively more reflective people and people with better political knowledge and higher education. In contrast, people with a preference towards intuitive thinking are worse on average at discerning truth. Similarly, Republican-leaning candidates mistrust true news and are more likely to believe in fake news. Finally, the concordance of the political bias of the news with the person's political ideology or the political preference of the county they live in increases the overall confidence in true news. When it comes to fake news, concordance increases accuracy judgment on the county level, but the effect is opposite at the individual level. Overall, the results overwhelmingly support the dual-process framework from the standpoint of the Elaboration Likelihood Model.

The rest of the thesis unfolds as follows: Section 2 provides an in-depth overview of multidisciplinary research on fake news and contextualizes them through the prism of dual processing theory. Next, Section 3 presents the experimental design and describes the methods. Further, Section 4 outlines the hypotheses. Section 5 summarizes the main findings from the experiment, while Section 6 links them back to the theory and discusses the implications. Finally, Section 7 concludes.

Section 2 - Literature Review

In this section, we first review the behavioral science of fake news - "fabricated information that mimics news media content in form but not in organizational process or intent" (Lazer et al., 2018). After discussing the evolution of fake news in the first subsection, we turn to motivational and cognitive aspects in the following two subsections. We demonstrate the overwhelming support of empirical results for the dual-process account of fake news influence. Furthermore, we review factors affecting human decision-making in the context of fake news through the lenses of the Elaboration Likelihood Model, a relevant dual-process theory. Finally, we discuss strategies proposed by researchers to reduce the impact of fake news and make a case for the importance of reviewing one in particular from the standpoint of dual-process theory - accuracy prompts in the form of a warning label.

2.1 The Evolution of Fake News

Organized deception, limited only by the virality of the mass media, can be traced back to the early beginnings of human society. After a failed siege of Kadesh in 1274 BC, Ramses II ordered news of his fake victory to be publicized through murals. Octavian waged a smear campaign against Anthony via short slogans imprinted on coins (Kaminska, 2017). Gutenberg's invention of the printing press in 1493 made spreading hoaxes, such as The Sun's 1835 series covering the discovery of (nonexistent) life on the moon (Andrews, 2015). Together with print, the radio-enabled disinformation campaigns against adversaries during both world wars mobilized the population in revolutions and justified atrocities such as Holocaust (Posetti & Matthews, 2018).

However, the onset of the internet and social media commenced the "post-truth" era, as coined by Oxford Dictionary (2016) after the US presidential election. Social media allows for sophisticated, effective, and coordinated dissemination of lies targeting specific population clusters or even individuals. This precision is driven by the fact that demographics and personality traits can be extracted from social media data (Volkova et al., 2015; Kosinski et al., 2013). Cambridge Analytica allegedly exploited such data to influence the results of more than 200 international elections from the USA to Argentina and the Czech Republic (Posetti & Matthews, 2018). State-backed troll farms (organized groups using fake social media accounts to amplify certain messages online) have been used to harass political opposition in Azerbaijan (Wong & Harding, 2021). Bots, automated accounts impersonating humans (Lazer et al., 2018), boosted political incumbent's support in Honduras and Brazil and increased support for the Russian-produced "Sputnik V" COVID-19 vaccine in Europe while undermining confidence in other vaccines (Gordon & Volz, 2021). Notably, some of the content is machine-generated - with recent advancements in machine learning able to bypass algorithmic fake news detection (Schuster et al., 2019; Schuster et al., 2020).

In the post-truth era, it indeed costs almost nothing to produce and disseminate disinformation (content with a deliberate intention to mislead people; Pennycook & Rand, 2021). Even high-ranking officials are not immune to the effect of fake news. Pakistan's Defence Minister, for instance, issued an official nuclear retaliation warning in response to a ‘fake news’ story of Israel planning a nuclear attack against the country he may have seen online (Goldman, 2016). Alarmingly, once accepted, falsehoods are challenging to correct and may influence related beliefs even when people no longer agree with them (see Lewandowsky et al., 2017). Hence, a growing multidisciplinary scientific effort was amassed to tackle fake news: researchers across behavioral economics (e.g., Jost et al., 2020; Tappin et al., 2020; Allcott & Gentzkow, 2017), psychology (e.g., Pennycook & Rand, 2021; Pennycook & Rand, 2019; Mosleh et al., 2021), political (Vegetti & Mancosu, 2020; Berinsky, 2017), communication (Amazeen et al., 2018) and information sciences (Zhao et al., 2020), as well as human-computer interaction studies (e.g., Kirchner & Reuter, 2020) started scrutinizing the fake news phenomenon.

In all, deception is as old as humankind. Nevertheless, the latest developments in technology have caused major shifts in the way falsehoods can propagate online and beyond. In the following subsections, we will attempt to synthesize multidisciplinary evidence to inform our main hypotheses.

2.2 Motivational Aspects of Fake News Influence

In this subsection, we briefly review the first of the two competing streams of research on fake news - one grounded in motivational processes and one attributing accepting, sharing, and correcting misinformation to cognitive processes. Since the main focus of this thesis is the latter, we will cover it in greater detail in the next section.

In 1998, a former British doctor, Andrew Wakefield, published what can be called one of the most famous "scientific fake news articles" in Lancet. The author doctored the data, linking autism to the MMR (Measles, Mumps, and Rubella) vaccine. Now retracted, the study's impact still lingers in the form of vaccine hesitancy, causing previously eradicated measles to rise again (Fact About Measles Outbreak, 2015). Why has this conspiracy permeated? In this subsection, we explore the motivational aspects of this phenomenon.

Greifeneder et al. (2020) argue that such false information gains plausibility and unfalsifiability due to social integration motivation and knowledge motivation. On the one hand, central to maintaining belief in falsehoods is a need for maintaining a need for belonging. On the other hand, social exclusion leads to superstitious thinking (Graeupner & Coman, 2017). A prime example of that is the Flat Earth Society - a group of people all around the globe supporting the idea that our planet is flat (Mirsky, 2020).

Similarly, knowledge motivation is satisfied by explainable patterns (Moulding et al., 2016). The MMR theory draws a false pattern between children diagnosed with autism after receiving the MMR shot. Next is uncertainty (an at least mildly unpleasant feeling of a doubt; Wichman et al., 2010). The feeling arises due to the discrepancy between the actual and the desired understanding of a phenomenon. Subsequently, people strive to reduce ambiguity (Greifeneder et al., 2020). Admittedly, fake news and conspiracies satisfy that need by oversimplifying the events and explaining the world in an organized and predictable way. For instance, "vaccines are harmful" is a more accessible worldview than "vaccines may have side effects on a negligibly small percentage of the population". The latter is also confirmed by the fact that individuals scoring high on the need for structure tend to believe that the media are trying to actively deceive their audience (Axt et al., 2020).

Another motivational process impacting the belief in fake news is identity-based. Identity-based motivation theory posits that agents make sense and act congruently with their salient social or personal identities, which are accessible in the moment of decision-making (Oyserman & Dawson, 2020). To illustrate, the "Leave" campaign in the UK ran a false ad claiming that "the European Union wants to kill cuppa" (i.e., the traditional "cup of tea"). The visual ad framed European Union secession as a precondition for the maintenance of British identity. As a result, the "Vote to Leave" call-to-action was linked to the need to adhere to a collectivistic, identity-based action. As a result, British voters no longer should have remembered the source of the information, credibility, or veracity. Instead, they may have thought: what is a "truly British" thing to do? Unable to reach a decision, a critical demographic such as 18-24 stayed at home (Douthat, 2015).

In sum, fake news acceptance and sharing have a motivational component, grounded in identity (especially political identity), the need to reduce uncertainty, find explainable patterns and belong.

2.3 Cognitive Aspects of Fake News Influence

This section reviews the second of the two competing streams of research on fake news, grounded in information processing, cognition, behavioral economics, and influence literature. We present the dual processing model of elaboration likelihood (ELM) and contextualize empirical findings within this framework.

2.3.1. Dual Processing Account of Fake News

We are hardwired to believe that new information is accurate (Marsh & Stanley, 2020). In a generally truthful everyday life, where any new piece of information is more likely to be true than false, this may be a more cognitively efficient approach (Gilbert, 1990). In the online world, where around 15% (47 million) of Twitter accounts are bots (Varol et al., 2017), with the majority spreading political disinformation (Ferrara, 2020), this cognitive shortcut gets more costly. A recent study by YouGov, for instance, indicated that only 4% of participants were able to fully discern fake news from true news (Channel 4, 2017). Dual-process theories, a cornerstone of research on reasoning, put this "lazy processing" at fault (De Neys, 2017; Evans & Stanovich, 2013; Kahneman et al., 2011; Petty & Cacioppo, 1986). In this subsection, we briefly summarize the principles posited by dual system theories and motivate the use of one of them - the elaboration likelihood model of persuasion (ELM; Petty & Cacioppo, 1986) to advance the argument.

Although dual-process theories differ slightly, they all agree on a central concept: human cognition can be separated into two fundamentally distinct types of processes that qualitatively differ from each other (Evans & Stanovich, 2013). System 1 (or peripheral route in ELM) processing can be described by automaticity in response to stimuli and reliance on heuristics. In contrast, System 2 (or central route in ELM) processing is more deliberate and effortful. Consider the following item from the Cognitive Reflection Test, for instance:

"In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?"

The incorrect, intuitive answer - 24 days - would come to most people immediately. However, the correct answer - 47 days - would be apparent only with additional thinking (Toplak et al., 2011). In the context of fake news, therefore, the dual processing perspective expects people who deliberate more to be more likely to discern truth, regardless of political concordance (as opposed to the motivated reasoning account). This prediction has solid empirical support. People who engage in higher-order processing do discern true news from fake news more (for review, see Pennycook & Rand, 2021), regardless of concordancy with partisanship (Ross et al., 2021; Bago et al., 2020; Pennycook & Rand, 2019), the fact they are judging just the headlines or the articles (Pehlivanoglu et al., 2020) or the exact measure of cognitive sophistication (Bronstein et al., 2019). Increased deliberation and attention to accuracy - components of the central processing route (System 2) seem to reduce sharing intentions for false headlines much higher than the rating of their veracity (e.g., 91% higher in Pennycook et al., 2021b). This finding is essential, considering that an analysis of 2.8 million news posts on Twitter showed that 59% of the news items were shared with no effortful deliberation (being opened; Gabielkov et al., 2016).

In all, deliberation and more analytic thinking may override the faulty predictions of intuitive cognition and increase truth discernment.

2.3.2 Confirmation Bias and Motivated Reasoning - Where Cognition and Motivation meet

Before we transition to outlining the research findings outlining the influence of fake news from the standpoint of the dual-process framework, it feels appropriate to review two important concepts between the motivational and cognitive aspects of fake news influence. These are confirmation bias and the closely related notion of (politically) motivated reasoning.

Firstly, confirmation bias is a cognitive bias, or rather a set of biases (Bryant, 2020) that generally work to ignore evidence that contradicts their preconceived notions (Kahneman et al., 2011) and have a "common propensity of over-belief in one’s preferred opinions" (Klayman, 1995). Importantly, confirmation bias has both motivated and unmotivated aspects (Sanitioso et al., 1990) that are hard to decouple non-experimentally (Bryant, 2020). On social media, users are restricted to a self-perpetuating cycle of being shown information they like, further increasing the confirmation bias towards what they believe is the only truth (Thornhill et al., 2019). When consuming news on social media, people do so individually and usually gravitate towards ideologically homogeneous groups called "echo chambers". This further increases their confirmation bias (Duffy & Ling, 2020; Masta & Shearer, 2018; Törnberg, 2018; Athey et al., 2017; Garrett, 2017; Del Vicario et al., 2017). As a result, people only see online what perpetuates their beliefs: after Donald Trump's 2016 election win, many Democrats were left surprised - since they did not see any notable support for Trump in their newsfeeds (Thornhill et al., 2019).

Low-effort, peripheral cognition is profoundly influenced by confirmation bias (Bago & De Neys, 2017; Kahneman et al., 2011) as it draws on experience and favors information reinforcing pre-existing beliefs (Nickerson, 1998). In addition, confirmation bias disregards opposing facts (Nickerson, 1998), blocking out the information inconsistent with one's preconceived notions (Moravec et al., 2019; Devine et al., 1990). It thus comes as no surprise that this variable affects sharing and belief in fake news (e.g., Britt et al., 2019; Kim et al., 2020),

In contrast, (politically) motivated reasoning works through higher-order cognition (Pennycook & Rand, 2021; Kunda, 1990). The phenomenon lies at the intersection of motivational and cognitive research on fake news. It refers to the disparity between assessing the veracity of information conditional on the political ideology or partisanship. Humans have a "partisan brain" that places loyalty over veracity, thus decreasing truth discernment (Van Bavel & Pereira, 2018). Indeed, people are motivated consumers of misinformation (Kahan, 2017). As a means of ego defense (Albarracin, 2020), they engage in 'identity-protective cognition' upon encountering politically valenced articles. Subsequently, they overestimate the veracity of politically concordant content and underestimate content inconsistent with their partisanship (Kahan, 2012). This is known as the Motivated System 2 Reasoning (MS2R; Kahan, 2012), a theoretical account positing that deliberation causes agents to believe information concordant with their ideological identity.

In a recent review of the fake news research, Pennycook and Rand (2021) challenged the motivated reasoning framework, demonstrating that the effect of political concordance on truth discernment is much smaller than the influence of the veracity itself. Moreover, they list a few challenges with identifying partisans' motivation in thinking. Firstly, there is a possible confounding effect with other relevant variables and apparent experimental design drawbacks biasing the results. Accounting for these flaws shows that cognitive sophistication does not trigger politically biased processing (Tappin et al., 2020a; Tappin et al., 2020b). Instead, the differences across the aisle may arise from unbiased, Bayesian inference built on prior knowledge and beliefs. Exposure to polarized and different information environments, users build a different factual knowledge base regarding specific facts or empirical evidence (Tappin et al., 2020), such as climate change (Kahan et al., 2012).

2.3.3 A Brief Introduction to the Elaboration Likelihood Model

In the previous subsection, we have explored dual-processing theories and how they relate to fake news. This subsection is dedicated to discussing one of such theories, the Elaboration Likelihood Model of persuasion (Petty & Cacioppo, 1986) - one of the prominent frameworks of persuasion literature. Understanding which factors influence the activation of the central processing route - a more effortful, analytic way of cognition is critical to design interventions against fake news that work.

ELM is a dual-process model developed to explain how persuasion and attitude change can occur. According to ELM, individuals process information via two routes: the central processing route (analogous to System 2), involving “elaboration” (or thoughtful consideration of the information), while the peripheral processing route (like System 1) does not consider all the relevant elements, instead of basing the decision on simple (peripheral) cues. Understandably, the latter requires less cognitive effort on behalf of the recipient to assess information, and the attitude change is typically less lasting.

Several factors lead to the activation of the central route, namely, motivation (e.g., need for cognition, personal responsibility, personal relevance) and ability to process information (e.g., message comprehensibility, prior knowledge, cognitive resources, repetition). When these factors are not satisfied, people are more likely to resort to a state of low elaboration. Consequently, individuals tend to adjust their attitude based on peripheral cues or extrinsic factors to the arguments presented but are quickly processed, such as affective state, expert endorsements, number of arguments. An aspect not explored in this study but essential for designing successful intervention is the temporal aspect of ELM. Typically, the stimulus will show stronger temporal resistance and more significant opposition to counter-persuasion if the central route is activated. Notably, some factors (e.g., message comprehensibility in our case) can serve both as a peripheral cue and a factor increasing motivation or ability to process (Kitchen et al., 2013; Petty & Cacioppo, 1986).

To illustrate how the framework works, Petty and Cacioppo (Petty & Cacioppo, 1986) asked students to read one of three messages to increase university tuition at either a distant but comparable university or the students' own university (personal relevance manipulation). The three messages contained either three weak, three solid arguments or the six arguments altogether. When the students had to decide the fate of the distant university, the personal relevance, hence the motivation to process the information, was low. Consequently, a significant proportion of students were persuaded by the arguments' quantity rather than their quality. Conversely, when the students had to agree with the enforcement of the new policy at their home university, they favored messages with strong arguments over agreed weak arguments. In contrast, the number of arguments did not significantly affect their decision.

Although ELM has been challenged over the years due to its better suitability for post-hoc analyses rather than predictive ability, the less clear applicability to persuasion online, or the true effect size of some variables, ELM is one of the popular frameworks of the persuasion literature (Kitchen et al., 2013). More importantly, recent studies on fake news started successfully applying ELM in the context of fake news, demonstrating strong support for the model. These studies used ELM to develop predictive models for fake news (Zhao et al., 2021; Janze & Risius, 2017), understand factors contributing to their spread (Horne & Adali, 2017), and identify features differing between fake and true news (Singh et al., 2021; Lee et al., 2018; Horne & Adali, 2017).

In all, the dual-route ELM posits that various factors influencing the motivation, ability, and opportunity to process, trigger the processing of the information via the peripheral (heuristic and low-effort) or central (analytic and cognitively more challenging) routes. In the next section, we will explore how this model applies to fake news.

2.4 Additional Factors Influencing Fake News Susceptibility Through the Lenses of the Elaboration Likelihood Model

In days prior and shortly after the US presidential election in 2016, almost half the news shared online was fake (Howard et al., 2018). One of the most popular stories - an alleged child abuse ring led by Hillary Clinton operating out of a pizza restaurant (Posetti & Matthews, 2018). As such, not only do fake news propagate differently from true news (Vosoughi et al., 2018), but they also tend to be distinct from accurate articles qualitatively (see Horne & Adali, 2017). In this part of the literature review, we present the traits of fake news in line with the dual-process ELM and act by increasing motivation and ability to process information or act as peripheral cues and trigger the lower-effort route.

2.4.1 Factors Affecting Peripheral Route Activation

This subsection summarizes factors contributing to peripheral route activation, such as emotional valence, cognitive fluency, source credibility, distraction, and familiarity.

Fake news is written to elicit a stronger affective response. In turn, reliance on emotion promotes belief in fake but not true news (Martel et al., 2020). Fake news triggers emotions like fear and disgust, which are much stronger in arousal than sadness, joy and trust evoked in users by real news (Osatuyi & Hughes, 2018; Vosoughi et al., 2018; Horne & Adali, 2017; Plutchik & Conte, 1997). They are also typically more novel than true news, making them more likely to be shared (Horne & Adali, 2017). Since the affective cues are more negative, they may attract more attention (Pratto & John, 1991), be perceived as more informative (Peeters & Czapinski, 1990), and generally rated more plausible due to the negative framing of the statement in itself (see Jaffe & Greifender, 2020).

Fake news is significantly more straightforward to process than true news. Targeting mainly people who read just the titles of the articles (Horne & Adali, 2017), fake news creators provide less information in the text to facilitate quicker processing of the information (Osatuyi & Hughes, 2018). While real news persuades through arguments - facts and figures, quotes, additional references, fake news persuades through simplicity. To increase cognitive fluency, fake news articles are shorter and use simpler, less technical language, fewer quotes, and fewer arguments. A higher amount of substance and claims are put into titles than in the body of the text (Horne & Adali, 2017). As a result, fake news is easier to process and may thus be more believable (Wang et al., 2016).

Source credibility is another factor affecting peripheral processing, and although mixed, the evidence broadly supports the importance of source cues on truth discernment. There are multiple source cues present in the social media context. The source is another important cue that may be used when evaluating news. Participants are more likely to believe the information provided by people they view as credible (Pornpitakpan, 2004). For instance, a false claim attributed to Trump increases his supporters' belief in the claim while reducing Democrats' belief (Swire et al., 2017). Social feedback, such as the number of "likes", also seems to increase perceived credibility, particularly for misinformation (Avram et al., 2020). However, removing source information or making it prominent does not seem to impact accuracy beliefs (Dias et al., 2020).

Finally, familiarity heuristic increases the impact of fake news (Pennycook & Rand, 2018). Consistent with a well-known influence of prior exposure on truth discernment (Pennycook & Rand; 2018) a single previous exposure to a fake news headline increases later belief in the story regardless of its partisanship (Rand et al., 2018) or plausibility (Fazio et al., 2019).

2.4.2 Factors Affecting Central Route Activation

To conclude the review of factors affecting deliberate judgment of fake news, we discuss factors affecting ability and motivation to process information relevant to fake news. We omit the discussion of the already presented aspects of message comprehensibility and cognitive sophistication (e.g., CRT performance and need for cognition; Petty & Cacioppo, 1986) for the sake of brevity of this section. We thus discuss the effects of distraction, prior knowledge, and repetition, as well as personal relevance.

Firstly, the impact of the cues listed in the previous subsection may be amplified by contributing to general distraction. As distraction increases, the strength of an argument becomes a less important determinant of persuasion, while the effect of peripheral cues increases (Petty & Cacioppo, 1986). Social media is abundant with simple cues, such as the number of reactions to the posts and the valence of the top comments (determined by the amount and type of the reactions of other users). Even a brief exposure to a related but non-probative photo (e.g., a photo of a syringe above the article on COVID-19 vaccines) can bias people in perceiving the claim as accurate (Newman & Zhang, 2020).

On the contrary, prior knowledge increases the ability to process information effortfully. Media literacy (Amazeen & Bucy, 2019), political knowledge (Brashier et al., 2021; Vegetti & Mancosu, 2020), and science knowledge (Pennycook et al., 2020) increase truth discernment. Individuals scoring high on these measures should have higher ability and motivation to pursue the central processing route. Interlinked with prior knowledge is repetition, although it serves as a double-edged sword: seeing the same fake news article multiple times increases perceived truth equally for both fake and true news (Corneille et al., 2020; Fazio et al., 2019).

Likewise, personal relevance (e.g., previously covered political concordance or topic interest) may increase the motivation to process the information. Based on the information individuals select as interesting, the social media algorithms optimize engagement by showing people what they like. The former effectively confines users in the “echo chambers” we have alluded to in Section 2.3.3, where they can increasingly polarize their opinion (Freelon, 2017; Del Vicario et al., 2016). Upon encountering a fake news article on their “newsfeed” (algorithmically optimized stream of updates from the topics and friends the user follows, as well as promoted posts), users are likely to get exposed to reactions from like-minded people, which influences their opinion on truthfulness and propensity to share content (Li & Sakamoto, 2014).

2.5 Fighting Against Fake News

When Facebook introduced a "disputed flag" on articles in December 2016, it backfired. Instead of curbing misinformation, it caused readers to engage with dubious content more often (Meixler, 2017). It is thus imperative to understand the theoretical underpinnings of various interventions. In the final part of the literature review, we finally discuss the strategies for mitigating the impact of fake news. Specifically, we will focus on user-level behavioral interventions and accuracy-based prompts, one of the most quickly developing research streams.

2.5.1 Brief Overview of Strategies To Fight Fake News

Van der Linden and Roozenbeek (2020) broadly categorize interventions against fake news into four distinct categories: (1) algorithmic, (2) corrective, (3) legislative, and (4) behavioral.

The first approach is based on automatic detection and downranking fake news websites to prevent false information from spreading (Elgin & Wang, 2018; Calfas, 2017). This approach is problematic for several reasons. Firstly, the truth is not black-and-white: disagreement in how to rate an article can occur even among professional fact-checkers (Allen et al., 2020; Lin, 2018). Furthermore, the artificial intelligence algorithms are imperfect, especially for qualitatively novel misinformation (Pennycook & Rand, 2021b). In addition, they may be ineffective, resulting in backfires (Wakefield, 2017) or unnecessary censorship (Wolley, 2020).

The next approach is a post-hoc correction of false information. However, there are multiple issues with fact-checkers debunking the information. The approach's efficacy is limited and subject to time decay (Nyhan & Ryfler, 2019), mainly due to the familiarity effects discussed previously (Pennycook et al., 2018). Furthermore, the approach is not readily scalable: it is far easier to produce a fake news article than to fact-check it. Although fact-checking may be crowdsourced (Pennycook & Rand, 2019), the audience for the debunked articles is limited (Kurtzleben, 2016), as fake news spreads farther and faster than the truth (Vosoughi et al., 2018).

Another way to curb fake news spread is through legislation. In France, for instance, the law puts tighter restrictions on what content can be publicized during the election season (Bremner, 2018). This approach is not broadly popular due to the concerns of freedom of speech, the time lag between legislation and technology trends, or its prescriptive nature. For instance, although there have been calls for social media platforms to curb fake news, Facebook removed only 1 in 10 fake news articles its fact-checkers rated false or misleading last year (Yaffe-Bellany, 2021).

Ultimately, the most promising are the behavioral, non-invasive interventions. Most prominent examples in the literature include various warning labels (e.g., Brashier et al., 2021), making the publisher information more or less visible (Dias et al., 2020; Jakesch et al., 2019), proactive 'inoculation' (or so-called 'prebunking' against fake news; e.g., Van der Linden and Roozenbeek, 2020). Together, these approaches encourage people to "think slowly" and consider the accuracy of the article they are exposed to (Pennycook & Rand, 2021; Lorenz-Spreen, 2020).

To sum up, various approaches utilize algorithms, fact-checkers, legislation, and behavioral interventions to stop the spread of fake news online. These approaches vary in efficacy and scalability, with behavioral interventions seemingly providing scalable alternative resolution to fake news influence by engaging people in higher-order thinking.

2.5.2 Fighting Fake News with Accuracy

Although empirical results show that people are good at rating headlines true or false, this does not translate into discerning between true and fake news in terms of sharing (Epstein et al., 2021; Pennycook et al., 2021; Pennycook et al., 2020). Naturally, researchers wondered if making the concept of accuracy salient is an effective way to reduce the impact of fake news. This subsection reviews the findings across these experiments.

One significant benefit of the accuracy prompts over most strategies mentioned in the previous subsection is that they are proactive rather than reactive. Shifting people's attention to accuracy increases their motivation to deliberate, as people value sharing only accurate news over fake news innately (Pennycook et al., 2021). In addition, the treatment makes them pause and think, which is beneficial for truth discernment and higher-order thinking in itself (Fazio, 2020). Most accuracy prompt studies concentrate on asking people to rate the accuracy of an ostensibly unrelated neutral headline before the experimental stimuli are shown (for review, see Pennycook & Rand, in press). Other studies (e.g., Epstein et al., 2021) manipulate the importance of the concept, asking to rate participants how important it is to them to share only true news (or not share fake news). In addition, participants might be turned to norms. For instance, they may be told that most people think it is crucial to share only true news (Pennycook & Rand, in press). Other, less popular treatments include showing brief videos targeted at the importance of accuracy (Pennycook & Rand, 2021a) or prompting participants to think in a deliberate, not emotional, way (Pennycook & Rand, in press). The effects were robust to the intervention method and were amplified when combined. Nonetheless, a notable gap in research within this field remains in combining accuracy prompts within warning labels (Pennycook & Rand, 2021a). Out of all treatments, this intervention strategy is most easily deployed on social media and is proven to work if implemented with caution online (Pennycook & Rand, 2021b).

