AIS Transactions on Human-Computer Interaction

Volume 14 | Issue 2 Article 6

6-2022

The Reasoning behind Fake News Assessments: A Linguistic Analysis

Lydia Manikonda
Rensselaer Polytechnic Institute, manikl@rpi.edu

Dorit Nevo
Rensselaer Polytechnic Institute, Troy, NY, United States., nevod@rpi.edu

Benjamin D. Horne
The University of Tennessee Knoxville, bhorne6@utk.edu

Clare Arrington
Rensselaer Polytechnic Institute, arrinj@rpi.edu

Sibel Adali
Rensselaer Polytechnic Institute, adalis@rpi.edu

Recommended Citation

Manikonda, L., Nevo, D., Horne, B. D., Arrington, C., & Adali, S. (2022). The Reasoning behind Fake News Assessments: A Linguistic Analysis. *AIS Transactions on Human-Computer Interaction, 14*(2), 230-253.
https://doi.org/10.17705/1thci.0010168
DOI: 10.17705/1thci.0010168

Transactions on Human-Computer Interaction

Volume 14 Issue 2

6-2022

The Reasoning behind Fake News Assessments: A Linguistic Analysis

Lydia Manikonda

Lally School of Management, Rensselaer Polytechnic Institute, maniklrpi.edu

Dorit Nevo

Lally School of Management, Rensselaer Polytechnic Institute, nevod@rpi.edu

Benjamin D. Horne

School of Information Sciences, The University of Tennessee Knoxville, bhorne6@utk.edu

Clare Arrington

Department of Computer Science, Rensselaer Polytechnic Institute, arrinj@rpi.edu

Sibel Adali

Department of Computer Science, Rensselaer Polytechnic Institute, adalis@rpi.edu

Recommended Citation

Manikonda, L., Nevo, D., Horne, B. D., Arrington, C., & Adali, S. (2022). The reasoning behind fake news assessments: A linguistic analysis. *AIS Transactions on Human-Computer Interaction*, *14*(2), pp. 230-253.

DOI: 10.17705/1thci.0010168

Available at http://aisel.aisnet.org/thci/vol14/iss2/6

Transactions on Human-Computer Interaction

Research Paper DOI: 10.17705/1thci.00168 ISSN: 1944-3900

The Reasoning behind Fake News Assessments: A Linguistic Analysis

Lydia Manikonda
Lally School of Management, Rensselaer Polytechnic Institute, Troy, New York
manikl@rpi.edu

Dorit Nevo
Lally School of Management,
Rensselaer Polytechnic Institute

Benjamin D. Horne
School of Information Sciences,
College of Communication and Information,
The University of Tennessee Knoxville

Clare Arrington
Department of Computer Science,
Rensselaer Polytechnic Institute

Sibel Adali
Department of Computer Science,
Rensselaer Polytechnic Institute

Abstract:

This paper investigates how individuals reason about the authenticity of the news content they consume. While researchers have conducted much work on fake news detection and prevention, we know relatively less about how news readers reason about the content that they read. Using data collected through Amazon Mechanical Turk, we analyzed over 1,000 justifications that news readers provided about why they believe (or fail to believe) given news articles. We included both fake and credible articles in our analyses and examined the novelty of the news topic as a possible contingency factor that differentiated the reasoning provided. Based on our psycholinguistic analyses, we found that news readers employ both cognitive and motivated reasoning and that agreement with the ground truth impacts the reasoning more than a news topic's novelty. Our insights contribute to the literature on news consumption and reasoning in the context of evaluating fake news. Furthermore, this knowledge contribution has implications in areas such as news veracity intervention and tool design. Lastly, we offer a methodological contribution via using linguistic analysis in a novel way to assess the quality of open-ended survey questions.

Keywords: Fake News, Reasoning, Natural Language Processing, Crowdsourcing.

Shuk Ying (Susanna) Ho was the accepting senior editor for this paper.