In conclusion, dual-process theory can be applied to investigate the behavioral aspects of fake news. With multiple cognitive and motivational factors involved in the fake news influence mechanism, it is appropriate to review them via a theoretical framework accounting for most of them - ELM. Understanding the underlying influence process makes it possible to design appropriate behavioral interventions - one of the most effective strategies to fight fake news online. A promising solution, accuracy prompts, is being increasingly scrutinized by researchers. However, to the best of the author's knowledge, the effect of accuracy prompts in the form of warning labels is yet to be seen.

Section 3 - Experimental Design

In this section, we explore the experimental design of the study. In Section 3.1, we review the sample selection details and the experimental procedure. Section 3.2 outlines the variables included in the research based on the literature review conducted in Section 2. Section 3.3 summarizes the quantitative methods used in the study, while Section 3.4 provides an overview of robustness and assumptions checks done in the study. Preregistration, data, and supplementary materials are available on the Open Science Framework at https://osf.io/wcq8b/. The Research Ethics Committee of the Faculty of Social Sciences, Charles University, has approved this study.

3.1 Participants and Procedure

3.1.1 Sample Selection and Exclusion Criteria

The pre-registered sample of participants for the pre-test and main study has been collected via the Prolific online survey platform (Palan & Schitter, 2018). All participants have been reimbursed for their time competitively according to Prolific policies. For our sample, we have chosen US residents that have completed at least 300 tasks with an approval rate of 95% and above, as suggested by Peer et al. (2021). Following Pennycook & Rand (2019a), we specify these additional pre-registered exclusion criteria. Participants were excluded from the study if they reported not having or never using a Facebook account. For the models studying the influence of fake news and other factors on sharing, we excluded participants reporting they never share political news online. Although we do not exclude participants based on inattention, we do control for it using two established attention checks following Berinsky et al. (2014), as well as one of our own (a multiple-choice task asking the respondent to identify which news item was shown to them during the study). Respondents were not excluded for inattentiveness to avoid selection effects that may undermine our analysis.

3.1.2 Experimental Procedure

Main study procedure

For the sake of brevity, we describe the main study procedure first, since the pre-test differed only by the number of items included and a few questions to verify the effectiveness of the treatment. Participants were recruited to provide opinions on social media posts. After screening, participants were randomly assigned to treatment or control conditions and presented with 18 actual fake and true news headlines in random order following Pennycook et al. (2020). The difference between the two conditions was that the treatment condition had an accuracy prompt displayed under each headline saying, "*Consider the accuracy of this article before sharing it. Learn why this is important*". The author created all article previews and the accuracy prompt to resemble the most current Facebook designs (Figure 1).

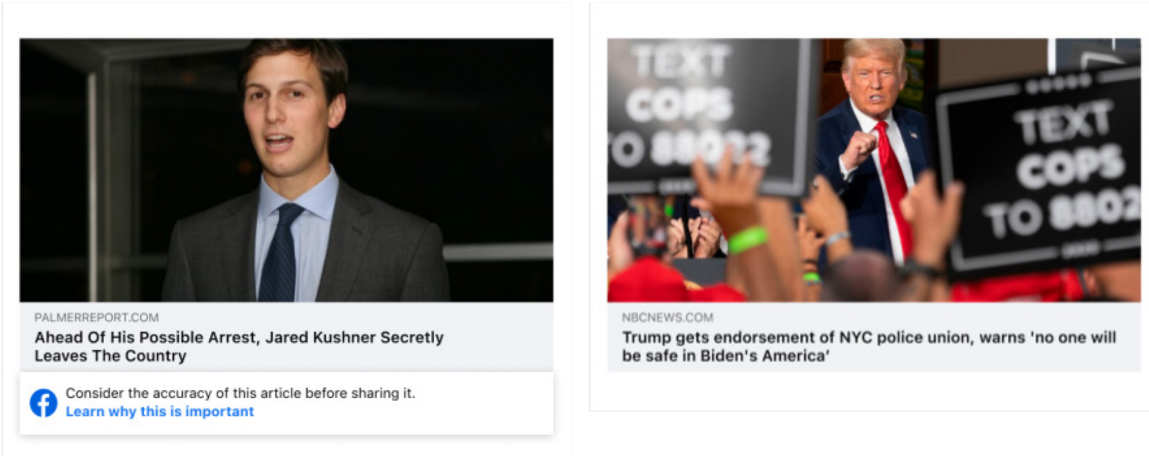


Figure 1: Treatment (left) versus control (right) design for news headlines

Subsequently, the respondents completed a five-item political knowledge questionnaire (from Pennycook & Rand, 2020) and a pooled seven-item Cognitive Reflection Test (CRT; Thomson & Oppenheimer, 2016; Frederick, 2005). In addition, participants also completed several demographic questions (age, gender, education, income, household size, political preference, and zip code). Based on the participant's support of the Republican versus Democratic Party, we classified the headlines as politically concordant versus discordant. Next, the participants also self-reported their preference for intuitive and effortful thinking (Newton et al., 2021) and how easy it was for them to identify fake news online. Finally, participants had to answer if they responded randomly, looked up answers online before responding to a question, or thought the news items we had selected were fabricated. We did not exclude those participants but wanted to ensure the control and treatment were balanced across these subgroups. The complete survey is available in the Appendix and online (https://osf.io/wcq8b/).

Pre-test procedure

Participants were recruited via the same platform and were given the same instructions. The only significant difference was the reduced number of CRT (three instead of seven) questions and experimental stimuli (4 news articles instead of 18). At the survey end, the control group participants were asked to provide opinions on the treatment efficacy (*In your opinion, how effective is the label used in this study in bringing attention to the accuracy of the news online?).* In contrast, participants assigned to the treatment were asked, "*How effective would you say the label used in this study in your opinion?".* Both groups then were asked: *"If you saw such a sign below a post, would that impact your assessment of the likelihood of the headline to be true?".*

Material selection

One of the essential parts of studies on fake news is the selection of correct experimental stimuli. We have thus closely followed the approach outlined by Pennycook et al. (2021a) and selected the articles directly from the list they have provided.

The fake news headlines were initially selected from a reputable third-party fact-checking website Snopes.com and were verifiably fake. The true news headlines were selected by Pennycook et al. (2020) from respected mainstream media and were roughly contemporary with the false headlines. Importantly, we have chosen headlines that are still relevant and balanced across the political spectrum, using the data on the bias score assigned to the article by Pennycook et al. (2020). Although there is no recommended approach for the number of pieces used in the experiment, studies conducted by Pennycook et al. (e.g., 2021, 2020a, 2020b, 2019, in press) typically recruit larger samples. They have a longer (e.g., 30-item, with random sampling from a larger pool of items) set of news headlines. Due to the cost and survey length constraints, we limited the number of items, ensuring we captured maximum information per item nonetheless. This meant equalizing the average bias of the headlines and selecting items of varying difficulty (i.e., the percentage of correct guesses in the sample provided by Pennycook et al., 2020). We analyzed the dataset supplied by Pennycook et al. (2020), scoring all the items and picking three items per category (veracity x left, neutral, or right-leaning). The resulting average shift from a mean of 3.5 on the scale from pro-Democrat to pro-Republican for groups of fake and true right and left-leaning articles was 0.8. The item difficulties were 49.37% for fake headlines and 59.4% for true headlines. The table in Appendix IV summarizes the headlines selected.

3.2 Variables of Interest

3.2.1 Dependent Variables

Our two main dependent variables of interest are *likely,* a 1 to 6 (1 = unlikely to be true) Likert-type variable measuring belief in the news (likelihood to rate news as true; *"Do you think that the headline is likely to be true?"*) and similarly-rated *share*, measuring sharing intention (*If you were to see the above article on Facebook, how likely would you be to share it?*). An alternative method applied in fake news research is *discernment. Truth discernment* is the extent to which respondents can distinguish true headlines from false headlines in their assessment (Pennycook & Rand, 2019a).

Effectively it is similar to 'sensitivity' or 'd' in signal detection theory (Wickens, 2002), as it explains the ‘overall’ accuracy of one's beliefs. Consider a person who gives a rating of 7 out of 9 (i.e., 0.78) true headlines and 3 out of 9 false headlines (i.e., 0.33) correctly. This means that their overall discernment is 0.78 - 0.33 = 0.45. The higher people score on the scale, the more sensitive they are to truth relative to falsity (Pennycook et al., 2020). Similarly, we compute sharing discernment. Hereafter, when interpreting results of the regressions with *likely* and *share* as dependent variables, we refer to the interaction effect of any independent variable and the news being fake (*fake_dummy,* equals one if fake) as “truth discernment” or “belief in fake news”.

3.2.1 Independent Variables

Confirmation bias

Given the mounting evidence that confirmation bias affects sharing and belief in fake news (e.g., Moravec et al., 2020; Britt et al., 2019), we measure confirmation bias following (Moravec et al., 2020). We construct the measure of confirmation bias through the multiplication of variables *importance* (*Do you find the issue described in the article important?*; 1 to 7; 7 = extremely important) and *position* (*What is your position on the topic covered by the article?*; -3 to +3, +3 = extremely positive). Thus this scale from −21 to +21 captures both direction (agree/disagree) and the magnitude (strongly/weakly) of the congruence between the respondent's preexisting beliefs and the news items (Moravec et al., 2020).

Cognitive Reflection Test

We will denote performance at the task with the variable *crt_score*. In total, *crt_score* can vary from 0 to 7. Since the measure is created by pooling the items from the classical (Frederick, 2005) and non-numeric (Thompson & Oppenheimer, 2016) versions, we verify the reliability of the scales. Cronbach's alpha, a measure for scale reliability, is equal to 0.74, which is satisfactory. Notably, we find that 301 respondents (57.88%) reported having completed these tasks before, and only 113 (21.73%) said they had never seen it. We find that prior exposure to CRT in our sample did affect the task performance but did not influence the study's main outcomes.

Political knowledge

We use a five-item political knowledge questionnaire (from Pennycook & Rand, 2020), comprising items like "W*hose responsibility is it to decide if a law is constitutional or not?*" with three multiple-choice answers shown in random order. Participants were encouraged to guess if they did not know the correct answer. After a 10 second countdown, the next question was displayed. Variable *political_knowledge* thus varies from 0 to 5.

Political identity

We have asked two questions to determine the political preferences of the respondents. The participants were asked to rate themselves on a seven-item scale ranging from strongly Democrat to strongly Republican (4 = Independent). If the participants identified as Independent, we used their answer to the binary "*If you absolutely had to choose, which political party would you vote for?*" to determine their preference to code the political concordance variables (explained in the next paragraph). Variable *prefers_republican,* ranging from 1 to 7, denotes the answers to the former question.

Individual's political concordance

As previously mentioned, if the headline bias favored the party the respondent was supporting, the *concordance_dummy* dummy was rated 1. If the headline was neutral, the concordance was rated as 0 either way. As a robustness check, we also compute *concordance2_dummy,* where neutrally-written articles were concordant (*concordance2_dummy* = 1) if the person identified as independent, with no significant differences between the methods of coding.

County-level political concordance

Minor addition to the fake news literature is also estimating the concordance of the political concordance of the article with the political preferences of the individual's zip code (county-level, in our case). There are three reasons to measure this. According to ELM, prior knowledge and personal relevance are essential factors in determining the ability and motivation of the individual to process the information thoughtfully. Living in a traditionally Democrat state, for instance, may boost the formation of factual knowledge about the party and Democrat-leaning topics (Tappin et al., 2018). In addition, county-level political concordance might make processing the news items easier, leading to less resource-consuming activation of more critical thinking (Petty & Cacioppo, 1986). Thus, we expect county-level political concordance leading to better truth discernment overall.

To estimate the concordance, we matched the self-reported zip code of the participant to the 2020 US presidential county-level voting data (MIT Lab, 2021). If most of the county voted for the party favored by the news article's political bias, the *concordance_zipcode_dummy* binary variable equals 1.

Living in a polarized county

Political polarization online translates into offline actions (Gallacher et al., 2021; Howard et al., 2018; Lane et al., 2017). Conversely, the state's political orientation affects the discourse online (Karami et al., 2021), with stark differences in what people in uncontested and swing states consume in terms of political news and misinformation (Howard et al., 2018). Although exploring what happens in swing states might be interesting, our sample size does not allow for that analysis. Instead, we want to control for polarization (vote margin more considerable than or equal to 5%). In states won by a landslide, the support for the candidate and the party is naturally higher. Higher support can be displayed in news coverage, outdoor campaigning, word of mouth, and increased sharing of partisan content (An et al., 2014) that we will control for in the study.

Preference for effortful (PET) or intuitive thinking (PIT)

To differentiate between the preference for effortful thinking (PET) and preference for intuitive thinking (PIT), we use two items with the highest correlation with their respective scale from the Comprehensive Thinking Styles Questionnaire (Newton et al., 2021). *"I often go by my instincts when deciding on a course of action,"* thus measuring the preference for intuitive thinking. At the same time, *"I try to avoid situations that require thinking in-depth about something"* is a reverse-scored item measuring the preference for effortful thinking. Both variables range from 1 to 7 (extremely agree). For better analysis, they are median-centered in the models and denoted as *pit_centered* and *pet_centered. Pit_centered* varies from -4 to +2, where +2 indicates 2 points higher than the median. Likewise, *pet_centered* is between -5 and +1, where -5 represents 5 points lower than the median.

Level of education

Meta-analytic evidence speaks in support of college education improving critical thinking skills and dispositions (Huber & Kuncel, 2016). Education also improves economic rationality and decision-making (Kim et al., 2018). Although not central to the goal of our thesis, we thus expect college-educated people to be better at discerning fake news. The data on education has been collected on an eight-item scale (ranging from "less than high school" to "doctoral degree") to verify that the sample was balanced across treatment and control conditions. For analysis, we will collapse the variable into binary *education_higher_dummy* (equals 1 when the person has received at least a Bachelor's degree).

3.2.3 Non-experimental data

We collect a variety of non-experimental data. Age, household size and level of income (a range in $10,000 increments from $10,000 to $150,000 and above; *income_high_dummy* is equal to 1 if income is above $70,000), frequency of Facebook usage (we code the daily usage as *fb_daily_dummy*, equal to 1 if the usage is daily), prior exposure to CRT (yes/maybe/no). As previously mentioned, we collect information on whether participants answered randomly at any point in time, thought the news items were fabricated or searched any of the items on the web.

3.3 Data Analysis

According to our pre-registered analysis plan (https://osf.io/wcq8b/), we follow an established procedure in similar experiments (e.g., Pennycook & Rand, 2021; Pennycook, Binnendyk, et al., 2020; Pennycook, McPhetres, et al., 2020), we utilize Ordinary Least Squares (OLS) regression, with cluster-robust errors clustered on participant and headline level. This approach allows the model errors to be independent across clusters but correlated with them. We need this since we will be evaluating the regressions on per news article per participant level. We follow this approach for all regressions apart from the two, where we use truth and sharing discernment as to the dependent variables. In the latter case, we report heteroskedasticity-robust errors.

In addition, as a robustness check, we conduct a linear mixed-effects model analysis that is well-suited to the hierarchical (participant - news item) nature of our dataset. The advantage of a mixed-effects model is that one can specify the data-generating process explicitly - and we can account for the random effects of the participant and news item (Brown, 2021).

Finally, we use a bivariate logit model to validate the robustness results and offer an additional layer of analysis. Since this is a non-linear estimation method (Wooldridge, 2010), we collapse our dependent variables *likely* and *share* into binary variables *likely_high_dummy* and *share_high_dummy* equal to 1 for outcomes of 4 up to 6 for *likely* and *share,* respectively.

3.4 Assumptions and Randomization Checks

The assumptions check will comprise four distinct parts - randomization, heteroskedasticity, multicollinearity, and results robustness check.

To begin, we validate the randomization of the assignment of the respondents to treatment and control. We do that by Wilcoxon rank-sum and independent samples t-tests across a multitude of variables. The Table presented in Appendix III shows that the populations are statistically equal across multiple factors.

Next, we use the Breusch-Pagan test heteroskedasticity on all our main models and report heteroskedasticity and cluster robust errors where computationally feasible. See Appendix XXII for details.

In addition, we conduct multicollinearity checks via correlation matrix and utilize the typical variance inflation factor cutoff value of 10, where collinearity is strong enough to require adjustments to the model (Craney & Surles, 2002). The only variable causing major issues was otherwise insignificant *age squared,* which we removed from the model specifications. Another variable with a high variance inflation factor was *fake_dummy* (no higher than 33.3) due to the abundance of interaction effects included in the model. The problem was typically resolved in the short but parsimonious (basing our decision on Bayesian information criterion - BIC; Burnham & Anderson, 2004) model specifications. Appendices XXVII - XXI are relevant in this case.

Finally, we provide proof of the robustness of obtained results via multiple estimation methods - simple OLS with clustered standard errors, linear mixed-effects models, as well as logistic and probit estimations. In addition, where possible, we also conduct a robustness check by measuring the variable's persistence when collapsed into a dummy based on its median (denoted as variable_name_high_dummy). Full regression tables are available in Appendices XII, XIII, and XV.

Section 4 - Hypotheses

Based on the literature reviewed, we propose the following hypotheses focused on the applicability of dual-route theory on fake news processing. Additionally, we hypothesize that dual-route theory applies to truth, sharing discernment, and the efficacy of accuracy prompts. Further, a group of secondary hypotheses is aimed at heterogeneity within specific subgroups in the responses to the accuracy prompts and their news sharing intentions and truth discernment or belief in fake news[1].

4.1 Primary Hypotheses

On the efficacy of accuracy warning labels to fight fake news

As presented in the literature review, we expect that accuracy prompts will be effective against fake news in two ways. In line with previous literature (e.g., Pennycook & Rand 2021a, 2021b), we expect the accuracy prompt to increase truth discernment and decrease the sharing intentions.

1.1 Prompting people to think about accuracy will make them better discern fake news from real news.

[1] The hypotheses have been paraphrased and regrouped into primary and secondary for better readability of the manuscript as compared to the pre-registration.

1.2 Prompting people to think about accuracy will decrease social media sharing of fake news (much more so than real news).

On the confirmation bias and its impact on fake news influence

Furthermore, in line with past research (Moravec et al., 2020; Kim & Dennis 2019; Moravec et al. 2019), we expect confirmation bias, symptomatic of lower-effort cognition, and an increase in inaccurate judgment.

2. Confirmation bias increases the inclination to share fake news and decreases the likelihood of truth discernment.

On the importance of deliberate thinking in fake news discernment

Section 2 provided strong support for more effortful thinking decreasing belief in fake news. In line with previous literature on this subject, we predict that engaging in the central processing route (i.e., in a more deliberate thought process) and factors influencing the central route's activation are negatively related to belief in fake news. In addition, we expect that accuracy prompts will amplify the effect of more effortful thinking.

3. People who engage in more effortful thinking according to dual processing theory are more likely to discern between fake and true news. The effect is more pronounced with the accuracy prompt.

Specifically,

3.a People scoring high on preference for effortful thinking and scoring lower on preference for intuitive thinking are better at discerning fake news; the effect is more pronounced via accuracy prompt.

3.b People scoring higher on the cognitive reflection test are better at discerning fake news; the effect is more pronounced via accuracy prompt.

3.c People scoring higher on political knowledge are better at discerning fake news; the effect is more pronounced via accuracy prompt.

3.d People are better at discerning politically concordant fake news (i.e., in tune with their political ideology); the effect is more pronounced via accuracy prompt.

4.2 Secondary Hypotheses

The dual-process ELM also predicts additional heterogeneity in fake news influence, sharing, and truth discernment attitudes among specific subgroups.

Additional Heterogeneity in Fake News Influence

4.a Living in areas with voter preferences more concordant with the article's political valence helps better discern fake news.

4.b. People with higher education are better at discerning fake news and prefer to share it less; the effect is more pronounced via accuracy prompt.

Heterogeneity in Sharing Intentions

Although people's accuracy judgments are not directly related to their sharing intentions (see Pennycook et al., 2021), we expect the dual-process model predictions to apply in this case. Specific subgroups of people (as posited by ELM and in line with literature presented in Section 2) should be more likely to reduce their sharing intentions when exposed to the accuracy prompt, as well as less likely to share fake news.

4.c. People with a higher preference for effortful thinking and a lower preference for intuitive thinking prefer to share fake news less; the effect is more pronounced via accuracy prompt.

4.d. People scoring higher on the cognitive reflection test prefer to share fake news less when exposed to accuracy prompt.

4.e. People scoring higher on political knowledge prefer to share fake news less when exposed to accuracy prompt.

4.f. People share fake news in tune with their political ideology more, but the accuracy prompt mitigates the impact.

4.g. People from more politically polarized areas prefer to share fake news more, but the accuracy prompt mitigates the impact.

Section 5 - Results

The following chapter outlines the results of the online experiment, including a brief discussion of the pre-study (Section 5.1) and the main study (Section 5.2). The primary analysis methods used in Section 5.2 comprise Ordinary Least Regression, mixed-level models, and logistic regressions. Subsequently, Section 5.3 reviews items necessary for the post-hoc analysis of the experimental manipulation.

5.1 Pretest Results

Overall, we recruited 72 US residents via Prolific. Since the goal of the pre-test was to verify the questionnaire and test the treatment efficacy. The supposed effect size would not be detectable in a small sample like this; we will not be reviewing the results from the pre-study on the small sample in detail. However, an essential outcome of the pre-test was that both treatment and control respondents rated the treatment as possibly effective in influencing their judgments of the accuracy of news articles on a five-item scale ($\mu = 4.4$ for control, $\mu = 4.67$ for treatment). The participants who saw the accuracy prompt were slightly but insignificantly more encouraging (Welch two-sample t-test, $p = 0.12$) regarding the treatment. No significant differences between the subgroups based on sharing intentions and belief in fake news were found on the pre-test sample. In addition, thanks to the pre-test, we have reworded some items to make them more straightforward (e.g., specified that an expected answer to a CRT question is numeric only unless otherwise specified).

5.2. Main Study Results: Primary Hypotheses

5.2.1 Sample Overview

Overall, 589 US residents were recruited via Prolific. 42 reported never using Facebook, and another 27 passed the screener but did not complete the survey. According to the pre-registered criteria, these observations were excluded from our study, resulting in a final sample size of 520. The average completion time of the survey was 16.15 minutes. 64% of our sample was female, while 35% was male, and 5% reported their gender as "other". The educational background of our sample was as follows: 58 (11%) completed high school or lower, 157 (30%) had some college or associate degree, 209 (40%) had bachelor's degree, while 96 (18%) had Master's or higher postgraduate degree. 405 (77.89%) were identified as white, while 36 (5%) as African American, 41 (7.88%) as Asian, 22 (4.23%) as Latino, and 3.07% as other (ethnicity was collapsed into white and non-white subcategories for more straightforward interpretability). The mean age of the respondents was 37 (σ = 18.4 years). Participants were randomly assigned to treatment ($n_{treatment}$ = 259) via Qualtrics surveying software and control ($n_{control}$ = 261). The vast majority (506 or 97.31%) of our respondents passed at least one attention check, while 472 (90.7%) respondents answered all of them correctly. Randomization checks showed no significant difference across a multitude of factors. Please see the summary of the randomization checks and sample descriptive statistics in Appendix I and II.

5.2.2 Ordinary Least Squares Analysis - Primary Hypotheses

In this subsection, we will explore the data via Ordinary Least Squares (OLS) regression with the variable *likely* denoting the reported likelihood of the news article being true (from 1 to 6, 6 being most likely to be true) or *share* (from 1 to 6, 6 being most likely to be shared). To build our final model, we firstly proceed by testing the multiple moderation hypotheses. The hypotheses will test the link between various variables signifying effortful thinking according to the dual-route model and (dis)belief in fake news. Later, these will be included in our final model. The effects of all variables persist even after controlling for other variables. We report cluster-robust standard errors for improved robustness, with clustering on the participant and news item level as suggested by Pennycook et al. (in press).

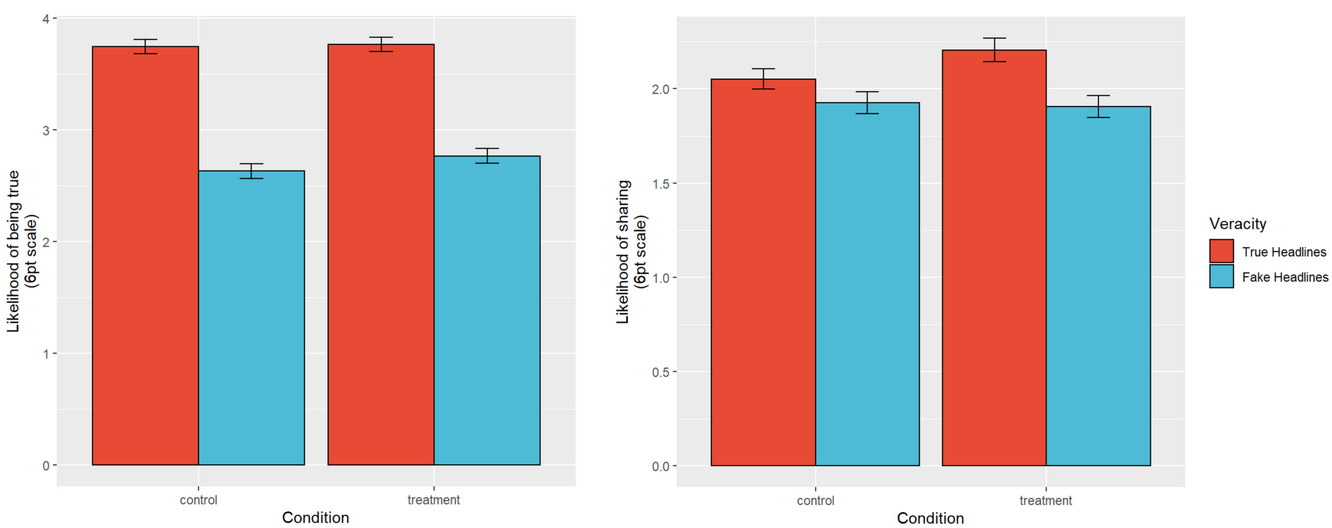


Figure 2: The effect of the treatment on sharing intention and news belief by condition and veracity

The role of the accuracy prompt on the fake news belief and sharing intentions

Table 5.1 The efficacy of the accuracy prompt treatment on belief in the news (likely; ranges from 1 - 6) and sharing intentions (share; ranges from 1-6)

Dependent variables: belief in the news (likely; ranges from 1 - 6) and sharing intentions (share; ranges from 1-6)

	Likely		Share	
(Intercept)	3.754***	3.744***	2.126***	2.052***
	(0.023)	(0.033)	(0.033)	(0.043)
fake_dummy	-1.056***	-1.113***	-0.210***	-0.126*
	(0.033)	(0.046)	(0.046)	(0.062)
treatment_dummy		0.020		0.154*
		(0.046)		(0.066)
fake_dummy × treatment_dummy		0.115+		-0.174+
		(0.066)		(0.093)
Num.Obs.	9360	9360	3816	3816
R2	0.099	0.100	0.005	0.007
R2 Adj.	0.099	0.100	0.005	0.006
BIC	35307.9	35317.6	13569.1	13579.9
F	1027.860	345.717	20.676	8.788

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

No significant effect of the treatment on ratings of the likelihood to be true ($\chi2 = 7.8907$; p-value = 0.1624) or sharing ($\chi2 = 1.964$; p-value = 0.8541) was found via the chi-squared test (as Figure

2 illustrates[2]). The estimates coming from the conducted regressions confirm this result. In further regression analyses, we include the interaction term to test out additional hypotheses. However, we do not find a meaningful impact of the treatment on any of the variables mentioned in our hypotheses for both likelihood to believe the news and share it (Table 5.1).

The role of the confirmation bias

As expected, confirmation bias (see Table 5.2), regardless of the treatment, slightly increases news believability (model 2.1; $\beta = 0.020$, $p < 0.001$), especially in case of fake news (model 2.1; $\beta = 0.051$, $p < 0.001$).

Table 5.2 The role of confirmation bias on the likelihood to believe in fake news

Dependent variable: belief in the news (likely; ranges from 1 - 6)		
	2.1	2.2
(Intercept)	3.819***	3.799***
	(0.024)	(0.034)
fake_dummy	-0.959***	-1.000***
	(0.033)	(0.047)
bias	0.020***	0.016***
	(0.003)	(0.005)
fake_dummy × bias	0.051***	0.055***
	(0.005)	(0.007)
treatment_dummy		0.039

[2] For bar plots, error bars represent means of dependent variables and associated 95% confidence intervals for various subgroups based on categorical variables.

		(0.047)
fake_dummy × treatment_dummy		0.085
		(0.065)
bias × treatment_dummy		0.007
		(0.006)
fake_dummy × bias × treatment_dummy		-0.007
		(0.009)
Num.Obs.	9359	9359
R2	0.156	0.157
R2 Adj.	0.156	0.156
BIC	34708.4	34735.9
F	577.350	248.874

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

The role of the Cognitive Reflection Test performance on believing in fake news

CRT performance seems to lead to greater belief in true news and better identification of fake news (model 3.a.1; $\beta = -0.121$, $p < 0.001$). Notably, the effect on fake news discernment is much more pronounced. This effect is replicated by treating CRT as a continuous (crt_score) and binary (crt_score_high_dummy) variable. Overall, people scoring high on CRT (Table 5.3) are better at discerning both true and fake news.

Table 5.3 The role of Cognitive Reflection Test performance on believing in fake news

Dependent variable: belief in the news (likely; ranges from 1 - 6)

	3a.1	3a.2	3a.3	3a.4
(Intercept)	3.620***	3.560***	3.671***	3.639***
fake_dummy	-0.622***	-0.744***	-0.858***	-0.971***

crt_score	0.037*	0.053*		
fake_dummy × crt_score	-0.121***	-0.106**		
treatment_dummy		0.130		0.067
fake_dummy × treatment_dummy		0.269		0.238+
crt_score × treatment_dummy		-0.033		
fake_dummy × crt_score × treatment_dummy		-0.035		
crt_score_high_dummy			0.162*	0.214*
fake_dummy × crt_score_high_dummy			-0.386***	-0.291*
crt_score_high_dummy × treatment_dummy				-0.107
fake_dummy × crt_score_high_dummy × treatment_dummy				-0.205
Num.Obs.	9360	9360	9360	9360
R2	0.105	0.107	0.102	0.105
R2 Adj.	0.104	0.106	0.102	0.104
F	364.947	159.745	355.633	155.917

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The role of the preference towards effortful and intuitive thinking on believing in fake news

Next, we review the effect of preference for effortful (PET) and preference for intuitive (PIT) thinking (Table 5.4). Both seem to influence news credibility in the supposed way. On the one hand, PIT is related to slightly increased belief when the news is fake. On the other hand, people scoring high on PET seem to be skeptical of fake news, but the effect seems to disappear after controlling for additional variables. Overall, effortful thinkers are more confident in true news, while intuitive thinkers are more doubtful of them. The table below shows that both effect sizes are moderately small, but PIT's impact is much more persistent.

Table 5.4 The role of the preference towards effortful and intuitive thinking on believing in fake news

Dependent variable: belief in the news (likely; ranges from 1 - 6)				
	3b.1	3b.2	3b.3	3b.4
(Intercept)	3.776***	3.723***	3.750***	3.834***
fake_dummy	-1.044***	-1.046***	-1.288***	-1.489***
pit_centered	-0.057*	-0.059		
pet_centered	0.057*	0.003		
fake_dummy × pit_centered	0.183***	0.199***		
fake_dummy × pet_centered	-0.065*	-0.016		
treatment_dummy		0.101		-0.168
fake_dummy × treatment_dummy		0.011		0.417+
pit_centered × treatment_dummy		0.000		
pet_centered × treatment_dummy		0.101+		
fake_dummy × pit_centered × treatment_dummy		-0.032		
fake_dummy × pet_centered × treatment_dummy		-0.093		
pit_high_dummy			-0.139	-0.125
pet_high_dummy			0.180*	-0.006
fake_dummy × pit_high_dummy			0.498***	0.600***
fake_dummy × pet_high_dummy			-0.211*	-0.073
pit_high_dummy × treatment_dummy				-0.032
pet_high_dummy × treatment_dummy				0.373**
				(0.145)
fake_dummy × pit_high_dummy × treatment_dummy				-0.213

fake_dummy × pet_high_dummy × treatment_dummy				-0.281
Num.Obs.	9360	9360	9360	9360
R2	0.108	0.110	0.106	0.109
R2 Adj.	0.108	0.109	0.106	0.108
BIC	35248.9	35284.0	35269.4	35293.0
F	226.795	105.033	222.200	104.114

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The role of the political knowledge on believing in fake news

Furthermore, in Table 5.5, we find that people scoring high political knowledge (model 3.c.1; $\beta = -0.121$, $p < 0.001$) are more confident in their ratings of news articles (model 3.c.1; $\beta = 0.117$, $p < 0.001$), and more sceptical of fake news articles (model 3.c.1; $\beta = -0.298$, $p < 0.001$). Political knowledge effect stays the same regardless of the accuracy prompt (model 3.c.2).

Table 5.5 The role of the political knowledge on believing in fake news

Dependent variable: belief in the news (likely; ranges from 1 - 6)				
	3c.1	3c.2	3c.3	3c.4
(Intercept)	3.407***	3.437***	3.672***	3.679***
fake_dummy	-0.176	-0.215	-0.701***	-0.754***
political_knowledge	0.117***	0.102+		
fake_dummy × political_knowledge	-0.298***	-0.298***		
treatment_dummy		-0.063		-0.014
fake_dummy × treatment_dummy		0.068		0.105
political_knowledge × treatment_dummy		0.033		
fake_dummy × political_knowledge × treatment_dummy		0.005		
political_knowledge_high_dummy			0.119	0.094

fake_dummy × political_knowledge_high_dummy			-0.516***	-0.518***
political_knowledge_high_dummy × treatment_dummy				0.051
fake_dummy × political_knowledge_high_dummy × treatment_dummy				0.007
Num.Obs.	9360	9360	9360	9360
R2	0.109	0.109	0.105	0.106
R2 Adj.	0.108	0.109	0.105	0.106
BIC	35225.7	35254.3	35258.1	35285.9
F	379.888	164.020	367.815	158.965

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The role of the concordance with political ideology on believing in fake news

Political concordance results in overall increased belief in the news (Table 5.6). The effect is positive both via measuring concordance of individual with the article topic (model 3.d.1; $\beta = 0.205$, $p < 0.001$), as well as the voting preference of the county the individual is residing in with the news headline (model 3.d.1; $\beta = 0.223$, $p < 0.001$). The results are robust regardless of treating neutral news as a separate variable (see models 3.d.3 and 3.d.4). Contrary to expectation, personal concordance decreases the likelihood of truth discernment. Personally concordant fake news tends to be evaluated as more credible than true news (model 3.d.1; $\beta = 0.883$, $p < 0.001$). However, when fake news is concordant with county-level political preferences, respondents better discern them from the truth (model 3.d.1; $\beta = 0.233$, $p < 0.01$). We will explore these results closer in the discussion section.

Table 5.6 The role of the political concordance on believing in fake news

Dependent variable: belief in the news (likely; ranges from 1 - 6)				
	3d.1	3d.2	3d.3	3d.4
(Intercept)	3.652***	3.638***	3.673***	3.681***
	(0.031)	(0.045)	(0.031)	(0.045)
fake_dummy	-1.318***	-1.338***	-1.165***	-1.184***
	(0.043)	(0.061)	(0.043)	(0.063)
concordance_dummy	0.205***	0.204**		
	(0.050)	(0.070)		
concordance_zipcode_dummy	0.223***	0.261**	0.203***	0.231**
	(0.061)	(0.085)	(0.061)	(0.086)
fake_dummy × concordance_dummy	0.883***	0.837***		
	(0.071)	(0.100)		
fake_dummy × concordance_zipcode_dummy	-0.233**	-0.378**	-0.333***	-0.491***
	(0.086)	(0.118)	(0.088)	(0.120)
treatment_dummy		0.027		-0.015
		(0.062)		(0.062)
fake_dummy × treatment_dummy		0.043		0.043
		(0.086)		(0.087)
concordance_dummy × treatment_dummy		0.005		
		(0.100)		
concordance_zipcode_dummy × treatment_dummy		-0.079		-0.066
		(0.123)		(0.123)
fake_dummy × concordance_dummy × treatment_dummy		0.081		
		(0.141)		
fake_dummy × concordance_zipcode_dummy × treatment_dummy		0.306+		0.335+
		(0.173)		(0.176)

concordance2_dummy			0.151**	0.091
			(0.049)	(0.070)
fake_dummy × concordance2_dummy			0.473***	0.437***
			(0.071)	(0.100)
concordance2_dummy × treatment_dummy				0.122
				(0.099)
fake_dummy × concordance2_dummy × treatment_dummy				0.061
				(0.142)
Num.Obs.	9234	9234	9234	9234
R2	0.149	0.150	0.118	0.120
R2 Adj.	0.149	0.149	0.117	0.119
BIC	34360.2	34401.8	34693.5	34730.5
F	323.519	148.367	246.620	113.859
Std. Errors	Robust	Robust	Robust	Robust

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

5.2.3 Ordinary Least Squares Analysis - Secondary Hypotheses

The role of education on believing or sharing fake news

We find that higher education (i.e., obtaining Bachelor's or Master's degree) is better linked to fake news discernment (model 4.a.1; $\beta = -0.294^{**}$, $p < 0.01$). See Appendix V for the table.

The role of the preference towards effortful and intuitive thinking on sharing fake news

When it comes to sharing, PIT impacts sharing intentions only in case the news is fake (model 4.b.2; $\beta = 0.129^{*}$, $p < 0.05$). PET does not seem to impact news sharing - in general, or when exposed to fake news. See Appendix VI for the table.

The role of the CRT performance on sharing fake news

In Appendix VII, we show that CRT performance is generally related to lower news sharing intentions (model 4.c; $\beta = -0.061^{**}$, $p < 0.001$). In the case of fake news, above median CRT performance seems to be predictive of a lower likelihood to share the news (model 4.c; $\beta = -0.234^{**}$, $p < 0.001$). Nevertheless, the effect seems to require closer investigation in the full model since its significance is variable.

The role of political knowledge in sharing fake news

Additionally, greater political knowledge decreases the likelihood of sharing (model 4.d.1; $\beta = -0.117^{***}$, $p < 0.001$), with a less pronounced effect on fake news articles (significant difference only between individuals scoring below and above the median - model 4.d.4; $\beta = -0.185$; $p < 0.1$) which will be further scrutinized in the full specification. See Appendix VIII for the table.

The role of the concordance with political ideology on sharing fake news

Regardless of the way it was coded, personal concordance significantly increases the inclination to share news (model 4.e.1; $\beta = 0.43^{***}$, $p < 0.001$), even those that are fake (model 4.e.1; $\beta = 0.242^{*}$,

$p < 0.05$). The county-level concordance does not seem to show any significant effect in the restricted specification of the model (see Appendix IX for detail).

The role of the politically polarized environment in sharing fake news

Finally, we find no effect of a polarized environment on sharing attitudes, not even in the case of fake news (Appendix X).

5.2.4 Ordinary Least Squares Analysis - Summary of the Models

Before we move to the final specification of the models, for illustrative purposes, we present the overview of all crucial variables without the pre-registered control variables and interaction terms. While the significance of some previously significant variables has declined (CRT, PIT, political knowledge, and county-level concordance), this is expected due to omitting some known significant interaction effects with the *fake* dummy variable that will be added later.

Table 5.7 The main variables of our models, no interactions

Dependent variables: belief in the news (likely; ranges from 1 - 6) and sharing intentions (share; ranges from 1-6)		
	likely	share
(Intercept)	3.700***	1.036***
	(0.074)	(0.103)
fake_dummy	-1.097***	0.166***
	(0.032)	(0.044)
treatment_dummy	0.061+	0.123**
	(0.032)	(0.041)

bias	0.037***	0.024***
	(0.002)	(0.003)
crt_score	-0.012	-0.061***
	(0.009)	(0.010)
political_knowledge	-0.012	-0.110***
	(0.015)	(0.019)
education_higher_dummy	-0.130***	-0.133**
	(0.034)	(0.042)
pit_centered	0.008	0.006
	(0.012)	(0.016)
pet_centered	0.038***	-0.021
	(0.011)	(0.015)
concordance_dummy	0.523***	0.201***
	(0.036)	(0.045)
prefers_republican	0.047***	0.092***
	(0.010)	(0.012)
concordance_zipcode_dummy	0.050	-0.012
	(0.043)	(0.053)
polarized_zipcode_dummy	-0.040	-0.104*
	(0.043)	(0.051)
likely		0.353***
		(0.014)
Num.Obs.	9233	3762
R2	0.168	0.293
R2 Adj.	0.167	0.290
F	155.181	119.307

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We now proceed with full and reduced specifications of our models, demonstrated below. In the entire specification, we also control age, income, Facebook usage, and respondent ethnicity[3]. The only deviation from our pre-registration is the inclusion of the interaction of *crt_score* and (*confirmation) bias* and an individual's political *concordance* to account for the motivated reasoning (reviewed in Section 2.2). The full version of the table is available in Appendix XI. The most prominent effects are summarized in Appendices XXIII and XXIV.

Table 5.8 Final specification: reduced and full models. What drives fake news belief and sharing intentions?

Dependent variables: belief in the news (likely; ranges from 1 - 6) and sharing intentions (share; ranges from 1-6)

	likely (full)	likely (reduced)	share (full)	share (reduced)
(Intercept)	3.545***	3.461***	1.402***	1.370***
	(0.119)	(0.103)	(0.175)	(0.126)
fake_dummy	-0.982***	-0.856***	-0.087	-0.071
	(0.157)	(0.130)	(0.229)	(0.102)
treatment_dummy	0.028	0.057+	0.189**	0.184**
	(0.046)	(0.031)	(0.062)	(0.061)
bias	0.020***	0.019***	0.035***	0.037***
	(0.005)	(0.003)	(0.007)	(0.007)
crt_score	0.019	0.021	-0.061***	-0.062***
	(0.015)	(0.014)	(0.018)	(0.011)
political_knowledge	0.098***	0.097***	-0.098**	-0.108***

[3] These variables are not central to the research, thus the tables with the full variable list are reported only in the Appendix.

	(0.022)	(0.022)	(0.030)	(0.019)
education_higher_dummy	-0.052	-0.068	-0.118+	-0.121**
	(0.050)	(0.048)	(0.066)	(0.044)
pit_centered	-0.047**	-0.049**	-0.011	
	(0.018)	(0.017)	(0.025)	
pet_centered	0.049**	0.036***	-0.022	-0.021
	(0.015)	(0.010)	(0.022)	(0.015)
concordance_dummy	0.239**	0.241**	0.347***	0.282***
	(0.082)	(0.081)	(0.103)	(0.065)
prefers_republican	-0.036*	-0.039**	0.110***	0.097***
	(0.015)	(0.015)	(0.019)	(0.012)
concordance_zipcode_dummy	0.196**	0.185**	-0.079	-0.084
	(0.063)	(0.061)	(0.080)	(0.079)
polarized_zipcode_dummy	-0.064		-0.149+	-0.098+
	(0.062)		(0.079)	(0.052)
fake_dummy × treatment_dummy	0.062		-0.122	-0.115
	(0.063)		(0.082)	(0.079)
fake_dummy × bias	0.034***	0.034***	0.001	
	(0.005)	(0.005)	(0.007)	
fake_dummy × crt_score	-0.052**	-0.049**	0.010	
	(0.018)	(0.017)	(0.021)	
fake_dummy × political_knowledge	-0.205***	-0.209***	-0.002	
	(0.030)	(0.029)	(0.040)	
fake_dummy × education_higher_dummy	-0.157*	-0.137*	-0.005	
	(0.068)	(0.065)	(0.088)	
fake_dummy × pit_centered	0.104***	0.109***	0.034	
	(0.024)	(0.023)	(0.032)	
fake_dummy × pet_centered	-0.026		0.003	
	(0.021)		(0.030)	

fake_dummy × concordance_dummy	0.696***	0.697***	-0.205*	-0.200*
	(0.071)	(0.071)	(0.092)	(0.091)
fake_dummy × prefers_republican	0.174***	0.179***	-0.024	
	(0.020)	(0.020)	(0.025)	
fake_dummy × concordance_zipcode_dummy	-0.304***	-0.292***	0.167	0.178+
	(0.084)	(0.082)	(0.107)	(0.104)
fake_dummy × polarized_zipcode_dummy	0.063		0.095	
	(0.084)		(0.105)	
bias × crt_score	0.000		-0.004*	-0.004*
	(0.001)		(0.002)	(0.002)
crt_score × concordance_dummy	-0.023	-0.024	-0.017	
	(0.018)	(0.017)	(0.021)	
likely			0.318***	0.317***
			(0.019)	(0.019)
fake_dummy × likely			0.079**	0.079**
			(0.029)	(0.028)
Num.Obs.	9215	9215	3744	3744
R2	0.215	0.214	0.299	0.298
R2 Adj.	0.212	0.212	0.293	0.294
F	75.998	119.079	45.253	75.232

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The reduced versions of the models were achieved by removing the least significant terms in such a way that does not lead to a significant drop in explanatory power as indicated by the F-test of linear restrictions (F = 0.6812914, p-value = 0.64 for the model with the dependent variable *likely* and F = 0.4591, p-value = 0.9386 for the model with the dependent variable *share*). Hereafter, however, we cover the full model results (since it contains some of the insignificant interactions we are interested in).

The results for the truth discernment can thus be summarized as follows. People do discern between fake and true news (β = -0.982***, $p < 0.001$), downrating them by 16.37% on average. We find a consistently negative impact of confirmation bias on overall belief in articles, especially fake ones (β = 0.02***, $p < 0.001$). To illustrate the impact of bias: an individual with the most positively "biased" attitude (+21) on average increases the truthfulness rating of the article by 7%. Political knowledge increases confidence in judging the political headlines overall (β = 0.098***, $p < 0.001$), offering a better discernment of fake news (β = -0.205***, $p < 0.001$). This effect is rather remarkable, meaning people scoring a maximum of 5 points on political knowledge assigned on average 17.08% lower trustworthiness scores to the fake articles versus the true articles. Furthermore, political concordance of the article with individual's political identity increases the overall belief in news (β = 0.239***, $p < 0.001$), and leads to worse truth discernment contrary to our expectations (β = 0.696***, $p < 0.001$). The effect is meaningful in magnitude: people, on average, rate discordant fake news 15.58% less trustworthy than identity-concordant fake news. Similar to individual-level concordance, county-level concordance of the fake news article promotes overall belief in news (β = 0.196**, $p < 0.01$), but its effect reverses and is significant in case of false articles (β = -0.304*** $p < 0.001$). Preference towards Republican party encourages true news scepticism (β = -0.036*, $p < 0.05$), and fake news belief (β = 0.174***, $p < 0.001$). An avid Republican (equals 7) rated fake news 17.4% more trustworthy than a strong Democrat (equals 1). Next, PIT and PET affect the overall belief in the news with significant and opposite virtually equal effects. The difference arises while interacting with fake news - the only PIT is significant and increases the believability of fake news ($\beta = 0.104^{***}$, $p < 0.001$). CRT performance affects truth discernment through better discernment of fake news ($\beta = -0.052^{*}$, $p < 0.01$). Best performers on CRT with a maximum score of 7 rated fake articles 3.64% less likely to be true, ceteris paribus. Finally, higher education (Bachelor's and Master's degrees) results in slightly better discernment of fake news: ($\beta = -0.157^{*}$, $p < 0.05$). All other variables are not significant. This model explains roughly 21.5% of the variability in our data.

When it comes to sharing predisposition, we find that the treatment effect is present, but only for true news ($\beta = 0.184^{***}$, $p < 0.001$). Apart from that, individual concordance has the most pronounced, positive effect ($\beta = 0.347^{***}$, $p < 0.001$), with a weak positive impact on fake news scepticism ($\beta = -0.205^{*}$, $p < 0.05$). Second strongest was the impact of the likelihood to consider the news true ($\beta = 0.318^{***}$, $p < 0.001$), even if it is fake ($\beta = 0.079^{**}$, $p < 0.01$). Similar to CRT performance ($\beta = -0.061^{***}$, $p < 0.001$) Political knowledge decreases sharing of the news overall ($\beta = -0.098^{**}$, $p < 0.01$). In contrast, confirmation bias is strengthens sharing, ($\beta = 0.035^{***}$, $p < 0.001$), but not for the fake news. Republicans also share more news, but no significant difference across the aisle was found for the fake news. As expected, fake news are shared more, but not by much ($\beta = 0.079^{**}$, $p < 0.01$). Finally, we find a very negligible ($\beta = -0.004^{*}$, $p < 0.05$) effect of interaction between confirmation bias and CRT. The model explains about 29.4% of the variability in the dataset.

Lastly, to verify the robustness of obtained OLS results, we conduct linear mixed model analysis to control both fixed and random effects given the non-independence in our data. The results are virtually the same and are presented in Appendix XII. We also provide the full version of the regression in Appendix XIII and the z-scaled estimation for the ease of interpretation in Appendix XIV.

5.2.5 Ordinary Least Squares analysis with alternative truth and sharing discernment measures

According to our pre-registration, we also use a different measure for the accuracy of respondents' judgments in truth and sharing likelihood ratings, *truth discernment,* and *sharing discernment.* Due to the nature of the dependent variable, our analysis was conducted on a per-participant (instead of per item per participant level). In the table below, we present the results of the regressions with heteroskedasticity-robust standard errors.

Table 5.9 Alternative measures. What drives truth discernment and sharing discernment?