1 Introduction

Fake news refers to news stories that contain fabricated content and that lack verifiable facts and/or sources (Tandoc et al., 2018; Wardle, 2017) to support them. Fake news has existed for a while but has gained more attention in recent years due to its negative impact on society and the rise of social media and alternate news distribution channels. Since these channels can publish content faster compared to traditional news media, many actors have a strong interest in preventing fake news from spreading. As Zhou et al. (2019) noted, fake news has weakened the public trust in governments and impacted significant events such as the 2016 U.S. presidential election (Allcott & Gentzkow, 2017). Fake news also impacts non-political domains such as stock markets and trade deals (Clarke et al., 2020; Kogan et al., 2019). Research on social media shows that, on platforms such as Twitter, fake news spreads far more rapidly than genuine or truthful news (Vosoughi et al., 2018). Motives across these alternate information sources range from profit to political and inevitably lead to increasingly available misleading and false information.

Multiple research efforts have focused on combating fake news through detecting false information and mitigating its spread. Some existing approaches have relied on fact checking (manual or automated, such as https://fullfact.org/), style (using different visual, semantic, and syntactic post features) (e.g., Wang, 2017; Ruchansky et al., 2017), and propagation (such as how news disseminates) (e.g., Braun & Eklund, 2019; Gruppi et al., 2021b; Starbird et al., 2018). Researchers have paid less attention to the other side of the equation; that is, how the people who read (and often accept) fake news interpret fake news and how they might respond to the interventions mentioned above. Nevertheless, we need to understand this aspect for two main reasons.

First, it entails important user information on how to design more effective interventions that target the heart of fake news beliefs. Multiple studies have found resistance to false information interventions and "after-the-fact" interventions (Roozenbeek & van der Linden, 2019). For example, experiments with after-the-fact interventions have resulted in interventions that have found mixed initial success (Horne et al., 2019) or, even when initially successful, interventions that become ineffective in the long term (often called the continued influence effect) (Nyhan & Reifler, 2010). Hence, current intervention designs likely have inadequacies overall (Chan et al., 2017). By adding to our knowledge about how news customers reason, one might design these after-the-fact interventions more effectively.

Second, understanding people's reasoning can provide insight on how and why they share fake news and how it propagates across societal layers. Research on fake news spread shows that false information can propagate across platforms in two ways: 1) through malicious, intentional disinformation campaigns that often involve political elites and automated bots (Marwick & Lewis, 2017; Broniatowski et al., 2018) and 2) through real news consumers sharing false content for various reasons (e.g., trust the information source and social media fatigue) (Talwar et al. 2019). Studies that elicit these rationales have focused on reasons for sharing rather than reasons for believing, which often (though not always) (Lazer et al., 2021) constitutes a precursor to sharing. Our study provides insights to the latter.

Previous work that has focused on the reasoning for believing fake news has looked at different information-processing paths. For example, Kahan (2013) examined how heuristic-driven reasoning, ideologically motivated reasoning, and cognitive reasoning impact people's willingness to accept views different from their own in the political context. He concluded that ideological motivation plays a role in information processing in such a context, a similar conclusion to what Thaler (2019) made. In contrast, Pennycook and Rand (2019) contrasted news readers' use of classical analytical reasoning (where belief in fake news has roots in an inability to elaborate on the news content) versus ideologically motivated reasoning (where belief in fake news has roots in prior (usually partisan) beliefs and biases). They found that people tend to believe fake news due to the former (laziness) and not the latter (partisanship) (Pennycook & Rand, 2019, 2021). In related literature, Moravec et al. (2018) found that the existence of fake news interventions (such as fake news flags) can lead readers to elaborate and reason about the actual news content. However, when such elaboration occurs, motivated reasoning (in the form of confirmation bias) may take over and override the advice that the fake news flag provides.