Dependent variables: truth (ranges from -100 - 100%) and sharing discernment (ranges from -100 - 100%)		
	Truth discernment	Sharing discernment
(Intercept)	30.071**	9.050
	(10.134)	(8.909)
treatment_dummy	-1.668	6.623*
	(2.340)	(2.684)
crt_score	2.130**	0.060
	(0.679)	(0.717)
political_knowledge	5.533***	1.237

	(1.352)	(1.480)
prefers_republican	-4.173***	-0.545
	(0.694)	(0.676)
pit_centered	-2.892***	-1.391
	(0.852)	(0.978)
pet_centered	0.744	-0.069
	(0.734)	(0.988)
age	-0.037	-0.111
	(0.246)	(0.121)
race_white_dummy	0.316	-0.848
	(3.493)	(3.471)
income_high_dummy	-2.666	-0.882
	(2.508)	(2.694)
education_higher_dummy	2.625	1.435
	(2.595)	(2.641)
facebook_daily_dummy	-0.987	0.488
	(2.385)	(2.968)
polarized_zipcode_dummy	-1.318	2.067
	(3.245)	(3.526)
Num.Obs.	512	208
R2	0.221	0.081
R2 Adj.	0.202	0.024
F	11.804	1.430

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

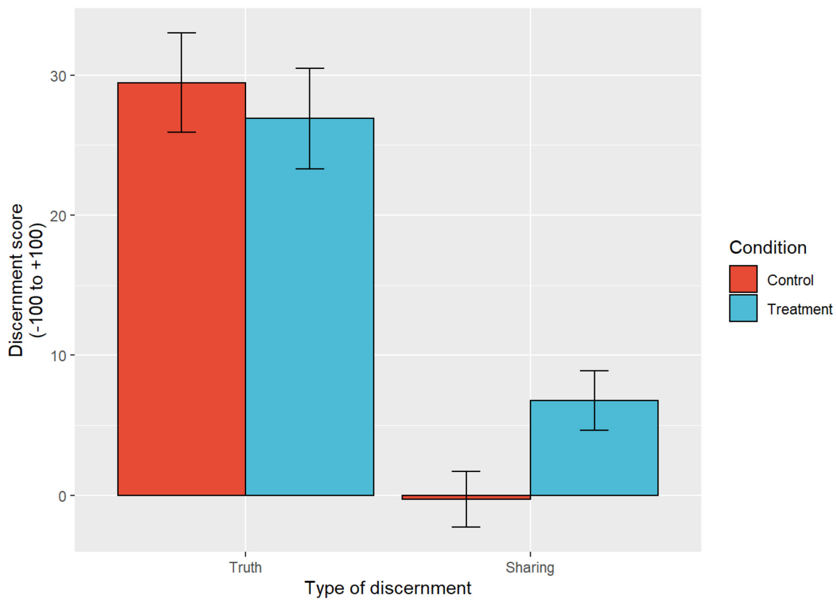


Figure 3: Sharing discernment versus truth discernment, by condition

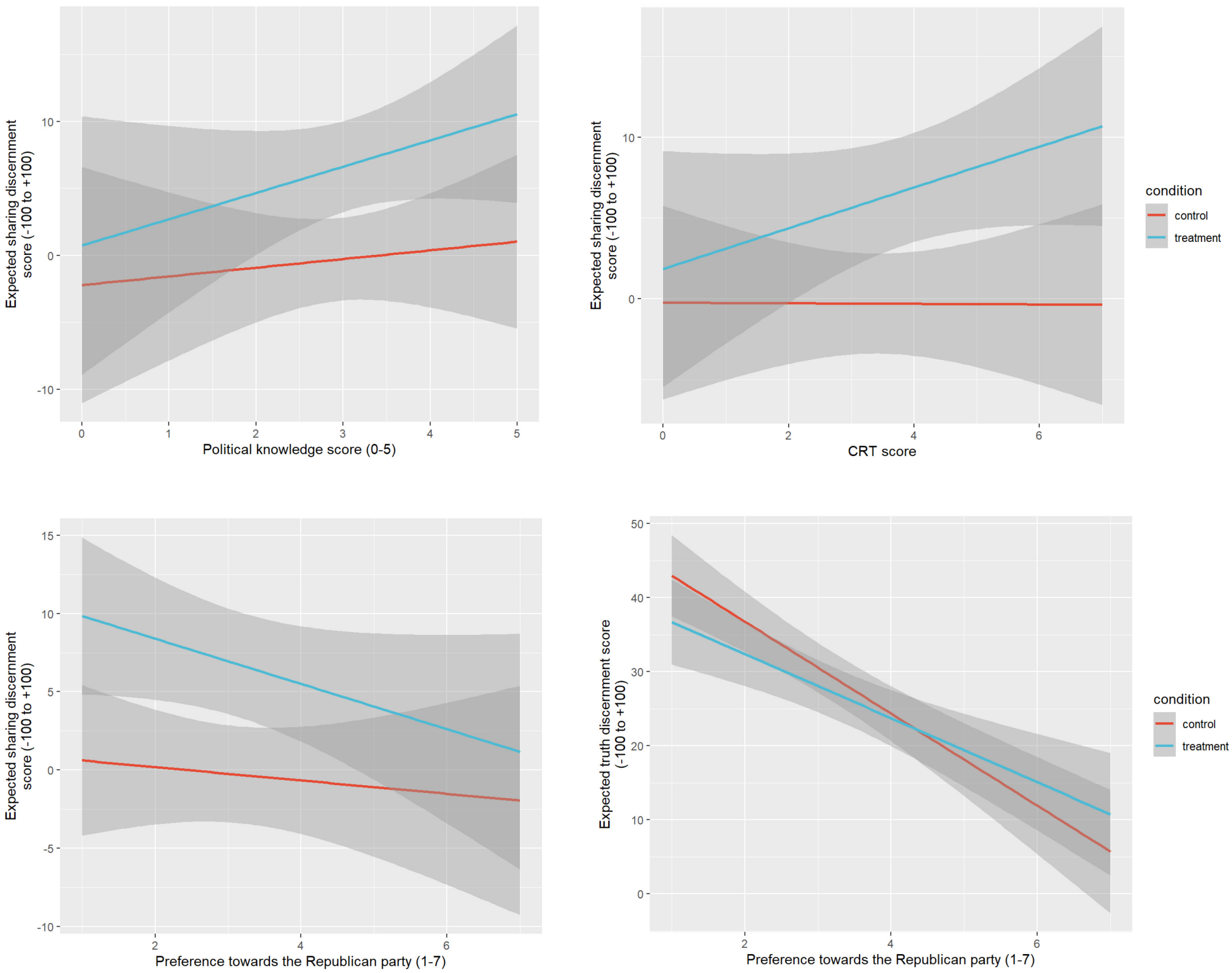
The results of these regressions virtually agree with the results obtained before. Notably, on average, truth discernment is much higher than sharing discernment (see Figure 3). Apart from the *education_higher* dummy, wherever the interaction between *fake_dummy* and *variable X* was significant in the previous OLS regressions, *variable X* is substantial in the discernment models. This is most probably because higher education had an additional effect only on discerning false (not true) news. In contrast, the other variables contributed to discernment by affecting both true and fake news's credibility (or sharing intention). The impact of political knowledge is strong: all factors fixed, maximum performance (5/5) could lead to 27.6% better choices. CRT seems to affect truth discernment only by better detecting fake articles, with a maximum gain of 14.91% if scoring maximum on the task. On the contrary, PIT reduces truth discernment, while preference towards

Republicans affects news consumption through distrust in true news and increased support of fake news (β = -4.173***, $p < 0.001$).

Ultimately, we find support that the treatment leads to better sharing, but not truth discernment, through increased belief in true news: participants exposed to the treatment recognized the truth 6.62% better than respondents assigned to control. In Figure 4, we present some of the most notable differences in truth and sharing discernment[4]. Furthermore, an in-depth breakdown of sharing and truth discernment by control and treatment conditions in case of fake and true news headlines is presented in Appendices XXIII and XXIV.

[4] The following plots represent smoothed conditional means of key dependent variables at various levels of some explanatory variables. Conditional means represent predictions of bivariate OLS regressions. 95% confidence bands around the regression lines were used to visualize uncertainty around the estimated conditional means.

Figure 4: The impact of CRT, political knowledge, and political ideology on sharing (first row, second row - left) and truth (second row - right) discernment



5.2.6 Two-value Analysis via Logistic Regressions

We now turn to study the propensity of the individuals to rate articles true or share them (binary variables *likely_high_dummy* and *share_high_dummy* with outcomes of 4 up to 6 for *likely* and *share* respectively equal to 1) in contrast to rate articles as false and not share them (ratings lower than 4) via logit estimation. This analysis is intended both as a robustness check for previously obtained results and an opportunity to scrutinize further personal characteristics of individuals who assign higher versus lower scores to the dependent variables of interest. A shortened version of the table is presented below.

Table 5.10 Robustness check: reduced and full logit models. What drives fake news belief and sharing intentions?

Dependent variables: likely_high_dummy (binary; 1 if likely >3) and share_high_dummy ((binary; 1 if share >3)

	likely (full)	likely (reduced)	share (full)	share (reduced)
(Intercept)	0.263+	0.141	-3.380***	-3.561***
	(0.149)	(0.137)	(0.416)	(0.294)
fake_dummy	-1.163***	-0.971***	0.107	0.533***
	(0.224)	(0.188)	(0.586)	(0.145)
treatment_dummy	0.041	0.073	0.574***	0.550***
	(0.063)	(0.046)	(0.138)	(0.130)
bias	0.026***	0.026***	0.035***	0.036***
	(0.004)	(0.004)	(0.008)	(0.006)
crt_score	0.023	0.021	-0.121***	-0.130***
	(0.017)	(0.017)	(0.037)	(0.026)

political_knowledge	0.118***	0.113***	-0.193**	-0.171***
	(0.030)	(0.030)	(0.065)	(0.047)
education_higher_dummy	-0.100	-0.122+	-0.076	-0.156
	(0.068)	(0.066)	(0.143)	(0.102)
pit_centered	-0.079***	-0.082***	0.007	
	(0.024)	(0.023)	(0.054)	
pet_centered	0.066**	0.053***	-0.042	
	(0.020)	(0.015)	(0.049)	
concordance_dummy	0.141*	0.140*	0.388**	0.355***
	(0.068)	(0.068)	(0.136)	(0.101)
prefers_republican	-0.030	-0.037*	0.169***	0.154***
	(0.019)	(0.019)	(0.038)	(0.028)
concordance_zipcode_dummy	0.234**	0.234**	-0.187	
	(0.088)	(0.088)	(0.177)	
polarized_zipcode_dummy	-0.079	-0.090	-0.181	-0.315*
	(0.084)	(0.084)	(0.188)	(0.143)
facebook_daily_dummy	0.116+	0.136**	-0.081	
	(0.064)	(0.047)	(0.157)	
fake_dummy × treatment_dummy	0.075		-0.416*	-0.436*
	(0.093)		(0.208)	(0.196)
fake_dummy × bias	0.048***	0.048***	0.002	
	(0.007)	(0.007)	(0.013)	
fake_dummy × crt_score	-0.096***	-0.089***	0.009	
	(0.026)	(0.025)	(0.055)	
fake_dummy × political_knowledge	-0.244***	-0.244***	0.054	
	(0.045)	(0.044)	(0.097)	
fake_dummy × education_higher_dummy	-0.196+	-0.169+	-0.123	
	(0.101)	(0.097)	(0.215)	
fake_dummy × pit_centered	0.147***	0.155***	0.066	

	(0.035)	(0.034)	(0.086)	
fake_dummy × pet_centered	-0.027		0.006	
	(0.030)		(0.075)	
fake_dummy × concordance_dummy	0.954***	0.953***	-0.130	
	(0.100)	(0.099)	(0.208)	
fake_dummy × prefers_republican	0.204***	0.214***	-0.045	
	(0.029)	(0.028)	(0.057)	
fake_dummy × concordance_zipcode_dummy	-0.480***	-0.477***	0.228	
	(0.133)	(0.133)	(0.278)	
fake_dummy × polarized_zipcode_dummy	0.171	0.183	-0.212	
	(0.125)	(0.125)	(0.297)	
likely			0.687***	0.720***
			(0.059)	(0.040)
fake_dummy × likely			0.084	
			(0.083)	
Num.Obs.	9215	9215	3744	3744
AIC	11100.1	11090.3	2794.8	2764.5
BIC	11328.2	11261.4	3006.5	2851.7
Log.Lik.	-5518.031	-5521.142	-1363.384	-1368.244

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Given that the results are in almost total agreement[5], we only discuss four interesting findings obtained during this estimation. For sharing, the treatment seems to be significant in the general increase of sharing propensity ($p < 0.001$) as well as in the case of fake news flagged with the accuracy prompt ($p < 0.05$). Although the sharing discernment of fake news somewhat increases, the result is neither highly significant (linear hypothesis test, $\chi 2 = 0.5997387$, $p = 0.4386$) nor replicated in other types of analyses. We thus may conclude that the effect of treatment on fake news is limited only to true news. In contrast, the effect on fake news is inconclusive but not substantially big in any case due to our considerable sample size.

Interestingly, while the veracity of the headlines matters when people make the accuracy judgment ($p < 0.001$), it does not have any impact on the sharing intentions (as corroborated by our simple OLS and mixed-effects estimation). Finally, county-level concordance seems to increase overall support for news aligning with ideology but decreases the influence of fake news. A full review of obtained results and their support of the hypotheses outlined in Section 4 can be found in Appendix XXV. All results will be further discussed in Section 6.

[5] There is a significant 5% (but not 1%) difference in the likelihood test for the reduced and the full *likely* logit models. We made sure that reduced models are more parsimonious (lower BIC), but have an insignificantly lower explanatory power (at the 1% level).

5.3 Post-hoc Analysis of the Treatment's Ineffectiveness

The post-hoc power analysis yields an effect size for the impact of treatment on belief in fake news is $d = 0.351$, which Cohen (1992) would classify as medium ($d < 0.4$). Similarly, for sharing discernment, $d = 0.307$. Although the 95% confidence intervals constructed with heteroskedasticity-robust errors show different results (significance for *likely* - [0.009, 0.145]; insignificance for *share* - [-0.024, 0.158]), we have demonstrated that this subsides and reverses when we control for other factors.

To inform our further discussion regarding the ineffectiveness of the treatment in the next chapter, we explore three results (one of them, prediction of the first click on the question's response window - first_click, pre-registered). First of all, we find no significant increase in response time and the addition of treatment (Appendix XVI). In a similar vein, there seems to be no significant increase in time taken from the display of the questions until the last click on the screen (which may indicate additional deliberation; Moravec, 2020). Lastly, the median for the variable *position*, a proxy for the affective stance of the respondent when exposed to a news item, does not differ by condition significantly ($W = 10703368$, p-value $= 0.05363$). There is, however, a significant difference in means according to the t-test ($p < 0.001$), but it is not practically substantial - just a fraction of the standard deviation. The significance may mainly be driven by the large sample size caused by repeated measures in the long dataset.

Section 6 - Discussion

To finalize the thesis, Section 6.1 reviews the results, while Section 6.2 explores the limited effect of the accuracy prompt as a treatment. Sections 6.3 and 6.4 discuss the practical and theoretical implications of the research. Finally, Section 6.5 lists the limitations of the study and provides suggestions for further research.

6.1 Results Summary

Before turning to the discussion of results, it feels appropriate to summarize them. Broadly speaking, our results speak in favor of the dual-process model predictions based on ELM.

Firstly, we find no significant evidence of the accuracy prompt helping to increase truth discernment. However, we do find an indication of its effect on sharing discernment, particularly for true news. Although people are generally good at distinguishing between fake and true news, they are worse at discerning which news can be shared. Importantly, they do not seem to care about the veracity of news while sharing it, in contrast to when judging the likelihood of the headline being true. The following may be symptomatic of "lazy thinking" - failing to consider the accuracy of the headline before sharing it. Next, confirmation bias negatively affects decision-making, more so when the news is fake.

Furthermore, ability components (political knowledge, CRT, higher education) indicated more effortful thinking. Effects were pronounced when considering the likelihood of something to be true but not sharing it. As such, the impact of political knowledge was more potent and consistent in increasing truth discernment. Preference for intuitive thinking robustly predicted worse truth discernment, while CRT led to better truth detection but not sharing. People with at least undergraduate education detected fake but not true news better, which resulted in much weaker effects of education not present across all of our robustness checks.

In their turn, identity-driven variables targeted at increasing motivation to elaborate offer promising insights. Preference towards Republicans yields consistent distrust in true news and increases trust in fake news. Next, an individual's political concordance increases overall sharing intention and belief in the news more strongly when it is fake. The interaction effect with fake news, however, does not propagate to sharing intentions. Besides, concordance of the article bias with the voting preferences of the county where the individual resides also increases belief in true news and increases skepticism of fake news. Similar to an individual's concordance, it does not amplify the intention to share fake news. Lastly, we find no effect of living in highly politically polarized areas on either truth or sharing discernment - although this could have been influenced by self-selection and self-misrepresentation bias (Aguinis et al., 2021).

Overall, the results generally support the notion that as the ability and motivation to process information increases, people can engage in more effortful thinking, resulting in better truth and news sharing discernment. Apart from differences in the effect of an individual's political concordance and the level of efficacy of the treatment, our findings support prior research. Importantly, factors influencing engagement in higher-order thinking seem to boost the belief in true news, while predictors of "lazy thinking" diminish trust in truth. Notably, the results demonstrate the applicability of the Elaboration Likelihood Model in the context of fake news in an experimental setting, enabling others to leverage the predictions of ELM to study interventions further mitigating the impact of fake news.

6.2 The Limited Effect of the Treatment

Similar to Roozenbeek et al. (2021), we failed to replicate the impact of the accuracy prompt previously reported by Pennycook and Rand (2021, 2020, in press). Since meta-analytic evidence by Pennycook & Rand (in press) seems to point towards a weaker than originally anticipated effect, we discuss a few reasons why this could have been the case.

Firstly, a drawback of this line of research on fake news is the need to use novel headlines that have a lower propensity to feel familiar to the respondents. We carefully selected the headlines from the set by Pennycook and Rand (2021) and followed their guidance. We ensured virtually equal average political bias of the articles and took additional steps to have variable complexity of the items to extract more information per item. It is, nonetheless, possible that individual headlines had a disproportionate impact on the study results (Roozenbeek et al., 2021).

Secondly, the subject recruitment platforms were different. Prolific, the crowdworking platform we have used, was found to have the highest data quality across multiple platforms (Peer et al., 2021), especially when compared to MTurk. 18 out of 20 of the studies that Pennycook and Rand (in press) inspect in their meta-analysis have been conducted on MTurk and Lucid. Both are notorious for high levels of inattention among participants (Aguinis et al., 2021; Aronow et al., 2020; Barends & de Vries, 2019). To illustrate, 90.7% of participants in our sample passed all attention checks. In comparison, the average success rate for the same check used in the study was 25% among Lucid workers. New MTurk workers seem to also do worse on the CRT than Prolific workers (Arechar & Rand, 2021). Thus, our samples' baseline attention and cognitive reflection levels could drive the insensitivity to the treatment, as more attentive participants are less responsive to accuracy prompts (Roozenbeek et al., 2021).

Thirdly, the treatment might not have worked because it did not elicit the more effortful response. ELM might help shed light on this particular issue. In a post-hoc analysis, we find that showing participants the accuracy prompt did not result in a statistically meaningful change in the *position* variable - the self-reported negative or positive stance on the topic. The variable can also be used as a proxy of favorable or unfavorable thoughts dominating when exposed to the news article. Since the treatment did not cause a noticeable change in the predominating favorable or unfavorable thoughts measured by the *position* proxy, the cognitive structure change driven by the accuracy prompt should be negligible. In effect, no attitude change would occur (i.e., opting to downrank or not to share fake news; Donohew et al., 2015; Petty & Cacioppo, 1986).

Likewise, the treatment did not seem to encourage more prolonged deliberation, as measured by two proxies - the time from the display of the question until the first and last click on it. On the one hand, shorter first click time may be interpreted as a preference towards providing intuitive responses. On the other - a longer last click time can indicate increased deliberation (a similar proxy for deliberation has been used by Moravec et al., 2020). However, as the table in Appendix XVI show, that is not the case. Since showing an accuracy prompt in the form of a label did not induce additional deliberation, this could be a reason why it failed to elicit the attitude changes as hypothesized.

In all, item- and treatment-level issues and stark contrast in the samples' preferences for accuracy may have influenced the difference in the results. Ultimately, although we did not find the treatment to work as expected, it seems effective in increasing trust in true news. In a media landscape where increasing trust in media, especially the "mainstream" ones, is a salient problem (Gottfried, 2021; Jurkowitz et al., 2020), this is an encouraging result.

6.3 Theoretical Implications of the Study

To finalize the thesis, this section puts the study results into a broader theoretical and practical perspective. It aims to integrate the results within the theoretical frameworks and empirical findings from the literature. With some exceptions discussed further in this subsection, our study results otherwise strongly support the dual-process account of fake news influence.

Bad concordance, good concordance

To the author's best knowledge, one of the novel contributions of this study to the existing literature was examining the effect of political polarization and voter preferences of respondent's place of residence. It allowed us to discern between the individual and group-level identity and their impact on truth discernment.

Consistent with prior research on the subject (e.g., Pennycook & Rand, 2021), both types of political concordance inflated the overall support for true news. In the context of ELM, both individual and county-level concordance may act through increasing personal relevance, the ideologically consistent prior factual knowledge. As a result, people think more deliberately and are subsequently more confident in their judgments of true news.

When it comes to discerning fake news, however, the results are mixed. County-level concordance of the news article worked as predicted. However, regarding individual-level concordance, our predictions were in agreement with the motivated thinking literature (e.g., Beck, 2017; Calvert, 2017; Kahan, 2017) rather than dual-process theory (for overview, Pennycook & Rand, 2021b). A speculative explanation for this difference might be due to the different mechanisms at play regarding ideological concordance. First, the county-level concordance may also work through the "personal responsibility" component of the ELM. Being shown a political news article that is concordant with the voting preferences of the respondent's county of residence could make their social group identity ("I am a resident of a Democrat-leaning county") more salient (Oyserman & Dawson, 2020). Responsibility towards the group, in its turn, might have helped the analytical system to supersede the possible motivated thinking.

Conversely, there was no such "override" in the case of an individual's concordance. On the bright side, this individual-level support does not seem to translate into spreading the information (both fake and true) further. Instead, it might be negatively associated with it, which could be driven again by the responsibility towards the larger group.

Lazy or politically biased brain?

The present results are relevant for the theoretical debate about the mechanism of influence of fake news. Two contrasting accounts - the politically motivated reasoning and the dual-process framework, positing that people are generally not biased but rather engage in "lazy thinking" (Pennycook & Rand, 2019). Our findings speak strongly in support of the latter. Consistently with the Elaboration Likelihood Model, we find that people that can engage more in deep thinking make better choices. Rather than driven by heightened overall skepticism, they seem to be better at detecting lies (as is the case, for instance, for CRT performance and political knowledge). Conversely, people who rely on heuristics and intuition share and believe in fake news more.

Nonetheless, to the best of our knowledge, studies that generally criticize the politically motivated reasoning mechanism (see Pennycook & Rand, 2021b) do not consider the motivational aspects of confirmation bias as a mechanism. We have shown that the effect of confirmation bias is consistently strong in our results. Together with the decrease in truth discernment due to political concordance, it highlights that more research needs to be conducted to untangle the motivational and cognitive aspects of fake news influence.

Partisanship and fake news

Although this was not a primary goal of our study, we find that the overall capacity to tell fake news from true news differs based on political ideology. Moreover, similar to other studies (e.g., Pennycook & Rand, 2019), the association between conservatism and media truth discernment is held independently of CRT performance. This finding is relevant for the ongoing debate on the ideological asymmetries in the way people process information (Ditto et al., 2018; Jost, 2017). This, however, sheds light on why Republican-consistent fake news was more common than Democrat-consistent fake news during the election period (Guess & Coppock, 2018; Allcott & Gentzkow, 2017), or why the US media ecosystem is more polarized on the political right than on the left (Faris et al., 2017). One potential explanation could be that Republicans tend to be more cognitively rigid than liberals (Jost, 2017). In addition, Republicans distrust the media more, especially if they classify it as "mainstream" (Gottfried, 2021; Jurkowitz et al., 2020). Together, these facts corroborate our findings that right-leaning respondents doubted the accuracy of true news and overestimated the accuracy of fake news.

6.4 Practical Implications of the Study

Importantly, our study contains several practical implications thanks to the practical relevance of the Elaboration Likelihood Model. This section will review broad recommendations based on our results, contrast them to the strategies pursued by some social media companies, and evaluate their possible efficacy based on our results.

Firstly, in tune with prior research (e.g., Pennycook et al., 2021), our study showed that while people generally are good at discerning fake news from true news, their sharing discernment is not as good. Considering our accuracy prompt's minor effect was only marginally successful with true news, complex strategies targeting sharing and truth discernment must be deployed.

An encouraging result was obtained concerning political knowledge. Fostering better political education among the population might thus be an excellent way to prevent misinformation. Both Facebook (Hatmaker, 2020) and Twitter (Coyne & Toizer, 2020) launched voter support hubs on their platforms, offering specific information regarding the elections. Similarly, YouTube showed voting information panels on searches related to 2020 elections and candidates (Hatmaker, 2020). Both Twitter and Facebook expanded to fight non-political misinformation, such as COVID-19 misinformation, similarly. In all, these are encouraging developments that need to be continued.

Furthermore, one of the predictions of ELM is that prior factual knowledge increases the probability of processing information effortfully. Apart from political knowledge, showing related articles for context enables better truth discernment (Alemanno, 2018). More information on the topic and better argumentation of why the topic is rated as false could also be effective according to ELM thanks to the better influence of higher-quality arguments (Petty & Cacioppo, 1986). High-quality evidence persuades more when it comes to true and fake news (Martire et al., 2020). Similarly, people seem to welcome more context provided to the news articles (Kirchner & Reuter, 2020). Thus adding brief facts from verified sources might help (Facebook and Twitter, for instance, started providing links to Encyclopedia Britannica or Wikipedia; Rosen et al., 2020).

Finally, platforms can infer users' political ideology and promote plurality and increased trust in true articles by highlighting they came from verified resources and encourage critical thinking when the content comes from untrustworthy resources, especially when they are politically concordant.

6.5 Limitations of the Study and Further Research

Like all studies, ours is not without limitations. Several obvious limitations come due to the sample. Firstly, our sample's generalizability is limited since the current work was not conducted on a nationally representative sample, balanced according to personal characteristics such as age, gender, income, and obtained with probability sampling. However, the goal of our study is not to make inferences regarding those variables but rather the underlying mechanism of influence of fake news. Furthermore, it has to be noted that most studies on this topic, including the current one, have been conducted on US samples - which might have implications for generalizability worldwide (Cheon et al., 2020).

Another significant limitation is the experimental setting. Although our study follows a paradigmatic design (see Pennycook et al., 2021 for review), our sharing intentions were only hypothetical. Although we have done our best to replicate the design of a Facebook warning label, our accuracy manipulation was still performed in the "lab" context, reducing the external validity of the results. After all, people consume social media for pleasure (Moravec et al., 2020) - thus, they might be in a different, more alert state of mind when asked to make a sharing judgment within a survey form. Individuals in such a hedonic state of mind are less likely to critically consider the information they see (Moravec et al., 2020; Kahneman, 2003). If anything, however, this would underestimate our study results. In addition, there is evidence that self-reported sharing intentions significantly correlate with actual social media engagement (Mosleh et al., 2021). Lastly, we have been using virtually the same methods as past studies showing external validity (survey results replicated in subsequent behavior on Twitter; Pennycook et al., 2020).

Further research should expand beyond the traditional US-based samples to explore cross-cultural generalizability. In addition, other content (e.g., climate change or vaccination-related fake news) should be tested to verify domain-wide generalizability. Furthermore, given that the accuracy prompt we have used did not have the intended effect, other studies may focus on related warning labels. Hence scrutinizing the impact of the treatments that influence both the central and peripheral routes (e.g., Moravec et al., 2020) might be a valuable addition to the literature. Moreover, the time decay of the treatments should be studied more closely due to emerging evidence that even the most successful interventions decay rapidly with time (Maertens et al., 2021).

Moreover, better experimental paradigms, simulating the environment of social media (e.g., Facebook interface), or conducting real-life interventions targeted on social media users that have previously shared misinformation, are needed to design intervention strategies that work. Another exciting avenue for research could be disentangling the influence of "more obvious" versus "less obvious" fake news and whether the effect of interventions varies with regards to those, given the emerging evidence consistent with ELM that people are more persuaded by higher-quality than lower-quality evidence (Martire et al., 2020). Ultimately, there is a growing consensus that dual-process theories should move away from solely discerning between two different types of processes toward models where analytic and intuitive thinking interact together (De Neys, 2017; Thompson, 2009). Hence, researchers should consider interventions that may affect both processing routes for a potentially increased impact (as suggested by early evidence, e.g., Moravec et al., 2020).

Section 7 - Conclusion

On an uneventful morning in December 2017, at 11:06 a.m, a fake news article surfaced online stating that National Security Adviser Michael Flynn would testify that Donald Trump was involved with Russian officials. By 11:34 a.m., the bombshell article signaling an imminent political crisis went viral, causing a sharp drop by 38 points in the S&P 500, equivalent to a $341 billion loss (Cavazos, 2019). This example illustrates that although it costs nothing to spread disinformation online, fake news does have real socioeconomic consequences. Therefore, it is crucial for researchers from various behavioral and social sciences, including economics, to develop a clear understanding of how fake news influences people and how to fight it. With this goal, the following pre-registered study combined an interdisciplinary approach to contribute to the literature on fake news in multiple ways.

First and foremost, applying the dual processing perspective, we demonstrated that fast and intuitive cognition antecedents promote belief in fake news and distrust in true news. Conversely, predictors of more deliberate thinking mitigate the impact of false content and increase trust in truth. Additionally, in support of the "lazy processing perspective" (Pennycook & Rand, in press), we find that the veracity of news articles does not impact sharing intentions, despite having a strong influence on article believability. In sum, these findings have implications for social media. These platforms are designed for rapid consumption of a mix of serious news and emotionally engaging content and provide instant gratification through social feedback on anything shared (Pennycook & Rand, 2019a). Hence, building upon the theoretical framework that we have validated with this research, one of the strategies to fight fake news could be interventions encouraging effortful thinking.

One possible intervention could be using accuracy prompts in the form of warning labels. To the best of the author's knowledge, the present research was the first study to shed light on how accuracy prompts might act in the form of warning labels. While the study did not confirm that accuracy prompts decrease belief in fake news, it partially substantiated that they could lead to better truth discernment. The approach used in the study may be used to promote trust in true news, especially in groups subject to overall skepticism in mainstream media, such as Republican-leaning individuals.

Finally, the current study advances the debate on two competing accounts - politically motivated reasoning and failing to think effortfully by demonstrating that other variables, such as confirmation bias or voter preferences in the individual's county of residence, may need to be considered while contrasting the two approaches.

In all, scalable interventions combating cognitive laziness could potentially increase the overall quality of the news shared online without the need to rely on a truth arbiter to certify truth and censor the fake news. Adapting to the tectonic shift in news consumption and sharing based on the dual-process theory may help us “slower” the thinking and make it more accurate.

Bibliography

Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk Research: Review and Recommendations. *Journal of Management*, *47*(4), 823–837. https://doi.org/10.1177/0149206320969787

Albarracín, D. (2020). Conspiracy Beliefs: Knowledge, ego defense, and social integration in the processing of fake news. In *The Psychology of Fake News* (pp. 196-219). Routledge.

Alemanno, A. (2018). How to Counter Fake News? A Taxonomy of Anti-fake News Approaches. *European Journal of Risk Regulation*, *9*(1), 1–5. https://doi.org/10.1017/err.2018.12

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Allen, J. N. L., Arechar, A. A., Pennycook, G., & Rand, D. (2020). *Scaling Up Fact-Checking Using the Wisdom of Crowds*. PsyArXiv. https://doi.org/10.31234/osf.io/9qdza

Amazeen, M. A., Thorson, E., Muddiman, A., & Graves, L. (2018). Correcting Political and Consumer Misperceptions: The Effectiveness and Effects of Rating Scale Versus Contextual Correction Formats. *Journalism & Mass Communication Quarterly*, *95*(1), 28–48. https://doi.org/10.1177/1077699016678186

An, J., Quercia, D., & Crowcroft, J. (2014). Partisan sharing: Facebook evidence and societal consequences. *Proceedings of the Second ACM Conference on Online Social Networks*, 13–24. https://doi.org/10.1145/2660460.2660469

Anderson, C. W. (2021). Fake News is Not a Virus: On Platforms and Their Effects. *Communication Theory*, *31*(1), 42–61. https://doi.org/10.1093/ct/qtaa008

Andrews, E. (2015). The Great Moon Hoax, *The History Channel*. Accessed 01/04/18: http://www.history.com/news/the-great-moon-hoax-180-years-ago?linkId=16545579

Arechar, A. A., & Rand, D. G. (2021). Turking in the time of COVID. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01588-4

Aronow, P. M., Kalla, J., Orr, L., & Ternovski, J. (2020). Evidence of rising rates of inattentiveness on Lucid in 2020. SocArXiv. https://doi.org/10.31235/osf.io/8sbe4 Barends, A. J., & de Vries, R. E. (2019).

Athey, S., Mobius, M. M., & Pál, J. (2017). The impact of aggregators on internet news consumption. Stanford University Graduate School of Business Research Paper Series, 17(8)

Austin, E. W., & Dong, Q. (1994). Source v. content effects on judgments of news believability. *Journalism quarterly*, 71(4), 973-983.

Avaaz. (2020). *Facebook's Algorithm: A Major Threat to Public Health*. Avaaz. https://secure.avaaz.org/campaign/en/facebook_threat_health/

Avram, M., Micallef, N., Patil, S., & Menczer, F. (2020). Exposure to Social Engagement Metrics Increases Vulnerability to Misinformation. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-033

Axt, J. R., Landau, M. J., & Kay, A. C. (2020). Fake news attributions as a source of nonspecific structure. *–Stephan Lewandowsky, Cognitive Science*, 220.

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90-109.

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, *149*(8), 1608–1613. https://doi.org/10.1037/xge0000729

Barends, A. J., & de Vries, R. E. (2019). Noncompliant responding: Comparing exclusion criteria in MTurk personality research to improve data quality. Personality and Individual Differences, 143, 84–89. https://doi.org/j .paid.2019.02.015

Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P., & Linden, S. V. D. (2021). Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. Big Data & Society, 8(1), 20539517211013868.

Bastick, Z. (2021). Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. *Computers in Human Behavior*, *116*, 106633. https://doi.org/10.1016/j.chb.2020.106633

Beck, J. (2017). This article won't change your mind: The fact on why facts alone can't fight false beliefs. The Atlantic Retrieved from https://www.theatlantic.com/science/ archive/2017/03/this-article-wont-change-your-mind/519093/.

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention to self-administered surveys. *American Journal of Political Science*, 58(3), 739-753.

Berinsky, A. J. (2017). Rumors and Health Care Reform: Experiments in Political Misinformation. *British Journal of Political Science*, *47*(2), 241–262. https://doi.org/10.1017/S0007123415000186

Bradshaw, S., & Howard, P. N. (n.d.). *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation*. 26.

Brashier, N. M., Pennycook, G., Berinsky, A. J., & Rand, D. G. (2021). Timing matters when correcting fake news. Proceedings of the National Academy of Sciences, 118(5).

Bremner, C. (2018). France aims to ban fake news at election time. The times.

Britt, M. A., Rouet, J.-F., Blaum, D., & Millis, K. (2019). A Reasoned Approach to Dealing With Fake News. *Policy Insights from the Behavioral and Brain Sciences*, *6*(1), 94–101. https://doi.org/10.1177/2372732218814855

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in Fake News is Associated with Delusionality, Dogmatism, Religious Fundamentalism, and Reduced Analytic Thinking. *Journal of Applied Research in Memory and Cognition*, *8*(1), 108–117. https://doi.org/10.1016/j.jarmac.2018.09.005

Brown, V. A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920960351. https://doi.org/10.1177/2515245920960351

Bryanov, K., & Vziatysheva, V. (2021). Determinants of individuals' belief in fake news: A scoping review determinants of belief in fake news. *PLOS ONE*, *16*(6), e0253717. https://doi.org/10.1371/journal.pone.0253717

Bryant, D. (2020). *Dopaminergic Genes Predictive of Unmotivated Confirmation Bias is not Predictive of Motivated Confirmation Bias*. https://shareok.org/handle/11244/325377

Burnham, K. P., & Anderson, D. R. (2004). *Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research,* 33(2), 261-304.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment*, *48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

Calfas, J. (2017). Google is changing its search algorithm to combat fake news.

Calvillo, D. P., Ross, B. J., Garcia, R. J., Smelter, T. J., & Rutchick, A. M. (2020). Political ideology predicts perceptions of the threat of COVID-19 (and susceptibility to fake news about it). Social Psychological and Personality Science, 11(8), 1119-1128.

Carrieri, V., Madio, L., & Principe, F. (2019). Vaccine hesitancy and (fake) news: Quasi-experimental evidence from Italy. *Health Economics*, *28*(11), 1377–1382. https://doi.org/10.1002/hec.3937

Cassidy, J. (2018). *Cambridge Analytica Whistleblower claims that cheating swung the Brexit vote.* The New Yorker. Accessed, 29, 03-18.

Cavazos, R., (2019). The economic cost of bad actors on the internet. *CHEQ*.

Channel 4. (2017). C4 study reveals only 4% surveyed can identify true or fake news. Retrieved August 29, 2017, from www.channel4.com/info/press/news/c4-studyreveals-only-4-surveyed-can-identify-true-or-fake-news

Chen, C. X., Pennycook, G., & Rand, D. (2021). *What Makes News Sharable on Social Media?* PsyArXiv. https://doi.org/10.31234/osf.io/gzqcd

Chen, C. Y., Kearney, M., & Chang, S. L. (2021). Comparative Approaches to Mis/Disinformation| Belief in or Identification of False News According to the Elaboration Likelihood Model. *International Journal of Communication,* 15, 23.

Cheon, B. K., Melani, I., & Hong, Y. Y. (2020). How USA-centric is psychology? An archival study of implicit assumptions of generalizability of findings to human nature based on origins of study samples. Social Psychological and Personality Science, 11(7), 928-937.

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, *1*(3), 98-101.

Corneille, O., Mierop, A., & Unkelbach, C. (2020). Repetition increases both the perceived truth and fakeness of information: An ecological account. *Cognition*, *205*, 104470. https://doi.org/10.1016/j.cognition.2020.104470

Coronavirus: Staying safe and informed on Twitter. (n.d.). Retrieved July 26, 2021, from https://blog.twitter.com/en_us/topics/company/2020/covid-19

Coyne, B., & Toizer, S. (2020). *Helping you find accurate US Election News and Information*. https://blog.twitter.com/en_us/topics/company/2020/2020-election-news

Craney, T. A., & Surles, J. G. (2002). Model-Dependent Variance Inflation Factor Cutoff Values. *Quality Engineering*, *14*(3), 391–403. https://doi.org/10.1081/QEN-120001878

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The Truth About the Truth: A Meta-Analytic Review of the Truth Effect. *Personality and Social Psychology Review*, *14*(2), 238–257. https://doi.org/10.1177/1088868309352251

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A, Caldarelli, G., Stanley, H. E., & Quattrociocci, W. (2016). The spreading of misinformation online. Proceedings of the National Academy of Sciences 113 (3), 554-559

Del Vicario, M., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2017). Modeling confirmation bias and polarization. Scientific Reports 7, 40391.

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on Facebook. *Scientific reports*, *6*, 37825.

De Neys, W. (Ed.). (2017). *Dual-process theory 2.0*. Routledge.

Dias, N., Pennycook, G., & Rand, D. G. (2020). Emphasizing publishers does not effectively reduce susceptibility to misinformation on social media. Harvard Kennedy School Misinformation Review, 1(1).

Dias, N., Pennycook, G., & Rand, D. G. (2020). Emphasizing publishers does not effectively reduce susceptibility to misinformation on social media*. Harvard Kennedy School Misinformation Review,* 1(1).

Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., Celniker, J. B., & Zinger, J. F. (2019). At Least Bias Is Bipartisan: A Meta-Analytic Comparison of Partisan Bias in Liberals and Conservatives. *Perspectives on Psychological Science*, *14*(2), 273–291. https://doi.org/10.1177/1745691617746796

Donohew, L., Sypher, H. E., & Higgins, E. T. (Eds.). (2015). *Communication, Social Cognition, and Affect (PLE: Emotion)*. Psychology Press. https://doi.org/10.4324/9781315743974

Douthat, R. (2015). For the South, against the Confederacy: On the distinctive appeal of a “backwards” region. *New York Times, June 24. douthat. blogs. nytimes. com/2015/06/24/for-the-south-against-the-confederacy.*

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, *9*(2), 111–119. https://doi.org/10.1038/s41558-018-0360-1

Duffy, A., and Ling, R. (2020). The gift of news: A shift from information to social in studies of news sharing. Journalism Studies 21(1), 72-87.

Epstein, Z., Berinsky, A., Cole, R., Gully, A., Pennycook, G., & Rand, D. (2021). *Developing an accuracy-prompt toolkit to reduce COVID-19 misinformation online.* PsyArXiv. https://doi.org/10.31234/osf.io/sjfbn

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, 8(3), 223-241. https://journals.sagepub.com/doi/abs/10.1177/1745691612460685

Facebook (2019). Helping to Protect the 2020 US Elections. *About Facebook*. https://about.fb.com/news/2019/10/update-on-election-integrity-efforts/

Facebook and Instagram will pin vote-by-mail explainers to the top of feeds. (n.d.). *TechCrunch*. Retrieved July 26, 2021, from https://social.techcrunch.com/2020/10/09/facebook-instagram-vote-by-mail-video-notifications/

Facebook Drops Fake News Warning Flags. (n.d.). Time. Retrieved July 21, 2021, from https://time.com/5077002/facebook-fake-news-articles/

Facts about the Measles Outbreak. (2015, February 2). The New York Times. Retrieved from www.nytimes.com/interactive/2015/02/02/us/measles-facts.html

Faris, R. M., Roberts, H., Etling, B., Bourassa, N., Zuckerman, E., & Benkler, Y. (2017). Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election. Berkman Klein Center for Internet & Society Research Paper. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id= 3019414

Fazio, L. (2020). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harvard Kennedy School Misinformation Review, 1*(2). https://doi.org/10.37016/mr-2020-009

Fazio, L. K., Rand, D. G., & Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review*, *26*(5), 1705–1710. https://doi.org/10.3758/s13423-019-01651-4

Ferrara, E. (2020). # covid-19 on Twitter: Bots, conspiracies, and social media activism. *arXiv preprint arXiv: 2004.09531.*

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Freelon, D. (2017). Personalized information environments and their potential consequences for disinformation. Understanding and addressing the disinformation ecosystem. Annenberg School for Communication, 38-44.

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social clicks: What and who gets read on Twitter?. ACM SIGMETRICS Performance Evaluation Review, 44(1), 179-192.

Gallacher, J. D., Heerdink, M. W., & Hewstone, M. (2021). Online Engagement Between Opposing Political Protest Groups via Social Media is Linked to Physical Violence of Offline Encounters. *Social Media + Society*, *7*(1), 2056305120984445. https://doi.org/10.1177/2056305120984445

Garrett, R. K. (2017). The “echo chamber” distraction: Disinformation campaigns are the problem, not audience fragmentation. Journal of Applied Research in Memory and Cognition 6(4), 370-376.

Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of personality and social psychology*, *59*(4), 601.

Goldman, R. (2016, December 24). Reading Fake News, Pakistani Minister Directs Nuclear Threat at Israel. *The New York Times*. https://www.nytimes.com/2016/12/24/world/asia/pakistan-israel-khawaja-asif-fake-news-nuclear.html

Gordon, M., & Volz, D. (202 C.E.). Russian Disinformation Campaign Aims to Undermine Confidence in Pfizer, Other Covid-19 Vaccines, U.S. Officials Say—WSJ. https://www.wsj.com/articles/russian-disinformation-campaign-aims-to-undermine-confidence-in-pfizer-other-covid-19-vaccines-u-s-officials-say-11615129200

Graeupner, D., & Coman, A. (2017). The dark side of meaning-making: How social exclusion leads to superstitious thinking. *Journal of Experimental Social Psychology*, *69*, 218–222. https://doi.org/10.1016/j.jesp.2016.10.003

Greifeneder, R., Jaffe, M., Newman, E., & Schwarz, N. (2021). *The psychology of fake news: Accepting, sharing, and correcting misinformation* (p. 252).

Guess, A., & Coppock, A. (2018). Does Counter-Attitudinal Information Cause Backlash? Results from Three Large Survey Experiments. *British Journal of Political Science*, *50*(4), 1497–1515. https://doi.org/10.1017/S0007123418000327

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

Hatmaker, T. (2020). Facebook wants content reviewers back ASAP, slows return plan for most employees.

Heller, A. (2018, October 31). Heller: How Orson Welles' 'War of The Worlds' broadcast influenced fake news and mass panic. *The Whit Online*. https://thewhitonline.com/2018/10/opinion/heller-how-orson-welles-war-of-the-worlds-broadcast-influenced-fake-news-and-mass-panic/

Horne, B., & Adali, S. (2017). This Just In: Fake News Packs A Lot In Title, Uses Simpler, Repetitive Content in Text Body, More Similar To Satire Than Real News. *Proceedings of the International AAAI Conference on Web and Social Media*, *11*(1), Article 1. https://ojs.aaai.org/index.php/ICWSM/article/view/14976

Howard, P. N., Kollanyi, B., Bradshaw, S., & Neudert, L.-M. (2018). Social Media, News and Political Information during the US Election: Was Polarizing Content Concentrated in Swing States? *ArXiv:1802.03573 [Cs]*. http://arxiv.org/abs/1802.03573

How Internet Platforms Are Combating Disinformation and Misinformation in the Age of COVID-19. (n.d.). New America. Retrieved July 26, 2021, from http://newamerica.org/oti/reports/how-internet-platforms-are-combating-disinformation-and-misinformation-age-covid-19/

Huber, C. R., & Kuncel, N. R. (2016). Does College Teach Critical Thinking? A Meta-Analysis. *Review of Educational Research*, *86*(2), 431–468. https://doi.org/10.3102/0034654315605917

Ioannidis, J. P. A. (2020). Coronavirus disease 2019: The harms of exaggerated information and non-evidence-based measures. *European Journal of Clinical Investigation*, *50*(4), e13222. https://doi.org/10.1111/eci.13222

Jaffé, M. E., & Greifeneder, R. (2020). Can that be True or is it Just Fake News?: New perspectives on the negativity bias in judgments of truth. *In The Psychology of Fake News* (pp. 115-130). Routledge.

Jakesch, M., Koren, M., Evtushenko, A., & Naaman, M. (2019). The role of source and expressive responding in political news evaluation. In *Computation and Journalism Symposium.*

Janze, C., & Risius, M. (2017). *Automatic Detection of Fake News on Social Media Platforms*. 16.

Jost, J. T. (2017). Ideological asymmetries and the essence of political psychology. Political Psychology, 38(2), 167–208. http://dx.doi.org/10.1111/pops.12407.

Jost, P. J., Pünder, J., & Schulze-Lohoff, I. (2020). Fake news—Does perception matter more than the truth? *Journal of Behavioral and Experimental Economics*, *85*, 101513. https://doi.org/10.1016/j.socec.2020.101513

Jurkowitz, M., Mitchell, A., Shearer, E., & Walker, M. (2020, January 24). 1. Democrats report much higher levels of trust in a number of news sources than Republicans. *Pew Research Center's Journalism Project*. https://www.journalism.org/2020/01/24/democrats-report-much-higher-levels-of-trust-in-a-number-of-news-sources-than-republicans/

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature climate change,* 2(10), 732-735.

Kahan, D. M. (2012). Ideology, motivated reasoning, and cognitive reflection: *An experimental study. Judgment and Decision making*, 8, 407-24.

Kahan, D. M. (2017). *Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition* (SSRN Scholarly Paper ID 2973067). Social Science Research Network. https://doi.org/10.2139/ssrn.2973067

Kahneman, D., Lovallo, D., & Sibony, O. (2011). *Before You Make That Big Decision*. http://ir.vnulib.edu.vn/handle/123456789/4136

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, 93(5), 1449-1475.

Kaminska, I. (2017). A lesson in fake news from the info-wars of ancient Rome. *Financial Times*.

Karami, A., Dahl, A. A., Shaw, G., Valappil, S. P., Turner-McGrievy, G., Kharrazi, H., & Bozorgi, P. (2021). Analysis of Social Media Discussions on (#)Diet by Blue, Red, and Swing States in the U.S. *Healthcare*, *9*(5), 518. https://doi.org/10.3390/healthcare9050518

Kim, A., Moravec, P. L., & Dennis, A. R. (2019). Combating fake news on social media with source ratings: The effects of user and expert reputation ratings. *Journal of Management Information Systems*, *36*(3), 931-968.

Kim, H. B., Choi, S., Kim, B., & Pop-Eleches, C. (2018). The role of education interventions in improving economic rationality. *Science*, *362*(6410), 83–86. https://doi.org/10.1126/science.aar6987

Kirchner, J., & Reuter, C. (2020). Countering Fake News: A Comparison of Possible Solutions Regarding User Acceptance and Effectiveness. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 140:1-140:27. https://doi.org/10.1145/3415211

Kitchen, P., Kerr, G., Schultz, D., Mccoll, R., & Pals, H. (2014). The elaboration likelihood model: Review, critique and research agenda. *European Journal of Marketing*, *48*, 2033–2050. https://doi.org/10.1108/EJM-12-2011-0776

Klayman, J. (1995). Varieties of Confirmation Bias. In J. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Psychology of Learning and Motivation* (Vol. 32, pp. 385–418). Academic Press. https://doi.org/10.1016/S0079-7421(08)60315-1

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences*, *110*(15), 5802-5805.

Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E. W., & Baddour, K. (2020). Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter. *Cureus*, *12*(3), e7255. https://doi.org/10.7759/cureus.7255

Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools. *Psychological Science in the Public Interest*, *21*(3), 103–156. https://doi.org/10.1177/1529100620946707

Kurtzleben, D. (2016). Do fact checks matter? Retrieved August 8, 2018, from www.npr.org/2016/09/27/495233627/do-fact-checks-matter

Lab, M. E. D. and S. (2021). *County Presidential Election Returns 2000-2020* [Data set]. Harvard Dataverse. https://doi.org/10.7910/DVN/VOQCHQ

Lane, D. S., Kim, D. H., Lee, S. S., Weeks, B. E., & Kwak, N. (2017). From Online Disagreement to Offline Action: How Diverse Motivations for Using Social Media Can Increase Political Information Sharing and Catalyze Offline Political Participation. *Social Media + Society*, *3*(3), 2056305117716274. https://doi.org/10.1177/2056305117716274

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Leetaru, K. (2017). The backfire effect and why Facebook's' Fake News' warning gets it all wrong. *Forbes.*

Levi, L. (2017). Real fake news and fake fake news. *First Amend. L. Rev.,* 16, 232.

Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of applied research in memory and cognition*, *6*(4), 353-369.

Lex, E., Wagner, M., & Kowald, D. (2018). Mitigating Confirmation Bias on Twitter by Recommending Opposing Views. *ArXiv:1809.03901 [Cs]*. http://arxiv.org/abs/1809.03901

Li, H., & Sakamoto, Y. (2014). Social impacts in social media: An examination of perceived truthfulness and sharing of information. *Computers in Human Behavior*, 41, 278-287.

Lim, C. (2018). Checking how fact-checkers check. Research & Politics, 5(3), 2053168018786848.

Linden, S. van der, Leiserowitz, A., Rosenthal, S., & Maibach, E. (2017). Inoculating the Public against Misinformation about Climate Change. *Global Challenges*, *1*(2), 1600008. https://doi.org/10.1002/gch2.201600008

Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C. R., & Hertwig, R. (2020). How behavioural sciences can promote truth, autonomy and democratic discourse online. *Nature Human Behaviour*, *4*(11), 1102–1109. https://doi.org/10.1038/s41562-020-0889-7

Luo, M., Hancock, J. T., & Markowitz, D. M. (2020). Credibility Perceptions and Detection Accuracy of Fake News Headlines on Social Media: Effects of Truth-Bias and Endorsement Cues. *Communication Research*, 0093650220921321. https://doi.org/10.1177/0093650220921321

Maertens, R., Roozenbeek, J., Basol, M., & van der Linden, S. (2020). Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments. *Journal of Experimental Psychology: Applied.*

Marsh, E. J., & Stanley, M. L. (2020). False Beliefs: Byproducts of an adaptive knowledge base?. In *The Psychology of Fake News* (pp. 131-146). Routledge.

Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, *5*(1), 47. https://doi.org/10.1186/s41235-020-00252-3

Martire, K. A., Growns, B., Bali, A. S., Montgomery-Farrer, B., Summersby, S., & Younan, M. (2020). Limited not lazy: A quasi-experimental secondary analysis of evidence quality evaluations by those who hold implausible beliefs. *Cognitive Research: Principles and Implications*, *5*(1), 65. https://doi.org/10.1186/s41235-020-00264-z

Massie, A. (2015, May 9). Reflections on the revolution in Scotland. The Spectator. Retrieved September 9, 2019, from www.spectator.co.uk/2015/05/this-election-hasbeen-

Masta, K. E., and E. Shearer. 2018. "News Use Across Social Media Platforms 2018." Pew Research Center, Journalism & Media.

Mattiuzzi, C., & Lippi, G. (2020). Which lessons shall we learn from the 2019 novel coronavirus outbreak? *Annals of Translational Medicine*, *8*(3), 48. https://doi.org/10.21037/atm.2020.02.06

Meixler, E. (2017). Facebook is dropping its fake news red flag warning after finding it had the opposite effect. *Time*.

Mirsky, S. (2020). *Flat Earthers: What They Believe and Why*. Scientific American. https://www.scientificamerican.com/podcast/episode/flat-earthers-what-they-believe-and-why/

Mitchell, Amy, Jeffrey Gottfried, Michael Barthel, and Elisa Shearer. "The modern news consumer: News attitudes and practices in the digital era." (2016).

MIT Election Data and Science (2020) https://www.nytimes.com/interactive/2020/12/20/us/politics/election-hispanics-asians-voting.html

Moravec, P. L., Kim, A., & Dennis, A. R. (2020). Appealing to Sense and Sensibility: System 1 and System 2 Interventions for Fake News on Social Media. *Information Systems Research*, *31*(3), 987–1006. https://doi.org/10.1287/isre.2020.0927

Mosleh, M., Pennycook, G., Arechar, A. A., & Rand, D. G. (2021). *Cognitive reflection correlates with behavior on Twitter. Nature communications, 12(1), 1-10.* Retrieved July 25, 2021, from https://doi.apa.org/doiLanding?doi=10.1037%2Ft49856-000

Moulding, R., Nix-Carnell, S., Schnabel, A., Nedeljkovic, M., Burnside, E. E., Lentini, A. F., & Mehzabin, N. (2016). Better the devil you know than a world you don't? Intolerance of uncertainty and worldview explanations for belief in conspiracy theories. *Personality and individual differences*, 98, 345-354.

Mustafaraj, E., & Metaxas, P. T. (2017, June). The fake news spreading plague: was it preventable?. In *Proceedings of the 2017 ACM on web science conference* (pp. 235-239). ACM.

Newman, E. J., & Zhang, L. (2020). Truthiness: How non-probative photos shape belief. In *The Psychology of Fake News* (pp. 90-114). Routledge.

New Study by Digital New Deal Finds Engagement with Deceptive Outlets Higher on Facebook Today Than Run-up to 2016 Election. (2020, October 12). The German Marshall Fund of the United States. https://www.gmfus.org/blog/2020/10/12/new-study-digital-new-deal-finds-engagement-deceptive-outlets-higher-facebook-today

Newton, C., Feeney, J., & Pennycook, G. (2021). *The Comprehensive Thinking Styles Questionnaire: A novel measure of intuitive-analytic thinking styles*. PsyArXiv. https://doi.org/10.31234/osf.io/r5wez

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, 2(2), 175-220.

Nyhan, B., & Reifler, J. (2019). The roles of information deficits and identity threat in the prevalence of misperceptions. Journal of Elections, Public Opinion and Parties, 29(2), 222-244.