The above work demonstrates that both low cognitive (analytical) reasoning and ideologically motivated reasoning may explain why people believe fake news. In this paper, we extend this line of work but add more nuance to it by considering news topic novelty. Novelty refers to the extent to which incoming information is similar to prior knowledge (Karkali et al., 2013; Vosoughi et al., 2018) and can represent news readers' familiarity with the subject matter. Research shows that novel news situations can become a hot bed for conspiracy theories and false content (Starbird et al., 2014); thus, we need to understand them better. We believe that news topic novelty might explain some discrepant findings about the role that analytical and ideologically motivated reasoning plays in fake news. Further, novel news situation can lead to information overload in which case people act irrationally and are more vulnerable to misleading information, which makes it difficult for them to differentiate fake and trustworthy news (Shao et al., 2017).

Based on the above information, we focus on understanding differences in how people reason about content in common (or everyday) versus novel (or emerging) news articles that they read and whether they employ different reasoning for fake versus credible news under each condition. Specifically, we address this question through an exploratory characterization of linguistic cues that news readers express when they evaluate news articles. To this end, we asked individuals to read and evaluate news articles, some false some not, and to comment on the extent to which they believed (or did not believed) the article and why. We then used their justifications to investigate how they reasoned about the articles using investigative linguistic cues such as grammatical tones, sentiment polarities, and emotions. We conducted our examination in two specific news contexts: 1) one that involved common news (specifically, climate change and general vaccinations before the coronavirus disease of 2019 (COVID-19) pandemic) and 2) one that involved novel news (specifically, at the beginning of the COVID-19 pandemic).

Our work offers several contributions. First, we identify the reasoning processes that news readers employ under different settings. We show that news readers use both cognitive and motivated reasoning but the reasoning processes differ for novel versus common news topics. Further, we found that respondents' agreement with the ground truth plays a role in the reasoning they provide. Second, from a methodological perspective, we employ several innovative natural language processing approaches to learn about individuals' behavior through their written responses. Third, we offer some advice to designers about designing technology interventions to minimize the extent to which fake news spreads while using these language processing approaches. We link our findings to the existing literature on motivated reasoning in the context of evaluating fake news articles and we add the use of social reasoning, when readers follow the opinions of significant others. Our contributions also include a quantitative approach to incorporating machine learning techniques that one could extend to a large-scale data analysis compared to traditional surveys and questionnaires.

2 Literature Review

2.1 Current Approaches to Detect and Mitigate Fake News

Researchers have extensively studied fake news detection in recent years, mostly with help from automated tools. We can broadly categorize studies on fake news detection into studies that use claim-level methods (e.g., Ciampaglia et al., 2015; Hassan et al., 2017; Popat et al., 2016), article-level methods (e.g., Potthast et al., 2017; Horne et al., 2018), and source-level methods (e.g., Horne et al., 2019; Baly et al., 2018; Starbird et al., 2018). The features that these methods employ to detect fake news closely mirror the tactics that have been used to lure people into reading and engaging with fake news. For example, studies have shown that content with moral-emotional language spreads much further than content without it by a factor of 20 percent per additional moral-emotional word (Brady et al., 2017); subsequently, the presence of moral-emotional language can indicate low content veracity. Similarly, false news articles often have longer headlines, which contain complete claims, which can allow readers to form false beliefs by reading news headlines alone (Horne et al., 2017). A related phenomenon called "clickbait" instead adopts a different headline structure that entices the consumer to click on the article link for more information in order to garner advertising revenue. While clickbait articles do not constitute misinformation themselves, many low veracity sources may rely on them (Shu et al., 2017; Chakraborty et al., 2016). Hence, the length and structure of the headline is another feature that one can use to assess news' veracity.

Researchers have also intensely studied methods to intervene in misinformation consumption (Lewandowsky et al., 2012). One of the most effective methods involves increasing readers' news and media literacy (Craft et al., 2017; Guess et al., 2020; Jones-Jang et al., 2021). Despite variations in approach, researchers usually focus on providing readers with skills to assess information's accuracy (Maksl et al., 2017) either through in-depth information processing or by invoking helpful heuristics (Horne et al., 2020). Heuristics constitute useful rules of thumb for forming a judgment (Colman, 2015)—mental shortcuts to ease the cognitive load on decision makers (Myers & Dewall, 2018). Another method called prebunking provides corrections to misinformation that one can anticipate in advance. Brashier and Schacter (2020) found a related approach, debunking (providing corrections to misinformation), more effective than prebunking (Brashier & Schacter, 2020). However, a correction may leave an inconsistency in one's memory or may cause a person to remember incorrect information better than the correction (Lewandowsky et al., 2012). Correcting misinformation that distorts the truth is especially difficult and requires significant mental effort from readers (Lewandowsky et al., 2017) and trust in the entities providing the correction. Another popular approach involves providing simple labels to alert users about problematic content or sources rather than detailed fact-checking statements (Horne et al., 2019, 2020; Sathe et al., 2020; Spezzano et al., 2021). Prior work has shown that AI-assisted labels for information accuracy are most helpful with explanations, but such feedback is not uniformly helpful for all users (Horne et al., 2019). Finally, other efforts have also concentrated on filtering or demoting problematic content or highly influential users as a problem's scale or urgency may necessitate it, especially in networked settings in the presence of bots or other information-manipulation tactics (Bak-Coleman et al., 2021).

While these interventions constitute important tools in the misinformation toolbox to help readers better calibrate their trust for the information that they encounter, they do not shed light on how readers consume and interpret fake news and, consequently, how they might respond to the different interventions.

2.2 The News Readers' Perspective

From the news consumer side, studies propose that people fail to make a distinction between true and false news headlines when they rely on their intuition and do not deliberate (Bago et al., 2020). Research suggests that people often accept fake news because they rely on peripheral cues. For instance, Kiely and Robertson (2016) found that individuals rely on elites' opinions to form opinions about a specific issue, a tactic that fake news often employs. Sundar et al. (2003) conducted an experiment that looked at political campaign websites design and found some evidence that users evaluate misinformation not just based on the content itself but also based on other contextual aspects such as its "look" and features. Fazio (2020) showed that mere exposure to a fake news headline would increase belief in that headline later. Related work from Pennycook and Rand (2019, 2021) suggests that lazy thinking and low cognitive reasoning drives susceptibility to fake news. In a similar study, Pennycook et al. (2015) found that people who more deeply reflected on fake news statements were more likely to reject them as true, but individuals with less critical reflection capacities accepted them more.

In contrast, other work shows that deliberation might actually cause people to fall for fake news under specific contexts (e.g., Kahan, 2013; Moravec et al., 2018). Bago et al. (2020) and Pennycook and Rand (2020) call this phenomenon "identity-protective cognition" and argue that people use it to rationalize content that concurs with their political ideology. These arguments also concur with the broader literature on confirmation bias, which means seeking or interpreting evidence in ways that agree with existing beliefs and expectations (Nickerson, 1998), especially when people have confidence in their knowledge and its correctness (Koriat et al., 1980). Due to confirmation bias, people consider new information that supports their pre-existing views and ignore information that challenges those views (Moravec et al., 2018; Lewandowsky et al., 2012). In some cases, even when one explicitly tells people that information is false, they remain prone to accepting it, which might cause people to develop even more ingrained beliefs (Shu et al., 2017; Vosoughi et al., 2018). Indeed, recent research found that social media users were more likely to believe, like, and share articles that matched their point of view (Kim et al., 2018), which, in turn, can propagate a broad belief in fake news (Vicario et al., 2019).

An interesting gap in the above literature concerns the contingencies under which cognitive reasoning takes over ideologically motivated reasoning. With prior findings supporting both information processing approaches, we need to study when one might take over the other. Kahan (2013) and Thaler (2019) proposed one such contingency as being political affiliation. Kahan (2013) showed that conservatives exhibit stronger ideologically motivated reasoning. Meanwhile, Brashier and Schacter (2020) found that age might also play a factor. Specifically, they found that older adults might be especially susceptible to fake news, but they argued that, although older adults forget where they learned information, fluency remains intact, and knowledge accumulated across decades helps them evaluate claims. Related to this finding, Lazer et al. (2021) paradoxically found that, while older people shared more misinformation on social media than younger people, older people had a lower inclination to believe it.