Osatuyi, B., & Hughes, J. (2018, January 3). *A Tale of Two Internet News Platforms-Real vs. Fake: An Elaboration Likelihood Model Perspective*. https://doi.org/10.24251/HICSS.2018.500

Oyserman, D., & Dawson, A. (2020). Your fake news, our facts: Identity-based motivation shapes what we believe, share

Ozdalga, D. (n.d.). *Fake News and Social Endorsement Cues: Do the Number of Likes, Shares and Comments Influence Engagement with Inaccurate News Articles?* 65.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Paul, K. (2020). Here are all the steps social media made to combat misinformation. *The Guardian*, 2. http://www.theguardian.com/technology/2020/oct/29/here-are-all-the-steps-social-media-made-to-combat-misinformation-will-it-be-enough

Peer, E., Rothschild, D. M., Evernden, Z., Gordon, A., & Damer, E. (2021). *Data Quality of Platforms and Panels for Online Behavioral Research* (SSRN Scholarly Paper ID 3765448). Social Science Research Network. https://doi.org/10.2139/ssrn.3765448

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European review of social psychology*, *1*(1), 33-60.

Pehlivanoglu, D., Lin, T., Deceus, F., Heemskerk, A., Ebner, N. C., & Cahill, B. S. (2020). *The Role of Analytical Reasoning and Source Credibility on the Evaluation of Real and Fake Full-Length News Articles*. PsyArXiv. https://doi.org/10.31234/osf.io/98zwc

Pennycook, G., & Rand, D. G. (2019a). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, *116*(7), 2521–2526. https://doi.org/10.1073/pnas.1806781116

Pennycook, G., & Rand, D. G. (2019b). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Pennycook, G., & Rand, D. (2021a). *Reducing the spread of fake news by shifting attention to accuracy: Meta-analytic evidence of replicability and generalizability*. PsyArXiv. https://doi.org/10.31234/osf.io/v8ruj

Pennycook, G., & Rand, D. G. (2021b). The Psychology of Fake News. *Trends in Cognitive Sciences*, *25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Pennycook, G., Binnendyk, J., Newton, C., & Rand, D. G. (2021). A Practical Guide to Doing Behavioral Research on Fake News and Misinformation. *Collabra: Psychology*, *7*(1). https://doi.org/10.1525/collabra.25293

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865–1880. https://doi.org/10.1037/xge0000465

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590-595.

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science*, *31*(7), 770–780. https://doi.org/10.1177/0956797620939054

Pennycook, G. (2019). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, *88*. https://doi.org/10.1111/jopy.12476

Pereira, A., Harris, E. A., & Bavel, J. J. V. (2018). *Identity concerns drive belief: The impact of partisan identity on the belief and dissemination of true and false news*. PsyArXiv. https://doi.org/10.31234/osf.io/7vc5d

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In R. E. Petty & J. T. Cacioppo (Eds.), *Communication and Persuasion: Central and Peripheral Routes to Attitude Change* (pp. 1–24). Springer. https://doi.org/10.1007/978-1-4612-4964-1_1

Pew Research Center (2020) U.S. Public Now Divided Over Whether To Get COVID-19 Vaccine. *Pew Research Center Science & Society*. https://www.pewresearch.org/science/2020/09/17/u-s-public-now-divided-over-whether-to-get-covid-19-vaccine/

Pew Research Center (2021). Republicans less likely to trust their main news source if they see it as 'mainstream'; Democrats more likely. *Pew Research Center*. Retrieved July 26, 2021, from https://www.pewresearch.org/fact-tank/2021/07/01/republicans-less-likely-to-trust-their-main-news-source-if-they-see-it-as-mainstream-democrats-more-likely/

Plutchik, R. E., & Conte, H. R. (1997). *Circumplex models of personality and emotions* (pp. xi-484). American Psychological Association.

Pornpitakpan, C. (2004). The Persuasiveness of Source Credibility: A Critical Review of Five Decades' Evidence. *Journal of Applied Social Psychology*, *34*(2), 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x

Posetti, J., & Matthews, A. (2018). *A short guide to the history of 'fake news' and disinformation*. 20.

Pratto, F., & John, O. P. (1991). Automatic vigilance: the attention-grabbing power of negative social information. *Journal of personality and social psychology*, *61*(3), 380.

Rapoza, K. (2017). Can 'fake news' impact the stock market?. *Forbes News*. Retrieved November 23, 2020, from https://cheddar.com/media/exclusive-fakenews-is-costing-the- world-billion-a-year

Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, *5*(1), 1-10.

Roozenbeek, J., Freeman, A. L. J., & van der Linden, S. (2021). How Accurate Are Accuracy-Nudge Interventions? A Preregistered Direct Replication of Pennycook et al. (2020). *Psychological Science*, *32*(7), 1169–1178. https://doi.org/10.1177/09567976211024535

Rosen, Guy; Harbath, Katie; Gleicher, Nathaniel; Leathern, Rob (2019). "Helping to protect the 2020 US elections". Facebook newsroom, October 21. https://about.fb.com/news/2019/10/update-on-election-integrity-efforts

Ross, R. M., Rand, D. G., & Pennycook, G. (2021). Beyond" fake news": Analytic thinking and the detection of false and hyperpartisan news headlines. Judgment & Decision Making, 16(2).

Russian Disinformation Campaign Aims to Undermine Confidence in Pfizer, Other Covid-19 Vaccines, U.S. Officials Say—WSJ. (n.d.). Retrieved July 18, 2021, from https://www.wsj.com/articles/russian-disinformation-campaign-aims-to-undermine-confidence-in-pfizer-other-covid-19-vaccines-u-s-officials-say-11615129200

Safadi, H., Li, W., Soleymani, S., Kursuncu, U., & Sheth, A. (2020). *Curtailing Fake News Propagation with Psychographics* (SSRN Scholarly Paper ID 3558236). Social Science Research Network. https://doi.org/10.2139/ssrn.3558236

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social psychology*, 59(2), 229.

Schuster, T., Schuster, R., Shah, D. J., & Barzilay, R. (2019). Are we safe yet? the limitations of distributional features for fake news detection. *arXiv preprint arXiv:1908.09805*.

Schuster, T., Schuster, R., Shah, D. J., & Barzilay, R. (2020). The limitations of stylometry for detecting machine-generated fake news. *Computational Linguistics*, 46(2), 499-510.

Schwartz, A. B. (2015). *Broadcast Hysteria: Orson Welles's War of the Worlds and the Art of Fake News*. Farrar, Straus and Giroux.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36. https://doi.org/10.1145/3137597.3137600

Shu, K., Wang, S., & Liu, H. (2019). Beyond news contents: The role of social context for fake news detection. In *Proceedings of the twelfth ACM international conference on web search and data mining* (pp. 312-320).

Silverman, C., Mac, R., & Dixit, P. (2020). I Have Blood on My Hands: a whistleblower says Facebook ignored global political manipulation. *BuzzFeed.*

Singh, V. K., Ghosh, I., & Sonagara, D. (2021). Detecting fake news stories via multimodal analysis. *Journal of the Association for Information Science and Technology*, *72*(1), 3-17.

Spring, M. (2020). Coronavirus: The human cost of virus misinformation. *BBC News*. Available online at: https://www. bbc. com/news/stories-52731624 (accessed May 27, 2020).

Stefanone, M. A., Vollmer, M., & Covert, J. M. (2019). In News We Trust?: Examining Credibility and Sharing Behaviors of Fake News. *Proceedings of the 10th International Conference on Social Media and Society*, 136–147. https://doi.org/10.1145/3328529.3328554

Swire, B., Berinsky, A. J., Lewandowsky, S., & Ecker, U. K. H. (2017). Processing political misinformation: Comprehending the Trump phenomenon. *Royal Society Open Science*, *4*(3), 160802. https://doi.org/10.1098/rsos.160802

Swire, B., Ecker, U. K. H., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(12), 1948–1961. https://doi.org/10.1037/xlm0000422

Swire-Thompson, B., Ecker, U. K., Lewandowsky, S., & Berinsky, A. J. (2020). They might be a liar but they're my liar: Source evaluation and the prevalence of misinformation. Political Psychology, 41(1), 21-34.

Tappin, B. M., Pennycook, G., & Rand, D. (2018). *Rethinking the link between cognitive sophistication and politically motivated reasoning*. PsyArXiv. https://doi.org/10.31234/osf.io/yuzfj

Tappin, B. M., Pennycook, G., & Rand, D. G. (2020a). Thinking clearly about causal inferences of politically motivated reasoning: Why paradigmatic study designs often undermine causal inference. *Current Opinion in Behavioral Sciences*, *34*, 81–87. https://doi.org/10.1016/j.cobeha.2020.01.003

Tappin, B. M., Pennycook, G., & Rand, D. G. (2020b). Bayesian or biased? Analytic thinking and political belief updating. *Cognition*, *204*, 104375. https://doi.org/10.1016/j.cognition.2020.104375

The Fake-News Fallacy | The New Yorker. (n.d.). Retrieved July 22, 2021, from https://www.newyorker.com/magazine/2017/09/04/the-fake-news-fallacy

Thompson, V. A. (2009). *Dual-process theories: A metacognitive perspective*. Oxford University Press.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 15.

Thornhill, C., Meeus, Q., Peperkamp, J., & Berendt, B. (2019). A Digital Nudge to Counter Confirmation Bias. *Frontiers in Big Data*, *0*. https://doi.org/10.3389/fdata.2019.00011

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, *39*(7), 1275-1289.

Törnberg, P. (2018). Echo chambers and viral misinformation: Modeling fake news as complex contagion. PLoS one, 13(9), e0203958.

Tsipursky, G., & Morford, Z. (2018). *Addressing behaviors that lead to sharing fake news*. Behavior and Social Issues, 27, AA6-AA10.

Twitter (2021). Helping you find accurate US Election News and Information. Retrieved July 26, 2021, from https://blog.twitter.com/en_us/topics/company/2020/2020-election-news

Vaccination: When 'Fake News' Has Lasting Consequences. (n.d.). Lehigh University Business Thought Leadership. Retrieved July 19, 2021, from https://player.captivate.fm/episode/bf59656f-7c34-4f87-b738-1a140597de4c

Van Bavel, J. J., & Pereira, A. (2018). The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences*, *22*(3), 213–224. https://doi.org/10.1016/j.tics.2018.01.004

van Der Linden, S., Roozenbeek, J., & Compton, J. (2020). Inoculating against fake news about COVID-19. *Frontiers in psychology*, *11*, 2928.

van der Linden, S., Roozenbeek, J., Maertens, R., Basol, M., Kácha, O., Rathje, S., & Traberg, C. S. (2021). How can psychological science help counter the spread of fake news?. The Spanish Journal of Psychology, 24.

Varol, O., Ferrara, E., Davis, C., Menczer, F., & Flammini, A. (2017, May). Online human-bot interactions: Detection, estimation, and characterization. In *Proceedings of the international AAAI conference on web and social media* (Vol. 11, No. 1).

Vegetti, F., & Mancosu, M. (2020). *The impact of political sophistication and motivated reasoning on misinformation. Political Communication,* 37(5), 678-695.

Vegetti, F., & Mancosu, M. (2020). The impact of political sophistication and motivated reasoning on misinformation. *Political Communication*, *37*(5), 678-695.

Volkova, S., Bachrach, Y., Armstrong, M., & Sharma, V. (2015, March). Inferring latent user properties from texts published in social media. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wakefield, J. (2017). Facebook's fake news experiment backfires. Retrieved August 8, 2018, from www.bbc.co.uk/news/technology-41900877

Wang, W.-C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2016). On Known Unknowns: Fluency and the Neural Mechanism

Appendix

Appendix I: Sample Used for Truth and Sharing Discernment Regressions

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
truth_discernment	520	28.18	29.20	-44.44	11.11	44.44	100.00
sharing_discernment	212	3.09	17.11	-44.44	0.00	11.11	66.67
political_knowledge	520	2.96	1.08	0	2	4	5
crt_score	520	3.59	1.98	0	2	5	7
pit	520	4.68	1.47	1	4	6	7
pet	520	5.28	1.59	1	4.8	6	7
age	518	39.61	12.68	19.00	30.00	48.75	79.00
identifies_fake	520	4.64	1.09	1	4	5	7
prefers_republican	520	3.22	1.69	1	2	4	7
rep_margin	513	-6.40	28.77	-74.47	-23.17	11.15	67.93
income	520	6.93	3.41	1	4	10	13
q_total_duration	520	969.23	487.51	337	666.8	1,134	4,193
rtlastclick	520	17.57	28.35	0.00	7.56	18.43	412.69
income_high_dummy	520	0.50	0.50	0	0	1	1
race_white_dummy	520	0.78	0.42	0	1	1	1
republican_dummy	520	0.28	0.45	0	0	1	1
google_dummy	520	0.01	0.08	0	0	0	1
crt_score_high_dummy	520	0.51	0.50	0	0	1	1
education_higher_dummy	520	0.59	0.49	0	0	1	1
facebook_daily_dummy	520	0.59	0.49	0	0	1	1

polarized_zipcode_dummy	513	0.17	0.38	0.00	0.00	0.00	1.00
chk_color_dummy	520	0.99	0.10	0	1	1	1
chk_news_dummy	520	0.95	0.23	0	1	1	1
chk_recognize_dummy	520	0.96	0.20	0	1	1	1

Appendix II: Sample Used for Likely and Share Regressions

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
likely	9,215	3.23	1.68	1	2	5	6
share	3,744	2.02	1.43	1.00	1.00	3.00	6.00
fake_dummy	9,215	0.50	0.50	0	0	1	1
treatment_dummy	9,215	0.50	0.50	0	0	1	1
bias	9,215	-2.75	7.68	-21	-6	-1	21
crt_score	9,215	3.58	1.98	0	2	5	7
political_knowledge	9,215	2.96	1.08	0	2	4	5
education_higher_dummy	9,215	0.59	0.49	0	0	1	1
pit_centered	9,215	-0.33	1.47	-4	-1	1	2
pet_centered	9,215	-0.72	1.58	-5	-1.5	0	1
concordance_dummy	9,215	0.33	0.47	0	0	1	1
prefers_republican	9,215	3.22	1.70	1	2	4	7
concordance_zipcode_dummy	9,215	0.17	0.37	0	0	0	1
polarized_zipcode_dummy	9,215	0.17	0.38	0	0	0	1
age	9,215	39.67	12.71	19	30	49	79
race_white_dummy	9,215	0.78	0.41	0	1	1	1
income_high_dummy	9,215	0.50	0.50	0	0	1	1
facebook_daily_dummy	9,215	0.59	0.49	0	0	1	1

Appendix III: Randomization Check

Variable	Overall, N = 520[1]	control, N = 261[1]	treatment, N = 259[1]	p-value[2]
truth_discernment	0.89 (0.22, 1.67)	1.11 (0.33, 1.67)	0.78 (0.22, 1.67)	0.2
sharing_discernment	0.11 (-0.11, 0.56)	0.11 (-0.22, 0.44)	0.11 (-0.11, 0.56)	0.2
Missing	308	151	157	
truth_discernment2	22 (11, 44)	22 (11, 56)	22 (11, 44)	0.4
sharing_discernment2	0 (0, 11)	0 (-11, 8)	0 (0, 11)	0.003
Missing	308	151	157	
political_knowledge	3.00 (2.00, 4.00)	3.00 (2.00, 4.00)	3.00 (2.00, 4.00)	0.2
crt_score	4.00 (2.00, 5.00)	3.00 (2.00, 5.00)	4.00 (2.00, 5.00)	0.2
pit	5.00 (4.00, 6.00)	5.00 (3.00, 6.00)	5.00 (4.00, 6.00)	0.2
pet	6.00 (4.75, 6.00)	6.00 (5.00, 6.00)	6.00 (4.00, 7.00)	0.8
age	37 (30, 49)	38 (31, 48)	36 (29, 50)	0.7
Missing	2	1	1	
identifies_fake	5.00 (4.00, 5.00)	5.00 (4.00, 5.00)	5.00 (4.00, 5.00)	0.051
prefers_republican	3.00 (2.00, 4.00)	3.00 (2.00, 4.00)	3.00 (2.00, 4.00)	0.6
rep_margin	-5 (-23, 11)	-4 (-26, 16)	-7 (-21, 8)	0.2
Missing	7	3	4	
income	7.0 (4.0, 10.0)	7.0 (4.0, 10.0)	7.0 (4.0, 10.0)	>0.9
q_total_duration	839 (667, 1,134)	837 (635, 1,092)	844 (690, 1,172)	0.2
rtlastclick	12 (8, 18)	12 (8, 20)	12 (7, 18)	0.7
income_high_dummy	258 (50%)	134 (51%)	124 (48%)	0.4
race_white_dummy	405 (78%)	202 (77%)	203 (78%)	0.8
republican_dummy	145 (28%)	68 (26%)	77 (30%)	0.3
google_dummy	3 (0.6%)	3 (1.1%)	0 (0%)	0.2
crt_score_high_dummy	267 (51%)	128 (49%)	139 (54%)	0.3

education__higher__dummy	305 (59%)	155 (59%)	150 (58%)	0.7
facebook__daily__dummy	305 (59%)	148 (57%)	157 (61%)	0.4
polarized__zipcode__dummy	89 (17%)	43 (17%)	46 (18%)	0.7
Missing	7	3	4	
chk__color__dummy	515 (99%)	258 (99%)	257 (99%)	>0.9
chk__news__dummy	492 (95%)	251 (96%)	241 (93%)	0.12
chk__recognize__dummy	498 (96%)	251 (96%)	247 (95%)	0.6
education4__factor				0.030
high school or lower	58 (11%)	36 (14%)	22 (8.5%)	
some college/associate degree	157 (30%)	70 (27%)	87 (34%)	
bachelor's degree	209 (40%)	114 (44%)	95 (37%)	
master's degree or higher	96 (18%)	41 (16%)	55 (21%)	
gender__factor				0.2
male	183 (35%)	98 (38%)	85 (33%)	
female	332 (64%)	162 (62%)	170 (66%)	
other	5 (1.0%)	1 (0.4%)	4 (1.5%)	
political__factor				0.7
Yes	212 (41%)	110 (42%)	102 (39%)	
No	250 (48%)	124 (48%)	126 (49%)	
Never share	58 (11%)	27 (10%)	31 (12%)	
fabricate__factor				>0.9
Yes	129 (25%)	65 (25%)	64 (25%)	
Not sure	210 (40%)	106 (41%)	104 (40%)	
No	181 (35%)	90 (34%)	91 (35%)	

[1]Median (IQR); n (%)

[2]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

Appendix IV: News Headlines, Their Believability and Descriptions

Type	Slant	Item	Mean Belief Control	Mean Belief Treatment	Identified Correct Control (%)	Identified Correct Treatment (%)
True	right	Trump gets endorsement of NYC police union, warns 'no one will be safe in Biden's America'	4.3	4.4	76.6	78
True	right	Kentucky Attorney General Daniel Cameron pushes back on Biden's Black voters comments in RNC speech	4.2	4.3	78.2	81.1
True	right	Chinese dissident brought to US by Obama administration praises Trump at RNC	3.2	3.4	47.9	49
Fake	right	Donald Trump Sent His Own Plane To Transport 200 Stranded Marines	1.9	2	85.8	82.2
Fake	right	Rosa Parks' Granddaughter BLASTS Liberals: 'She Would Have Stood For The Anthem If It Played On That Bus'	2.9	3.1	62.1	56.4
Fake	right	Trump Reveals Which Democratic President Was Also a KKK Member, Liberals In Meltdown Mode	2.9	3.1	61.3	59.1
True	neutral	Democrats Fend Off Attempts to Back Medicare for All in Platform	3.6	3.8	58.2	65.6
True	neutral	Melania Trump statue set on fire in Slovenia	3.3	3.2	50.6	47.5
True	neutral	Trump says top Republican told him Congress would not force Pentagon to change Confederate names of military bases	4	4	74.3	70.3
Fake	neutral	Kamala Harris: 'White lab coats a sign of doctors' racism'	1.9	2.1	83.9	82.2
Fake	neutral	Michigan House Passes Human Microchipping Legislation - Repub.Li	2	2.1	82.8	78.4
Fake	neutral	Rutgers declared Grammar Racist	2.6	2.6	67.8	68
True	left	Kamala Harris Crystallizes Trump's View of Women: They are 'Nasty' or Housewives	3.8	3.7	62.1	62.2

True	left	Dem. senator accusing Trump of 'killing people' with rallies quiet on health risk from protests	3.8	3.7	61.3	61.4
True	left	Facebook removes Trump ads with symbols once use by Nazis	3.4	3.4	53.3	54.1
Fake	left	Trump pays Giuliani to stay silent	2.7	2.9	66.7	60.2
Fake	left	W.H. Staffers Defect, Releasing Private Tape Recording That Has Trump Dead Silent	3.1	3	56.3	58.7
Fake	left	"Wounded social justice warrior" project gives hope	3.7	4	36	27.8

Appendix V: The Role of Education on Believing or Sharing Fake News

Dependent variable: belief in news (likely; ranges from 1 - 6)		
	4a.1	4a.2
(Intercept)	2.251***	2.239***
	(0.050)	(0.064)
fake_dummy	-0.144*	-0.092
	(0.072)	(0.095)
education_higher_dummy	-0.228***	-0.356***
	(0.066)	(0.085)
fake_dummy × education_higher_dummy	-0.120	-0.065
	(0.093)	(0.124)
treatment_dummy		0.027
		(0.102)
fake_dummy × treatment_dummy		-0.115
		(0.145)
education_higher_dummy × treatment_dummy		0.252+
		(0.134)
fake_dummy × education_higher_dummy × treatment_dummy		-0.097
		(0.188)
Num.Obs.	3816	3816
R2	0.016	0.019
R2 Adj.	0.015	0.017
AIC	13514.1	13510.5
BIC	13545.3	13566.7
F	20.437	10.438

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix VI: The Role of The Preference Towards Effortful and Intuitive Thinking on Sharing Fake News

Dependent variable: sharing inclination (share; ranges from 1 - 6)				
	4b.1	4b.2	4b.3	4b.4
(Intercept)	2.126***	2.048***	2.135***	2.104***
	(0.035)	(0.048)	(0.082)	(0.106)
fake_dummy	-0.195***	-0.111	-0.331**	-0.295+
	(0.050)	(0.070)	(0.113)	(0.151)
pit_centered	0.048+	0.042		
	(0.027)	(0.034)		
pet_centered	-0.023	-0.012		
	(0.023)	(0.026)		
fake_dummy × pit_centered	0.121***	0.129**		
	(0.036)	(0.049)		
fake_dummy × pet_centered	-0.024	-0.002		
	(0.033)	(0.039)		
treatment_dummy		0.167*		0.160
		(0.071)		(0.172)
fake_dummy × treatment_dummy		-0.159		-0.101
		(0.100)		(0.237)
pit_centered × treatment_dummy		0.019		
		(0.052)		
pet_centered × treatment_dummy		-0.072		
		(0.053)		
fake_dummy × pit_centered × treatment_dummy		-0.017		

		(0.071)		
fake_dummy × pet_centered × treatment_dummy		-0.028		
		(0.073)		
pit_high_dummy			0.094	0.024
			(0.075)	(0.101)
pet_high_dummy			-0.124+	-0.135
			(0.069)	(0.087)
fake_dummy × pit_high_dummy			0.249*	0.240+
			(0.101)	(0.142)
fake_dummy × pet_high_dummy			-0.102	-0.032
			(0.098)	(0.127)
pit_high_dummy × treatment_dummy				0.154
				(0.149)
pet_high_dummy × treatment_dummy				-0.111
				(0.151)
fake_dummy × pit_high_dummy × treatment_dummy				0.009
				(0.203)
fake_dummy × pet_high_dummy × treatment_dummy				-0.080
				(0.211)
Num.Obs.	3816	3816	3816	3816
R2	0.021	0.025	0.016	0.019
R2 Adj.	0.020	0.022	0.014	0.017
AIC	13498.3	13495.2	13518.4	13515.9
BIC	13542.0	13576.5	13562.2	13597.1
Log.Lik.	-6742.128	-6734.622	-6752.221	-6744.951
F	16.293	8.786	12.186	6.872
Std. Errors	Robust	Robust	Robust	Robust

Dependent variable: share (1-6)

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix VII: The Role of the CRT Performance on Sharing Fake News

Dependent variable: sharing inclination (share; ranges from 1 - 6)				
	4c.1	4c.2	4c.3	4c.4
(Intercept)	2.347***	2.308***	2.227***	2.125***
	(0.068)	(0.084)	(0.051)	(0.064)
fake_dummy	-0.002	-0.025	-0.082	-0.034
	(0.097)	(0.120)	(0.073)	(0.093)
crt_score	-0.061***	-0.075***		
	(0.016)	(0.019)		
fake_dummy × crt_score	-0.057**	-0.030		
	(0.022)	(0.027)		
treatment_dummy		0.114		0.238*
		(0.141)		(0.105)
fake_dummy × treatment_dummy		0.031		-0.112
		(0.200)		(0.150)
crt_score × treatment_dummy		0.020		
		(0.032)		
fake_dummy × crt_score × treatment_dummy		-0.049		
		(0.044)		
crt_score_high_dummy			-0.185**	-0.147+
			(0.066)	(0.085)
fake_dummy × crt_score_high_dummy			-0.234*	-0.184
			(0.094)	(0.124)
crt_score_high_dummy × treatment_dummy				-0.115
				(0.135)
fake_dummy × crt_score_high_dummy × treatment_dummy				-0.074

				(0.190)
Num.Obs.	3816	3816	3816	3816
R2	0.024	0.026	0.018	0.021
R2 Adj.	0.023	0.024	0.017	0.019
AIC	13484.1	13482.5	13505.5	13503.4
BIC	13515.3	13538.7	13536.7	13559.6
Log.Lik.	-6737.057	-6732.226	-6747.728	-6742.680
F	30.622	14.522	23.365	11.471
Std. Errors	Robust	Robust	Robust	Robust

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.00$

Appendix VIII: The Role of Political Knowledge in Sharing Fake News

Dependent variable: sharing inclination (share; ranges from 1 - 6)

	4d.1	4d.2	4d.3	4d.4
(Intercept)	2.479***	2.723***	2.381***	2.359***
	(0.096)	(0.129)	(0.064)	(0.083)
fake_dummy	0.127	-0.005	-0.078	-0.079
	(0.137)	(0.187)	(0.093)	(0.121)
political_knowledge	-0.117***	-0.227***		
	(0.030)	(0.039)		
fake_dummy × political_knowledge	-0.112**	-0.041		
	(0.042)	(0.056)		
treatment_dummy		-0.472*		0.055
		(0.192)		(0.131)
fake_dummy × treatment_dummy		0.245		0.004
		(0.274)		(0.189)
political_knowledge × treatment_dummy		0.212***		
		(0.060)		
fake_dummy × political_knowledge × treatment_dummy		-0.135		
		(0.084)		
political_knowledge_high_dummy			-0.357***	-0.451***
			(0.074)	(0.096)
fake_dummy × political_knowledge_high_dummy			-0.185+	-0.069
			(0.106)	(0.140)
political_knowledge_high_dummy × treatment_dummy				0.176
				(0.151)
fake_dummy × political_knowledge_high_dummy × treatment_dummy				-0.229

					(0.216)
Num.Obs.		3816	3816	3816	3816
R2		0.026	0.032	0.026	0.029
R2 Adj.		0.025	0.030	0.025	0.027
BIC		13505.5	13516.2	13504.9	13528.3
F		33.976	17.826	34.196	16.038

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix IX: The Role of the Concordance with Political Ideology on Sharing Fake News

Dependent variable: sharing inclination (share; ranges from 1 - 6)				
	4e1	4e.2	4e.3	4e.4
(Intercept)	1.985***	1.932***	1.991***	1.934***
	(0.042)	(0.055)	(0.042)	(0.056)
fake_dummy	-0.325***	-0.211**	-0.278***	-0.155+
	(0.057)	(0.079)	(0.057)	(0.080)
concordance_dummy	0.430***	0.353***		
	(0.075)	(0.097)		
concordance_zipcode_dummy	0.011	-0.063	-0.010	-0.063
	(0.091)	(0.109)	(0.091)	(0.110)
fake_dummy × concordance_dummy	0.242*	0.158		
	(0.106)	(0.140)		
fake_dummy × concordance_zipcode_dummy	0.163	0.115	0.139	0.091
	(0.128)	(0.160)	(0.129)	(0.163)
treatment_dummy		0.107		0.120
		(0.084)		(0.083)
fake_dummy × treatment_dummy		-0.233*		-0.248*
		(0.115)		(0.115)
concordance_dummy × treatment_dummy		0.162		
		(0.150)		
concordance_zipcode_dummy × treatment_dummy		0.203		0.145
		(0.189)		(0.189)
fake_dummy × concordance_dummy × treatment_dummy		0.172		
		(0.212)		

fake_dummy × concordance_zipcode_dummy × treatment_dummy		0.078		0.062
		(0.262)		(0.265)
concordance2_dummy			0.421***	0.348***
			(0.074)	(0.096)
fake_dummy × concordance2_dummy			0.112	0.005
			(0.105)	(0.139)
concordance2_dummy × treatment_dummy				0.147
				(0.149)
fake_dummy × concordance2_dummy × treatment_dummy				0.219
				(0.210)
Num.Obs.	3762	3762	3762	3762
R2	0.040	0.045	0.031	0.036
R2 Adj.	0.039	0.042	0.030	0.033
AIC	13220.1	13213.1	13256.1	13249.6
BIC	13263.7	13294.1	13299.8	13330.6
Log.Lik.	-6603.038	-6593.531	-6621.067	-6611.797
F	31.603	16.142	24.136	12.691
Std. Errors	Robust	Robust	Robust	Robust

Dependent variable: share (1-6)

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix X: The Role of the Politically Polarized Environment on Sharing Fake News

Dependent variable: sharing inclination (share; ranges from 1 - 6)		
	4g.1	4g.2
(Intercept)	2.151***	2.076***
	(0.036)	(0.046)
fake_dummy	-0.226***	-0.121+
	(0.051)	(0.068)
polarized_zipcode_dummy	-0.133	-0.268*
	(0.090)	(0.121)
fake_dummy × polarized_zipcode_dummy	0.045	-0.124
	(0.125)	(0.162)
treatment_dummy		0.160*
		(0.073)
fake_dummy × treatment_dummy		-0.225*
		(0.102)
polarized_zipcode_dummy × treatment_dummy		0.235
		(0.179)
fake_dummy × polarized_zipcode_dummy × treatment_dummy		0.345
		(0.246)
Num.Obs.	3762	3762
R2	0.007	0.012
R2 Adj.	0.006	0.010
AIC	13345.9	13333.1
BIC	13377.0	13389.2
F	8.429	6.603

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix XI: Full and Reduced Final Regressions

Dependent variables: belief in the news (likely; ranges from 1 - 6) and sharing intentions (share; ranges from 1-6)

	likely (full)	likely (reduced)	share (full)	share (reduced)
(Intercept)	3.545***	3.461***	1.402***	1.370***
	(0.119)	(0.103)	(0.175)	(0.126)
fake_dummy	-0.982***	-0.856***	-0.087	-0.071
	(0.157)	(0.130)	(0.229)	(0.102)
treatment_dummy	0.028	0.057+	0.189**	0.184**
	(0.046)	(0.031)	(0.062)	(0.061)
bias	0.020***	0.019***	0.035***	0.037***
	(0.005)	(0.003)	(0.007)	(0.007)
crt_score	0.019	0.021	-0.061***	-0.062***
	(0.015)	(0.014)	(0.018)	(0.011)
political_knowledge	0.098***	0.097***	-0.098**	-0.108***
	(0.022)	(0.022)	(0.030)	(0.019)
education_higher_dummy	-0.052	-0.068	-0.118+	-0.121**
	(0.050)	(0.048)	(0.066)	(0.044)
pit_centered	-0.047**	-0.049**	-0.011	
	(0.018)	(0.017)	(0.025)	
pet_centered	0.049**	0.036***	-0.022	-0.021
	(0.015)	(0.010)	(0.022)	(0.015)
concordance_dummy	0.239**	0.241**	0.347***	0.282***
	(0.082)	(0.081)	(0.103)	(0.065)
prefers_republican	-0.036*	-0.039**	0.110***	0.097***
	(0.015)	(0.015)	(0.019)	(0.012)
concordance_zipcode_dummy	0.196**	0.185**	-0.079	-0.084
	(0.063)	(0.061)	(0.080)	(0.079)

polarized_zipcode_dummy	-0.064		-0.149+	-0.098+
	(0.062)		(0.079)	(0.052)
age	-0.002		-0.004	
	(0.001)		(0.003)	
race_white_dummy	0.036		-0.135	-0.188***
	(0.058)		(0.087)	(0.056)
income_high_dummy	-0.053		-0.108+	-0.114+
	(0.049)		(0.064)	(0.061)
facebook_daily_dummy	0.089+	0.103**	-0.106	-0.090*
	(0.048)	(0.032)	(0.068)	(0.044)
fake_dummy × treatment_dummy	0.062		-0.122	-0.115
	(0.063)		(0.082)	(0.079)
fake_dummy × bias	0.034***	0.034***	0.001	
	(0.005)	(0.005)	(0.007)	
fake_dummy × crt_score	-0.052**	-0.049**	0.010	
	(0.018)	(0.017)	(0.021)	
fake_dummy × political_knowledge	-0.205***	-0.209***	-0.002	
	(0.030)	(0.029)	(0.040)	
fake_dummy × education_higher_dummy	-0.157*	-0.137*	-0.005	
	(0.068)	(0.065)	(0.088)	
fake_dummy × pit_centered	0.104***	0.109***	0.034	
	(0.024)	(0.023)	(0.032)	
fake_dummy × pet_centered	-0.026		0.003	
	(0.021)		(0.030)	
fake_dummy × concordance_dummy	0.696***	0.697***	-0.205*	-0.200*
	(0.071)	(0.071)	(0.092)	(0.091)
fake_dummy × prefers_republican	0.174***	0.179***	-0.024	
	(0.020)	(0.020)	(0.025)	

fake_dummy × concordance_zipcode_dummy	-0.304***	-0.292***	0.167	0.178+
	(0.084)	(0.082)	(0.107)	(0.104)
fake_dummy × polarized_zipcode_dummy	0.063		0.095	
	(0.084)		(0.105)	
fake_dummy × age	0.001		0.002	
	(0.002)		(0.004)	
fake_dummy × race_white_dummy	-0.008		-0.077	
	(0.079)		(0.114)	
fake_dummy × income_high_dummy	0.089		0.125	0.137+
	(0.066)		(0.084)	(0.078)
fake_dummy × facebook_daily_dummy	0.034		0.052	
	(0.064)		(0.091)	
bias × crt_score	0.000		-0.004*	-0.004*
	(0.001)		(0.002)	(0.002)
crt_score × concordance_dummy	-0.023	-0.024	-0.017	
	(0.018)	(0.017)	(0.021)	
likely			0.318***	0.317***
			(0.019)	(0.019)
fake_dummy × likely			0.079**	0.079**
			(0.029)	(0.028)
Num.Obs.	9215	9215	3744	3744
R2	0.215	0.214	0.299	0.298
R2 Adj.	0.212	0.212	0.293	0.294
AIC	33551.8	33536.1	12025.0	12004.0
BIC	33801.3	33700.1	12255.4	12147.2
Log.Lik.	-16740.905	-16745.058	-5975.480	-5978.992
F	75.998	119.079	45.253	75.232
Std. Errors	Robust	Robust	Robust	Robust

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix XII: Mixed-level Linear Regressions

Dependent variables: belief in the news (likely; ranges from 1 - 6) and sharing intentions (share; ranges from 1-6)				
	likely (full) mixed	likely (reduced) mixed	share (full) mixed	share (reduced) mixed
(Intercept)	3.533***	3.575***	1.293***	1.239***
	(0.185)	(0.171)	(0.321)	(0.290)
fake_dummy	-0.921***	-0.887***	0.066	0.207
	(0.198)	(0.191)	(0.272)	(0.220)
treatment_dummy	0.024	0.026	0.192	
	(0.155)	(0.159)	(0.235)	
bias	0.020***	0.018***	0.031***	0.032***
	(0.005)	(0.003)	(0.005)	(0.005)
crt_score	0.019	0.016	-0.060+	-0.069*
	(0.019)	(0.017)	(0.031)	(0.030)
political_knowledge	0.099**	0.094**	-0.099+	-0.100+
	(0.032)	(0.031)	(0.053)	(0.053)
education_higher_dummy	-0.063	-0.079	-0.116	-0.113
	(0.070)	(0.067)	(0.118)	(0.113)
pit_centered	-0.048*	-0.043+	-0.010	0.007
	(0.024)	(0.023)	(0.044)	(0.042)
pet_centered	0.049*	0.051*	-0.023	-0.018
	(0.021)	(0.021)	(0.040)	(0.038)
concordance_dummy	0.324***	0.249***	0.566***	0.498***
	(0.078)	(0.055)	(0.086)	(0.062)
prefers_republican	-0.035+	-0.037+	0.112***	0.105***
	(0.020)	(0.019)	(0.031)	(0.030)
concordance_zipcode_dummy	0.005		-0.048	
	(0.062)		(0.072)	

polarized_zipcode_dummy	-0.026		-0.154	-0.118
	(0.087)		(0.157)	(0.149)
age	-0.002		-0.004	-0.004
	(0.002)		(0.005)	(0.005)
race_white_dummy	0.048		-0.138	
	(0.081)		(0.148)	
income_high_dummy	-0.053		-0.107	-0.050
	(0.068)		(0.114)	(0.109)
facebook_daily_dummy	0.092		-0.104	-0.062
	(0.066)		(0.125)	(0.119)
fake_dummy × treatment_dummy	0.066	0.068	-0.126	0.026
	(0.207)	(0.214)	(0.299)	(0.227)
fake_dummy × bias	0.035***	0.035***	0.003	0.004
	(0.004)	(0.004)	(0.004)	(0.004)
fake_dummy × crt_score	-0.052**	-0.051**	0.011	0.003
	(0.016)	(0.016)	(0.018)	(0.016)
fake_dummy × political_knowledge	-0.206***	-0.205***	-0.010	-0.009
	(0.028)	(0.028)	(0.031)	(0.028)
fake_dummy × education_higher_dummy	-0.144*	-0.112+	-0.021	
	(0.063)	(0.060)	(0.069)	
fake_dummy × pit_centered	0.104***	0.106***	0.044+	
	(0.021)	(0.021)	(0.026)	
fake_dummy × pet_centered	-0.027	-0.026	0.007	
	(0.019)	(0.019)	(0.023)	
fake_dummy × concordance_dummy	0.469***	0.466***	-0.065	-0.058
	(0.078)	(0.078)	(0.089)	(0.088)
fake_dummy × prefers_republican	0.173***	0.176***	-0.016	-0.013
	(0.018)	(0.017)	(0.018)	(0.017)
fake_dummy × concordance_zipcode_dummy	-0.089		0.066	

	(0.085)		(0.097)	
fake_dummy × polarized_zipcode_dummy	0.020		0.111	
	(0.078)		(0.093)	
fake_dummy × age	0.001		0.003	
	(0.002)		(0.003)	
fake_dummy × race_white_dummy	-0.023		-0.067	
	(0.072)		(0.086)	
fake_dummy × income_high_dummy	0.089		0.117+	
	(0.061)		(0.066)	
fake_dummy × facebook_daily_dummy	0.031		0.048	
	(0.059)		(0.072)	
bias × crt_score	0.000		-0.003**	-0.004***
	(0.001)		(0.001)	(0.001)
crt_score × concordance_dummy	-0.022		-0.018	
	(0.016)		(0.017)	
SD (Intercept)	0.172	0.178	0.252	0.248
	0.172	0.178	0.252	0.717
	0.172	0.178	0.715	0.248
	0.172	0.178	0.715	0.717
	0.172	0.564	0.252	0.248
	0.172	0.564	0.252	0.717
	0.172	0.564	0.715	0.248
	0.172	0.564	0.715	0.717
	0.566	0.178	0.252	0.248
	0.566	0.178	0.252	0.717
	0.566	0.178	0.715	0.248
	0.566	0.178	0.715	0.717
	0.566	0.564	0.252	0.248

	0.566	0.564	0.252	0.717
	0.566	0.564	0.715	0.248
	0.566	0.564	0.715	0.717
SD (Observations)	1.174	1.174	0.982	0.982
likely			0.324***	0.330***
			(0.016)	(0.012)
fake_dummy × likely			0.012	
			(0.023)	
Num.Obs.	9215	9215	3744	3744
R2 Marg.	0.205	0.203	0.294	0.290
R2 Cond.	0.328	0.327	0.563	0.561
AIC	33019.2	32925.5	11103.1	11030.2
BIC	33282.9	33089.5	11346.0	11185.9
ICC	0.2	0.2	0.4	0.4
RMSE	1.35	1.35	0.94	0.94

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix XIII: Full and Reduced Logit Regressions

Dependent variables: likely_high_dummy (binary; 1 if likely >3) and share_high_dummy ((binary; 1 if share >3)

	likely (full) logit	likely (reduced) logit	share (full) logit	share (reduced) logit
(Intercept)	0.263+	0.141	-3.380***	-3.561***
	(0.149)	(0.137)	(0.416)	(0.294)
fake_dummy	-1.163***	-0.971***	0.107	0.533***
	(0.224)	(0.188)	(0.586)	(0.145)
treatment_dummy	0.041	0.073	0.574***	0.550***
	(0.063)	(0.046)	(0.138)	(0.130)
bias	0.026***	0.026***	0.035***	0.036***
	(0.004)	(0.004)	(0.008)	(0.006)
crt_score	0.023	0.021	-0.121***	-0.130***
	(0.017)	(0.017)	(0.037)	(0.026)
political_knowledge	0.118***	0.113***	-0.193**	-0.171***
	(0.030)	(0.030)	(0.065)	(0.047)
education_higher_dummy	-0.100	-0.122+	-0.076	-0.156
	(0.068)	(0.066)	(0.143)	(0.102)
pit_centered	-0.079***	-0.082***	0.007	
	(0.024)	(0.023)	(0.054)	
pet_centered	0.066**	0.053***	-0.042	
	(0.020)	(0.015)	(0.049)	
concordance_dummy	0.141*	0.140*	0.388**	0.355***
	(0.068)	(0.068)	(0.136)	(0.101)
prefers_republican	-0.030	-0.037*	0.169***	0.154***
	(0.019)	(0.019)	(0.038)	(0.028)
concordance_zipcode_dummy	0.234**	0.234**	-0.187	
	(0.088)	(0.088)	(0.177)	

polarized_zipcode_dummy	-0.079	-0.090	-0.181	-0.315*
	(0.084)	(0.084)	(0.188)	(0.143)
age	-0.002		-0.011+	-0.011*
	(0.002)		(0.007)	(0.005)
race_white_dummy	0.115	0.125*	-0.339+	-0.353**
	(0.077)	(0.056)	(0.179)	(0.132)
income_high_dummy	-0.090		-0.158	
	(0.065)		(0.137)	
facebook_daily_dummy	0.116+	0.136**	-0.081	
	(0.064)	(0.047)	(0.157)	
fake_dummy × treatment_dummy	0.075		-0.416*	-0.436*
	(0.093)		(0.208)	(0.196)
fake_dummy × bias	0.048***	0.048***	0.002	
	(0.007)	(0.007)	(0.013)	
fake_dummy × crt_score	-0.096***	-0.089***	0.009	
	(0.026)	(0.025)	(0.055)	
fake_dummy × political_knowledge	-0.244***	-0.244***	0.054	
	(0.045)	(0.044)	(0.097)	
fake_dummy × education_higher_dummy	-0.196+	-0.169+	-0.123	
	(0.101)	(0.097)	(0.215)	
fake_dummy × pit_centered	0.147***	0.155***	0.066	
	(0.035)	(0.034)	(0.086)	
fake_dummy × pet_centered	-0.027		0.006	
	(0.030)		(0.075)	
fake_dummy × concordance_dummy	0.954***	0.953***	-0.130	
	(0.100)	(0.099)	(0.208)	
fake_dummy × prefers_republican	0.204***	0.214***	-0.045	
	(0.029)	(0.028)	(0.057)	

fake_dummy × concordance_zipcode_dummy	-0.480***	-0.477***	0.228	
	(0.133)	(0.133)	(0.278)	
fake_dummy × polarized_zipcode_dummy	0.171	0.183	-0.212	
	(0.125)	(0.125)	(0.297)	
fake_dummy × age	0.002		0.003	
	(0.003)		(0.010)	
fake_dummy × race_white_dummy	0.057		-0.060	
	(0.117)		(0.270)	
fake_dummy × income_high_dummy	0.135		0.277	
	(0.098)		(0.204)	
fake_dummy × facebook_daily_dummy	0.053		-0.075	
	(0.095)		(0.234)	
likely			0.687***	0.720***
			(0.059)	(0.040)
fake_dummy × likely			0.084	
			(0.083)	
Num.Obs.	9215	9215	3744	3744
BIC	11328.2	11261.4	3006.5	2851.7
Log.Lik.	-5518.031	-5521.142	-1363.384	-1368.244
Std. Errors	Robust	Robust	Robust	Robust

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix XIV: Z-scaled Regressions

Dependent variables: belief in news and sharing intentions (both z-scaled)

	likely (full) z	likely (reduced) z	share (full) z	share (reduced) z
(Intercept)	-0.005	3.461***	-0.001	1.370***
	(0.009)	(0.103)	(0.016)	(0.126)
fake_dummy	-0.328***	-0.856***	0.052**	-0.071
	(0.009)	(0.130)	(0.016)	(0.102)
treatment_dummy	0.018+	0.057+	0.045**	0.184**
	(0.009)	(0.031)	(0.014)	(0.061)
bias	0.165***	0.019***	0.123***	0.037***
	(0.011)	(0.003)	(0.018)	(0.007)
crt_score	-0.016	0.021	-0.072***	-0.062***
	(0.010)	(0.014)	(0.015)	(0.011)
political_knowledge	-0.003	0.097***	-0.075***	-0.108***
	(0.010)	(0.022)	(0.015)	(0.019)
education_higher_dummy	-0.038***	-0.068	-0.041**	-0.121**
	(0.010)	(0.048)	(0.015)	(0.044)
pit_centered	0.004	-0.049**	0.006	
	(0.010)	(0.017)	(0.016)	
pet_centered	0.034***	0.036***	-0.023	-0.021
	(0.010)	(0.010)	(0.017)	(0.015)
concordance_dummy	0.142***	0.241**	0.060***	0.282***
	(0.010)	(0.081)	(0.015)	(0.065)
prefers_republican	0.052***	-0.039**	0.116***	0.097***
	(0.010)	(0.015)	(0.015)	(0.012)
concordance_zipcode_dummy	0.010	0.185**	0.001	-0.084
	(0.009)	(0.061)	(0.014)	(0.079)

polarized_zipcode_dummy	-0.007		-0.027+	-0.098+
	(0.009)		(0.014)	(0.052)
age	-0.013		-0.031	
	(0.010)		(0.024)	
race_white_dummy	0.008		-0.050**	-0.188***
	(0.010)		(0.017)	(0.056)
income_high_dummy	-0.002		-0.016	-0.114+
	(0.010)		(0.015)	(0.061)
facebook_daily_dummy	0.031***	0.103**	-0.028+	-0.090*
	(0.009)	(0.032)	(0.016)	(0.044)
fake_dummy × treatment_dummy	0.009		-0.021	-0.115
	(0.009)		(0.014)	(0.079)
fake_dummy × bias	0.078***	0.034***	0.004	
	(0.011)	(0.005)	(0.018)	
fake_dummy × crt_score	-0.030**	-0.049**	0.007	
	(0.010)	(0.017)	(0.015)	
fake_dummy × political_knowledge	-0.066***	-0.209***	-0.001	
	(0.010)	(0.029)	(0.015)	
fake_dummy × education_higher_dummy	-0.023*	-0.137*	-0.001	
	(0.010)	(0.065)	(0.015)	
fake_dummy × pit_centered	0.045***	0.109***	0.017	
	(0.010)	(0.023)	(0.016)	
fake_dummy × pet_centered	-0.012		0.002	
	(0.010)		(0.017)	
fake_dummy × concordance_dummy	0.098***	0.697***	-0.034*	-0.200*
	(0.010)	(0.071)	(0.015)	(0.091)
fake_dummy × prefers_republican	0.088***	0.179***	-0.014	
	(0.010)	(0.020)	(0.015)	

fake_dummy × concordance_zipcode_dummy	-0.034***	-0.292***	0.022	0.178+
	(0.009)	(0.082)	(0.014)	(0.104)
fake_dummy × polarized_zipcode_dummy	0.007		0.013	
	(0.009)		(0.014)	
fake_dummy × age	0.004		0.016	
	(0.010)		(0.024)	
fake_dummy × race_white_dummy	-0.001		-0.011	
	(0.010)		(0.017)	
fake_dummy × income_high_dummy	0.013		0.022	0.137+
	(0.010)		(0.015)	(0.078)
fake_dummy × facebook_daily_dummy	0.005		0.009	
	(0.009)		(0.016)	
bias × crt_score	-0.003		-0.038*	-0.004*
	(0.011)		(0.017)	(0.002)
crt_score × concordance_dummy	-0.013	-0.024	-0.011	
	(0.010)	(0.017)	(0.014)	
likely			0.419***	0.317***
			(0.017)	(0.019)
fake_dummy × likely			0.046**	0.079**
			(0.017)	(0.028)
Num.Obs.	9215	9215	3744	3744
R2	0.215	0.214	0.299	0.298
R2 Adj.	0.212	0.212	0.293	0.294
AIC	24007.6	33536.1	9339.9	12004.0
BIC	24257.1	33700.1	9570.4	12147.2
Log.Lik.	-11968.811	-16745.058	-4632.973	-5978.992
F	75.998	119.079	45.253	75.232
Std. Errors	Robust	Robust	Robust	Robust

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix XV: Robustness Check - Probit Regression

Dependent variables: likely_high_dummy (binary; 1 if likely >3) and share_high_dummy ((binary; 1 if share >3)				
	likely (full) probit	likely (reduced) probit	share (full) probit	share (reduced) probit
(Intercept)	0.167+	0.091	-1.857***	-1.982***
	(0.092)	(0.084)	(0.233)	(0.163)
fake_dummy	-0.730***	-0.614***	0.026	0.314***
	(0.136)	(0.114)	(0.323)	(0.082)
treatment_dummy	0.025	0.047+	0.332***	0.321***
	(0.039)	(0.028)	(0.077)	(0.073)
bias	0.017***	0.017***	0.020***	0.018***
	(0.002)	(0.002)	(0.005)	(0.004)
crt_score	0.014	0.013	-0.072***	-0.079***
	(0.011)	(0.010)	(0.020)	(0.015)
political_knowledge	0.073***	0.070***	-0.103**	-0.089***
	(0.019)	(0.018)	(0.038)	(0.027)
education_higher_dummy	-0.060	-0.073+	-0.072	-0.109+
	(0.042)	(0.040)	(0.081)	(0.057)
pit_centered	-0.048***	-0.050***	0.008	
	(0.014)	(0.014)	(0.030)	
pet_centered	0.040**	0.031***	-0.019	
	(0.012)	(0.009)	(0.027)	
concordance_dummy	0.086*	0.086*	0.229**	0.200***
	(0.041)	(0.041)	(0.077)	(0.056)
prefers_republican	-0.018	-0.023+	0.095***	0.090***
	(0.012)	(0.012)	(0.022)	(0.016)

concordance_zipcode_dummy	0.144**	0.144**	-0.120	
	(0.054)	(0.054)	(0.098)	
polarized_zipcode_dummy	-0.049	-0.055	-0.103	-0.180*
	(0.052)	(0.051)	(0.105)	(0.079)
age	-0.001		-0.006+	-0.006*
	(0.001)		(0.004)	(0.003)
race_white_dummy	0.071	0.072*	-0.218*	-0.215**
	(0.048)	(0.034)	(0.103)	(0.074)
income_high_dummy	-0.055		-0.103	
	(0.040)		(0.078)	
facebook_daily_dummy	0.070+	0.082**	-0.044	
	(0.039)	(0.029)	(0.089)	
fake_dummy × treatment_dummy	0.049		-0.246*	-0.254*
	(0.056)		(0.116)	(0.109)
fake_dummy × bias	0.025***	0.025***	-0.002	
	(0.004)	(0.004)	(0.008)	
fake_dummy × crt_score	-0.058***	-0.054***	0.002	
	(0.016)	(0.015)	(0.031)	
fake_dummy × political_knowledge	-0.147***	-0.147***	0.029	
	(0.028)	(0.027)	(0.055)	
fake_dummy × education_higher_dummy	-0.121*	-0.103+	-0.045	
	(0.061)	(0.059)	(0.120)	
fake_dummy × pit_centered	0.088***	0.093***	0.037	
	(0.021)	(0.021)	(0.048)	
fake_dummy × pet_centered	-0.018		0.004	
	(0.018)		(0.042)	
fake_dummy × concordance_dummy	0.587***	0.586***	-0.095	
	(0.060)	(0.060)	(0.115)	
fake_dummy × prefers_republican	0.126***	0.132***	-0.018	