In this paper, we explore a different contingency: a news topic's novelty. To understand such novelty, we borrow from the literature on information novelty. Several factors characterize information novelty: its diversity or variation from previous information, non-redundant information volume, and its uniqueness (Aral & Dhillon, 2022). Itti and Baldi (2009) provide a Bayesian definition for surprising information and measure it by the extent to which it changes the receiver's posterior beliefs. Novelty attracts human attention and encourages information sharing because it updates how one understands the world (Vosoughi et al., 2018).

Thaler (2019) argues that novelty induces motivated reasoning. At the same time, novelty often introduces uncertainty, which researchers have shown to increase reliance on systematic cognitive reasoning (Tiedens & Linton, 2001). Hence, the effect that news topic novelty has on the reasoning that news readers employ can benefit from future research (Horne et al., 2020).

To better understand the effect that news topic novelty has on news consumption and veracity judgments, we conducted an exploratory study in which we applied linguistic analysis to justifications that news readers provided about the degree to which they accepted false and credible news articles. Although we focused on specific conditions (false vs. true and novel vs. common), we used an open-ended dependent variable that focused on the reasoning that respondents provided. For this reason, we opted for an exploratory study as we describe in Section 3.

3 Empirical Study

In our empirical study, we tasked respondents with reading one randomly assigned article from our pre-defined article set that contained both false and credible news and both common and novel news topics. After the respondents read the article, we asked them whether or not they believed it and why. Our main data analysis (i.e., our linguistic analysis) focused on the reasoning that the respondents provided. Linguistic analysis has become an increasingly popular approach in social science research. The analysis often uses textual features, which include several lexicon, syntax, discourse, and semantics aspects. As linguists define it, a lexicon refers to an entity's (here, written assessments of news articles) vocabulary. At the lexicon level, the main tasks involve measuring different frequency statistics, which one can do with approaches such as a bag of words model or a unigram-bigram model (Jin et al., 2014; Zhou et al., 2020). For syntax, one explores the frequency with which different parts of speech appear in text to measure low-level syntax operations and probabilistic context-free grammar parse trees for deep syntax operations (Conroy et al., 2015; Feng et al., 2012; Pérez-Rosas et al., 2017). At the semantic level, one can assign different frequencies to lexicons or phrases that fall into different psycho-linguistic categories that one can leverage with Linguistic Inquiry and Word Count (LIWC) (Bond et al., 2017; Jordan et al., 2018).

We conducted the study on Amazon's Mechanical Turk (AMT), a commonly used platform for data collection. Although AMT might not always be the most fitting platform, studies show that one can suitably use it in settings that involve open-ended subjective assessments and that it generates quality data (Bates & Lanza, 2013). Given that cultural differences may impact our finding, especially given our news context, we limited responses to workers from only the United States. To ensure quality responses, we also limited the human intelligence task (HIT) approval rate (a common quality measure on AMT) to 99 percent. We collected the data in several rounds beginning in November, 2019 and ending in April, 2020. We received 1,212 usable responses in total.

3.1 Selection of Articles and Ground Truth

Table 1 presents information on the 19 articles that we used in this study (their source, headline, and ground truth). These articles spanned three broad topics: climate change, vaccinations (before COVID-19), and COVID-19. We classified the first two topics as common news since they have been around for quite some time. We classified the third as novel. Recall that we collected data between November, 2019, and April, 2020, just as the COVID-19 pandemic began.