	(0.017)	(0.017)	(0.032)	
fake_dummy × concordance_zipcode_dummy	-0.279***	-0.278***	0.148	
	(0.079)	(0.079)	(0.152)	
fake_dummy × polarized_zipcode_dummy	0.103	0.111	-0.111	
	(0.076)	(0.075)	(0.163)	
fake_dummy × age	0.001		0.003	
	(0.002)		(0.005)	
fake_dummy × race_white_dummy	0.021		-0.007	
	(0.071)		(0.151)	
fake_dummy × income_high_dummy	0.085		0.160	
	(0.059)		(0.114)	
fake_dummy × facebook_daily_dummy	0.028		-0.060	
	(0.057)		(0.131)	
likely			0.379***	0.394***
			(0.033)	(0.022)
fake_dummy × likely			0.045	
			(0.045)	
Num.Obs.	9215	9215	3744	3744
BIC	11334.2	11267.5	3011.6	2857.6
Log.Lik.	-5521.032	-5524.183	-1365.904	-1371.202

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix XVI: Post-hoc Analysis - Response Time Regressions

first_click (time until first clicked on the questions) and last_click (time until the last click on the question); both in seconds				
	First click (full)	First click (reduced)	Last click (full)	Last click (reduced)
(Intercept)	13.341***	14.051***	18.242***	18.027***
	(1.403)	(0.867)	(1.720)	(1.347)
fake_dummy	-0.733	-1.284+	0.629	-0.273
	(1.928)	(0.749)	(2.681)	(1.405)
treatment_dummy	-0.672	-0.583	-0.555	-0.541
	(0.649)	(0.631)	(0.711)	(0.692)
bias	-0.078+	-0.081*	-0.073+	-0.075+
	(0.042)	(0.040)	(0.044)	(0.043)
crt_score	0.248		0.273	0.268
	(0.158)		(0.167)	(0.175)
political_knowledge	-0.669*	-0.415+	-1.071***	-0.855***
	(0.290)	(0.213)	(0.314)	(0.228)
education_higher_dummy	0.552		0.680	0.930*
	(0.621)		(0.697)	(0.469)
pit_centered	-0.099		-0.057	
	(0.261)		(0.287)	
pet_centered	-0.100		0.002	
	(0.307)		(0.317)	
concordance_dummy	-0.146		-0.431	
	(0.702)		(0.749)	
concordance_zipcode_dummy	0.049		0.356	
	(0.759)		(0.921)	
polarized_zipcode_dummy	0.352		0.013	
	(0.797)		(0.885)	

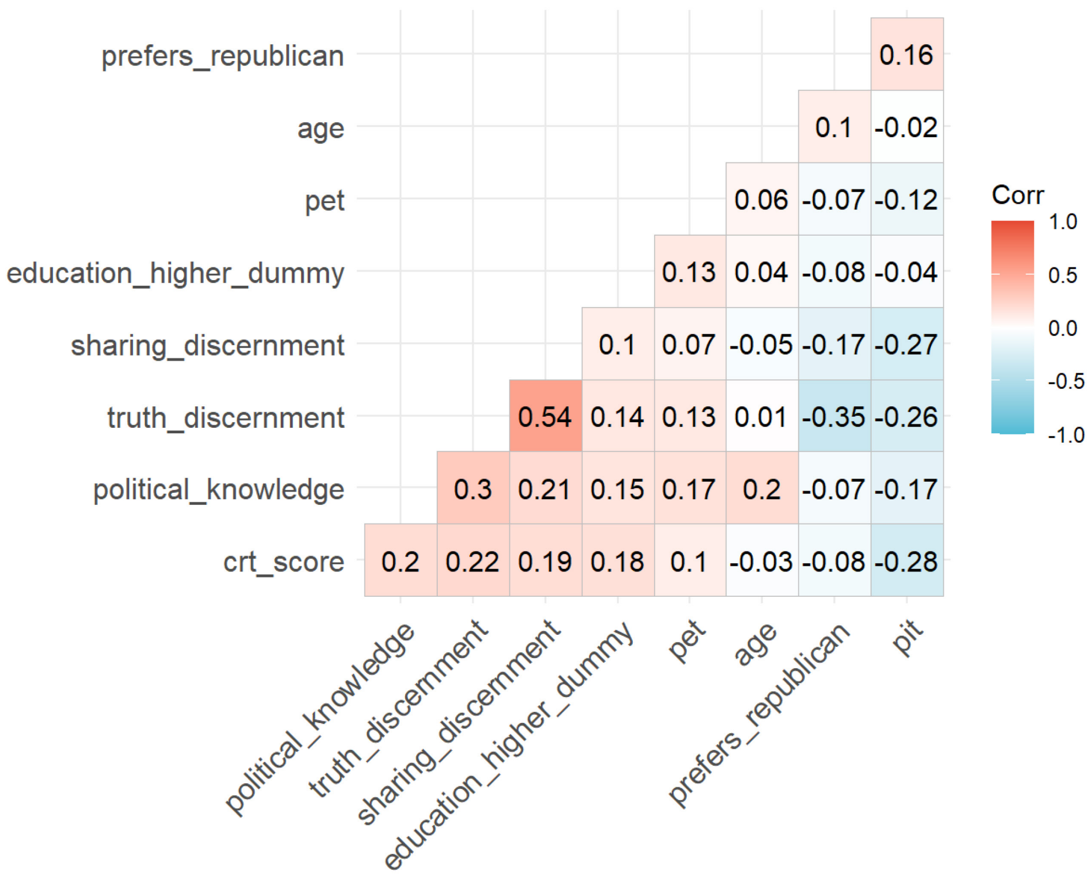
age	0.010		0.001	
	(0.011)		(0.011)	
race_white_dummy	-1.676+	-1.703**	-1.838+	-2.536**
	(0.856)	(0.621)	(0.969)	(0.777)
income_high_dummy	-0.456		-0.975	-1.037
	(0.637)		(0.706)	(0.717)
facebook_daily_dummy	1.046+	1.092+	0.739	0.722
	(0.622)	(0.602)	(0.698)	(0.661)
fake_dummy × treatment_dummy	1.548+	1.454+	1.582	1.616
	(0.871)	(0.854)	(1.010)	(0.989)
fake_dummy × bias	0.147*	0.148*	0.203*	0.194**
	(0.066)	(0.063)	(0.079)	(0.075)
fake_dummy × crt_score	-0.321		-0.483*	-0.534*
	(0.218)		(0.240)	(0.238)
fake_dummy × political_knowledge	0.325		0.437	
	(0.441)		(0.484)	
fake_dummy × education_higher_dummy	0.020		0.347	
	(0.840)		(0.968)	
fake_dummy × pit_centered	0.314		0.504	
	(0.333)		(0.370)	
fake_dummy × pet_centered	0.237		0.410	
	(0.380)		(0.419)	
fake_dummy × concordance_dummy	-0.038		-0.212	
	(0.998)		(1.131)	
fake_dummy × concordance_zipcode_dummy	-0.221		-1.180	
	(1.053)		(1.226)	
fake_dummy × polarized_zipcode_dummy	-0.652		-0.441	
	(1.030)		(1.159)	

fake_dummy × age	-0.003		-0.009	
	(0.015)		(0.016)	
fake_dummy × race_white_dummy	-0.258		-1.489	
	(1.283)		(1.561)	
fake_dummy × income_high_dummy	0.718		1.499	1.694+
	(0.877)		(1.009)	(1.016)
fake_dummy × facebook_daily_dummy	-1.590+	-1.521+	-2.127*	-1.916*
	(0.837)	(0.844)	(0.977)	(0.976)
Num.Obs.	9215	9215	9215	9215
R2	0.006	0.005	0.010	0.008
R2 Adj.	0.003	0.004	0.006	0.007
AIC	81966.1	81934.6	84599.2	84580.0
BIC	82187.1	82013.0	84820.2	84694.0
Log.Lik.	-40952.043	-40956.281	-42268.609	-42273.986
F	1.982	5.455	3.065	5.584
Std. Errors	Robust	Robust	Robust	Robust

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix XVII: Multicollinearity Check - Correlation Plot

A correlation matrix plot of selected variables was used in the regressions.



Appendix XVIII: Variance Inflation Factor Analysis for All Models - OLS

	Model			
Variable	**Model Likely**	**Likely reduced**	**Share**	**Share reduced**
age	2.2		2.5	
bias	5.4	2.1	5.6	4.3
bias:crt_score	4.3		4.4	4.2
concordance_dummy	5.4	5.3	5.3	2.1
concordance_zipcode_dummy	2.2	2.1	2.2	2.1
condition				
crt_score	3.2	2.8	3.4	1.2
crt_score:concordance_dummy	4.9	4.8	4.8	
education_higher_dummy	2.3	2.1	2.4	1.2
facebook_daily_dummy	2.1	1	2.2	1
fake_dummy	23.6	16.9	31	7.7
fake_dummy:age	7.6		17.5	
fake_dummy:bias	2.2	2.2	2.3	
fake_dummy:concordance_dummy	2.7	2.7	2.8	2.7

fake_dummy:concordance_zipcode_dummy	2.4	2.3	2.4	2.2
fake_dummy:crt_score	6.4	6	6.8	
fake_dummy:education_higher_dummy	4	3.6	3.8	
fake_dummy:facebook_daily_dummy	3.5		4.8	
fake_dummy:income_high_dummy	3.4		3.2	2.9
fake_dummy: likely			6	5.7
fake_dummy:pet_centered	2.3		2.4	
fake_dummy:pit_centered	2.4	2.3	2.4	
fake_dummy:polarized_zipcode_dummy	2.3		2.4	
fake_dummy:political_knowledge	10.9	10.2	11	
fake_dummy:prefers_republican	6.1	5.9	5.5	
fake_dummy:race_white_dummy	6		7.2	
fake_dummy:treatment_dummy	3		3.2	2.9
income_high_dummy	2.2		2.2	2.1
likely			2.5	2.5
pet				

pet_centered	2.1	1.1	2.3	1.1
pit				
pit_centered	2.3	2.3	2.4	
polarized_zipcode_dummy	2.1		2.2	1.1
political_knowledge	2.3	2.2	2.4	1.1
prefers_republican	2.2	2.1	2.3	1.1
race_white_dummy	2.2		2.3	1.1
treatment_dummy	2	1	2.2	2.1

Appendix XIX: Variance Inflation Factor Analysis for All Models - Mixed OLS

	Model			
Variable	**Model Likely**	**Likely reduced**	**Share**	**Share reduced**
age	1.4		1.4	1.2
bias	5.2	2	5.4	5.3
bias:crt_score	4.1		4.2	4
concordance_dummy	4.2	2.1	4.1	2.1
concordance_zipcode_dummy	1.9		1.9	
condition				
crt_score	1.7	1.4	1.5	1.4
crt_score:concordance_dummy	3.3		3	
education_higher_dummy	1.4	1.3	1.3	1.2
facebook_daily_dummy	1.3		1.2	1.1
fake_dummy	3.7	3.2	3.3	2.2
fake_dummy:age	1.8		2	
fake_dummy:bias	2	2	2.1	2
fake_dummy:concordance_dummy	2.2	2.1	2.2	2.1
fake_dummy:concordance_zipcode_dummy	2		1.9	
fake_dummy:crt_score	1.8	1.7	1.6	1.3
fake_dummy:education_higher_dummy	1.6	1.4	1.3	
fake_dummy:facebook_daily_dummy	1.4		1.3	

fake_dummy:income_high_dummy	1.5		1.2	
fake_dummy: likely			2.4	
fake_dummy:pet_centered	1.3	1.3	1.2	
fake_dummy:pit_centered	1.4	1.4	1.3	
fake_dummy:polarized_zipcode_dummy	1.3		1.2	
fake_dummy:political_knowledge	2.1	2	1.7	1.5
fake_dummy:prefers_republican	1.7	1.6	1.4	1.3
fake_dummy:race_white_dummy	1.7		1.5	
fake_dummy:treatment_dummy	2.8	2.8	2.7	1.6
income_high_dummy	1.4		1.2	1.1
likely			2.1	1.1
pet				
pet_centered	1.3	1.3	1.2	1.1
pit				
pit_centered	1.4	1.4	1.3	1.2
polarized_zipcode_dummy	1.3		1.1	1
political_knowledge	1.4	1.4	1.3	1.3
prefers_republican	1.3	1.3	1.2	1.2
race_white_dummy	1.4		1.3	
treatment_dummy	1.8	1.8	1.7	

Appendix XX: Variance Inflation Factor Analysis for All Models - Logistic Regression

	Model			
Variable	**Model Likely**	**Likely reduced**	**Share**	**Share reduced**
age	2		2.3	1.2
bias	1.8	1.8	1.9	1.1
bias:crt_score				
concordance_dummy	2	2	1.9	1.1
concordance_zipcode_dummy	1.9	1.9	1.9	
condition				
crt_score	2.2	2.1	2.3	1.1
crt_score:concordance_dummy				
education_higher_dummy	2.1	2	2.1	1.1
facebook_daily_dummy	1.9	1	2	
fake_dummy	23	16.5	33.3	2.1
fake_dummy:age	7.2		18.1	
fake_dummy:bias	1.8	1.8	2	
fake_dummy:concordance_dummy	2.7	2.7	2.9	
fake_dummy:concordance_zipcode_dummy	2.1	2.1	2.1	
fake_dummy:crt_score	5.9	5.6	5.5	
fake_dummy:education_higher_dummy	3.6	3.3	3.1	

fake_dummy:facebook_daily_dummy	3.4		4.6	
fake_dummy:income_high_dummy	3.2		2.8	
fake_dummy:likely			13.3	
fake_dummy:pet_centered	2.2		2.2	
fake_dummy:pit_centered	2.1	2	2	
fake_dummy:polarized_zipcode_dummy	2.1	2.1	1.9	
fake_dummy:political_knowledge	10.2	9.5	9	
fake_dummy:prefers_republican	6.2	5.9	5.9	
fake_dummy:race_white_dummy	5.8		6.6	
fake_dummy:treatment_dummy	2.9		2.9	2.7
income_high_dummy	2		1.9	
likely			2.4	1.1
pet				
pet_centered	1.9	1.1	2	
pit				
pit_centered	2.1	2	2	
polarized_zipcode_dummy	1.9	1.9	1.8	1
political_knowledge	2.1	2	2.2	1.2
prefers_republican	2	1.9	2.1	1.1
race_white_dummy	2	1	2.1	1.2
treatment_dummy	1.9	1	2	1.8

Appendix XXI: Variance Inflation Factor Analysis for All Models - Discernment

	Model	
Variable	**Truth Discernment**	**Sharing Discernment**
age	1.2	1.1
condition	1.1	1
crt_score	1.3	1.2
crt_score:concordance_dummy		
education_higher_dummy	1.2	1.2
facebook_daily_dummy	1.1	1
income_high_dummy	1.1	1.1
likely		
pet	1.1	1.1
pet_centered		
pit	1.2	1.2
pit_centered		
polarized_zipcode_dummy	1	1
political_knowledge	1.2	1.2
prefers_republican	1.1	1.1
race_white_dummy	1.2	1.1

Appendix XXII: Breusch-Pagan Results

data: m_likely
BP = 449.66, df = 33, p-value < 2.2e-16

data: m_likely_short
BP = 411.05, df = 21, p-value < 2.2e-16

data: m_share
BP = 559.48, df = 35, p-value < 2.2e-16

data: m_share_short
BP = 552.99, df = 21, p-value < 2.2e-16

data: m_likely_logit
BP = 663.7, df = 31, p-value < 2.2e-16

data: m_likely_short_logit
BP = 653.82, df = 23, p-value < 2.2e-16

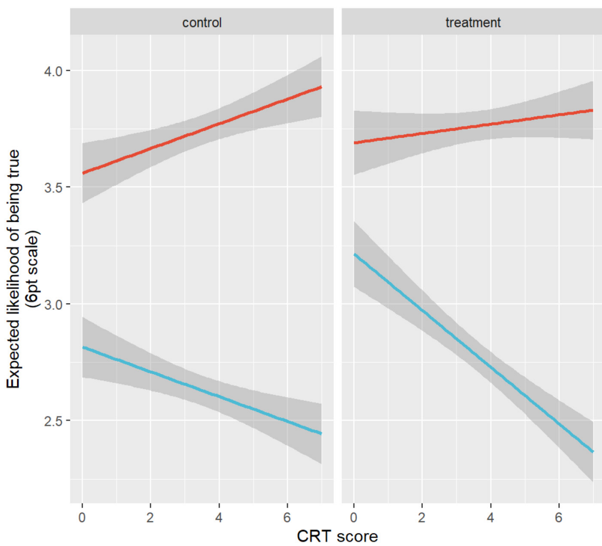
data: m_share_logit
BP = 671.13, df = 33, p-value < 2.2e-16

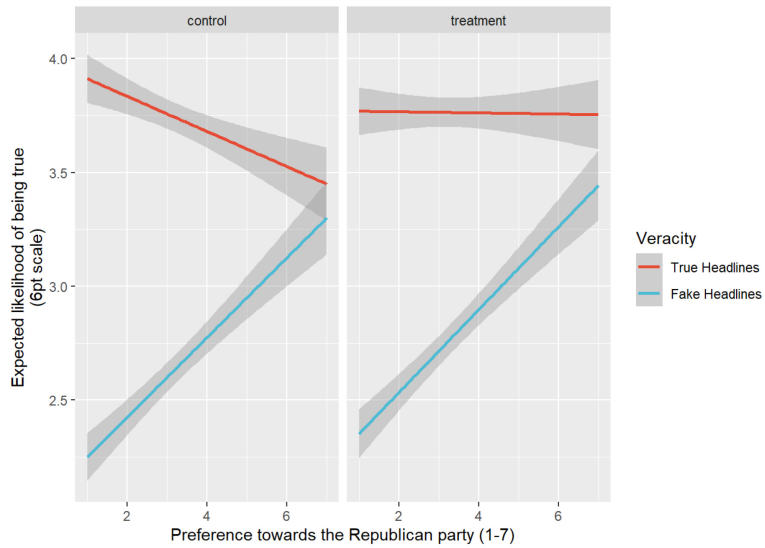
data: m_share_short_logit
BP = 645.92, df = 13, p-value < 2.2e-16

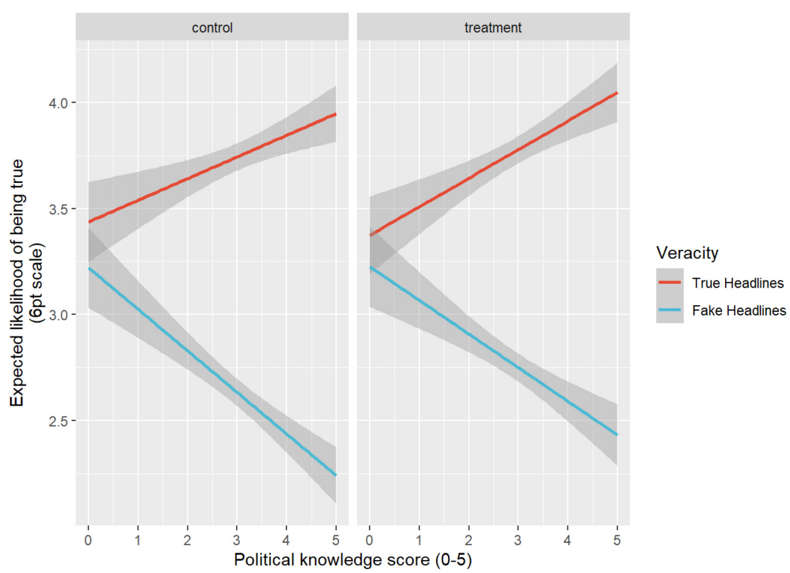
data: m_sharing_discernment
BP = 6.4397, df = 12, p-value = 0.8923

data: m_truth_discernment
BP = 9.4968, df = 12, p-value = 0.66

Appendix XXIII: Treatment Versus Control. Comparison of Effects of Variables on Belief in Fake News across True and False Headlines[6]

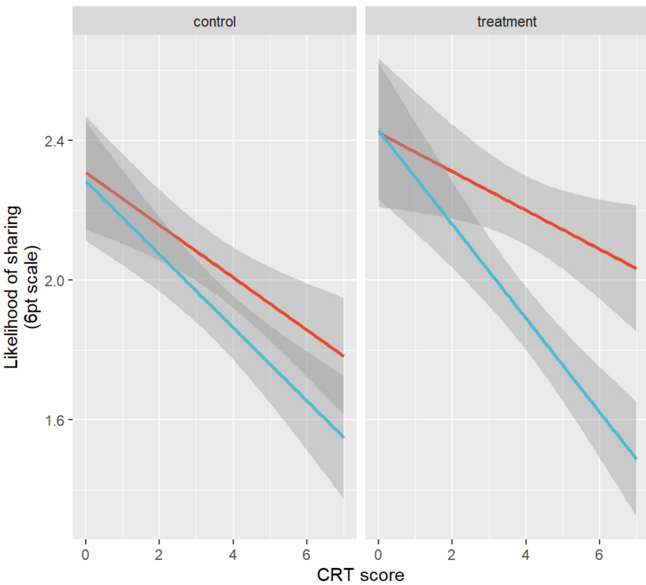


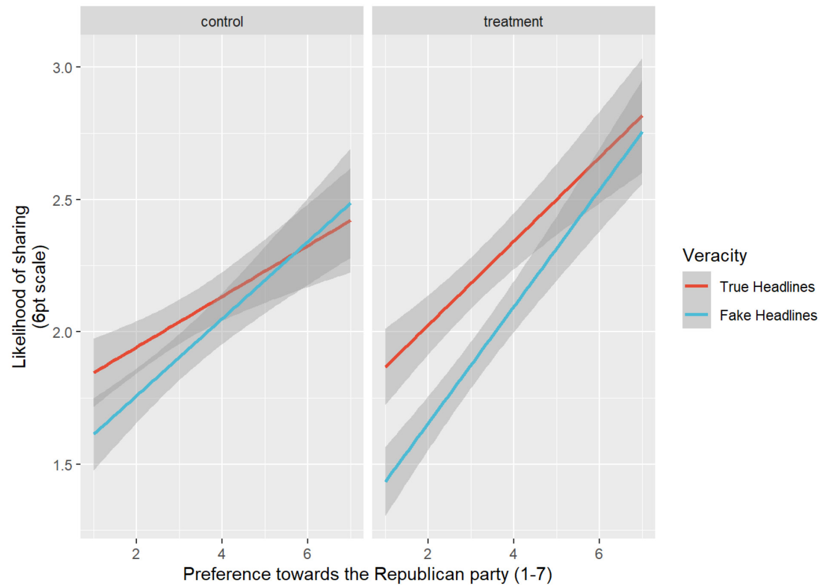


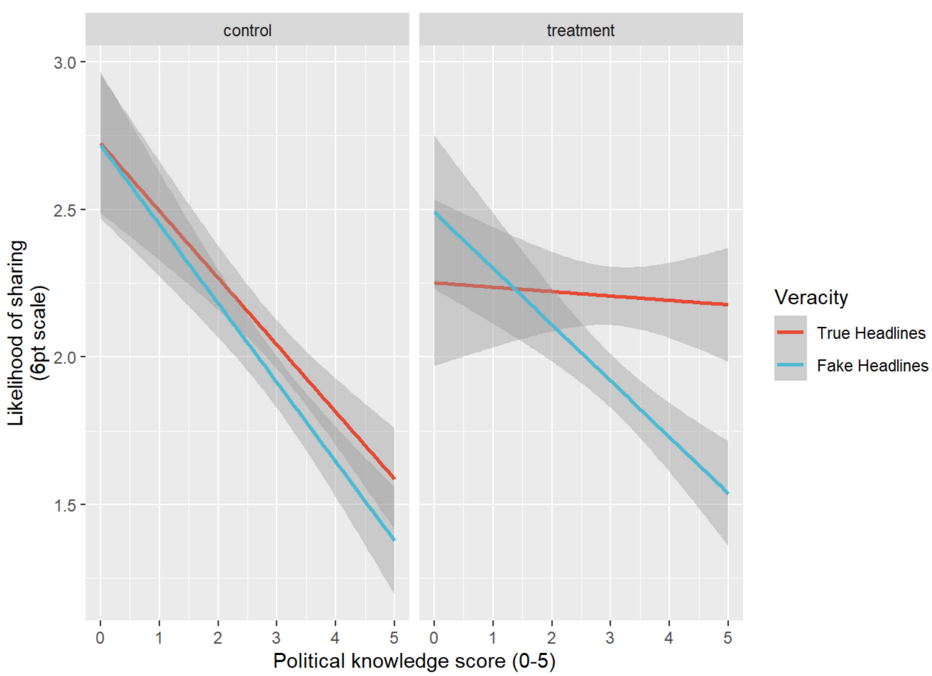


[6] The plots in Appendix XXIII and XXIV represent smoothed conditional means of key dependent variables at various levels of some explanatory variables. Conditional means represent predictions of bivariate OLS regressions. 95% confidence bands around the regression lines were used to visualize uncertainty around the estimated conditional means.

Appendix XXIV: Treatment Versus Control. Comparison of Effects of Variables on Sharing Inclination Across True and False Headlines







Appendix XXV: Summary Table for the Hypotheses

Hypothesis	**Verdict**	**Comment**
1.1	No support	No evidence of treatment leading to better truth discernment.
1.2	*Partial support*	Sharing discernment improved when exposed to treatment, but the effect is driven by increased confidence in true news.
2	*Partial support*	Confirmation bias decreases the likelihood of truth discernment but does not increase the inclination to share fake news.
3	*Partial support*	In line with dual processing theory, more effortful thinking leads to better truth discernment. The effect, however, is not amplified via treatment.
3.a	*Partial support*	Preference for intuitive thinking increases belief in fake news, but preference for effortful thinking does not decrease belief in fake news. Neither interacts with the treatment.
3.b	*Partial support*	Higher CRT scores lead to decreased belief in fake news, but the effect is not amplified via accuracy prompt.
3.c	*Partial support*	Better political knowledge leads to decreased belief in fake news. The effect is not amplified via the treatment.
3.d	No support	Contrary to expectations, political concordance increases belief in fake news.
4.a	*Partial support*	County-level concordance decreases the belief in fake news, but we find no support for impact on sharing. The effect is not amplified via the treatment.
4.b	*Partial support*	Higher education does not impact sharing intentions for fake news but decreases belief in fake news. The effect is not amplified via the treatment.

4.c	No support	Neither preference for intuitive thinking nor preference for effortful thinking impact sharing intentions for fake news. The effect is not amplified via the treatment.
4.d	No support	A higher CRT score does not lead to decreased sharing of fake news. The effect is not amplified via the treatment.
4.e	No support	Better political knowledge does not decrease sharing of fake news. The effect is not amplified via the treatment.
4.f	No support	No link between an individual's concordance with the political slant of fake news and sharing intentions. The effect is not amplified via the treatment.
4.g	No support	No link between living in politically polarized states and sharing intentions or belief in fake news.