The last column in Table 1 shows the ground truth that we assigned to each article. In selecting articles that varied in the ground truth, we used third-party organizations, similar to what previous studies have done (Nørregaard et al., 2019; Gruppi et al., 2021a). Since our study involved participants reading individual news articles, our ground truth also needed to be at the article level. To this end, we used a three-step process in which we 1) found sources that Media Bias/Fact Check (MBFC) labeled as reliable and unreliable, 2) found topic specific articles from those sources (climate change, vaccination, and COVID-19), 3) selected articles a third-party journalistic organization, such as Snopes, PolitiFact, FactCheck.org, Washington Post Fact Check, or AP Fact Check, fact-checked.

Table 1. News Articles that We Randomly Assigned to Respondents

Common news		
Source	**Title**	**Ground Truth**
Natural News	Climate change HOAX has literally convinced a member of Congress that “the world is going to end in 12 years”	Not credible
Freedom Bunker	Fight illness with this ancient immune booster	Not credible
Jew World Order	Greenpeace founder: global warming is a hoax pushed by corrupt scientists “hooked on government grants”	Not credible
The Gateway Pundit	NOAA ruins assertions by unhinged democrats that global warming has caused increase in hurricane activity	Not credible
Natural News	World Health Organization declares anti-vax movement to be a top “global health threat” just like the climate change hoax… the vaccine deep state grows desperate	Not credible
BBC	“Completely avoidable” measles outbreak hits 25-year high in US	Credible
NPR	Climate change was the engine that powered Hurricane Maria’s devastating rains	Credible
Chicago-Sun Times	Kentucky governor exposed his kids to chickenpox instead of getting vaccine	Credible
NPR	New U.S. measles cases break 25-year-old record, health officials say	Credible
Fortune	U.S. carbon emissions soared in 2018. Here’s why	Credible
Novel News		
Source	**Title**	**Ground truth**
The New York Times	Open windows. Don’t share food. Here’s the government’s coronavirus advice.	Credible
Reuters	World faces chronic shortage of coronavirus protective equipment: WHO	Credible
The Guardian	Can a face mask stop coronavirus? Covid-19 facts checked	Credible
Breitbart	Hillary Clinton falsely claims to Jimmy Fallon that Trump called coronavirus outbreak “a hoax”	Not credible
Natural News	Vitamin C infusions being studied in China as possible treatment for coronavirus-related pneumonia	Not credible
Natural News	Spirulina found to boost the body’s type 1 interferon response to fight RNA viral infections “including coronavirus”, new science finds	Not credible
The Russophile	coronavirus hoax: Fake virus pandemic fabricated to cover-up global outbreak of 5g syndrome	Not credible
The Russophile	Coronavirus special report: worldwide outbreaks of 5G syndrome and 5G flu driving pandemic	Not credible
The Liberty Daily	Coronavirus: Chinese espionage behind Wuhan bioweapon?	Not credible

3.2 Respondents

Table 2 overviews how many responses we received for each article type.

Table 2. Summary Statistics of the data

Article topic	Number of responses	Length (avg. words per response)
Climate Change (credible)	194	23.58
Climate Change (false)	183	23.09
Vaccination (credible)	198	21.07
Vaccination (false)	191	21.85
COVID-19 (credible)	142	22.05
COVID-19 (false)	304	28

The majority of respondents (67%) were between 25 to 44 years old, and nearly 60 percent were male. Over 75 percent had either an associate degree or a four-year college degree. Respondents indicated that they primarily consumed news on news websites (42%), social media (28%) and television (25%). Nearly 90 percent of respondents indicated they consumed news daily or multiple times a day. Finally, 60 percent indicated they sometimes share news on social media compared with 30 percent who indicated they never shared news on social media and about six percent who always shared.

3.3 Data-quality Assessment

When one uses Amazon Mechanical Turk to collect data, one risks involving respondents that generate low-quality data (Buhrmester et al., 2016; Kennedy et al., 2020). Even though we chose only respondents with high approval ratings, we needed to first perform an initial data analysis to ensure that we collected reliable data that we could use for further analysis. In order to assess the data's reliability, we evaluated it based on two aspects: 1) the nature of responses at a high level using n-gram analysis (Banko & Vanderwende, 2004; Yannakoudakis et al., 1990) and 2) the similarity of neural embeddings (Devlin et al., 2018) between the credible and false articles, which measures the linguistic variations expressed via responses (Carlebach et al., 2020; Vakulenko et al., 2017).

3.3.1 N-gram Analysis

Researchers have used n-grams extensively to understand the nature of free text (Abidin et al., 2017; Banko & Vanderwende, 2004; Wressnegger et al., 2013; Zak et al., 2017). An n-gram refers to a contiguous sequence of *n* elements from a given text or speech sample. These elements can be letters, words, or base pairs according to the task at hand. We extracted n-grams from the corresponding news article categories using the readers' evaluations. Due to n-grams' contiguous nature, we could identify important aspects in the text that readers highlighted as they provided their reasoning for believing a specific article.

Table 2. Article Type and Its Corresponding Tri-, Bi-, and Uni-grams

Article type	Tri-grams	Bi-grams	Uni-grams
Climate change (credible)	believe-climate-change, climate-change-real, climate-change-happening, article-well-written, trust-source-information	climate-change, believe-climate, makes-sense, change-real, carbon-emissions, news-source	climate, change, believe, article, information, seems, source, think, trust
Climate change (false)	climate-change-real, climate-change-hoax, global-warming-real, believe-climate-change, smart-ai-system	global-warming, climate-change, ai-system, change-real, seems-like, change-hoax	believe, article, biased, climate, global, warming, seems, change, source
Vaccination (credible)	sounds-like-lot, measles-cases-us, sun-times-reputable, believe-people-get	news-source, seems-like, chicken-pox, measles-cases	believe, news, article, people, seems, measles, would, source, heard
Vaccination (false)	trusted-news-source, smart-ai-system, source-freedom-bunker, seems-like-article, trying-sell-product	seems-like, anti-vax, news-article, believe-vaccines, flu-shots, reads-like	article, vaccines, like, people, believe, seems, information, know, anti, claims
COVID-19 (credible)	trusted-news-source, smart-ai-system, believe-news-article, kills-elderly-younger	news-source, believe-information, seems-like, ai-system, news-article, trusted-news, makes-sense	article, information, news, seems, believe, source, know, sense
COVID-19 (false)	smart-ai-system, trusted-news-source, already-proven-5g, ai-system-says, article-inaccurate-unreliable	ai-system, news-source, smart-ai, conspiracy-theory, news-article, never-heard	article, believe, source, news, 5g, ai, information, seems

The uni-grams in Table 3 show that, at an aggregated level across all the six different categories, news readers used vocabulary such as “seem”, “think”, “trust”, “believe”. This terminology reflects respondents' focus on the task to provide their reasoning about the news articles and indicates that they provided relevant responses. The bi- and tri-grams highlight the aspects of how individuals used their preexisting beliefs and experiences to the content that in the article along with their opinions about the article source itself. Overall, the top n-grams focused on articles' content (e.g., “global-warming”, “chicken-pox”, “news-source”) and reasoning that the respondents provided through using vocabulary such as “seems like”, “makes sense”, “common sense”, which again demonstrates relevance to the study’s task. From this analysis, we concluded that we obtained suitable data to examine differences in how consumers reason about news articles under different contexts.

3.3.2 Neural Embeddings

Before we delved deeper into the empirical analysis, we investigated if the responses we collected showed any high-level distinctions between justifications for credible and false articles. Current literature on natural language processing (Řehůřek & Sojka, 2011; Devlin et al., 2018) shows that neural embeddings can capture both the semantic and syntactic characteristics of sentences in order to find any high-level distinctions between groups. One extracts this information by representing each sentence in a given corpus (corpus means the entire dataset that comprises different sentences) as a data point in a vector space. Research has found this approach highly efficient compared to the traditional bag of words (BOW) (Zhang et al., 2010) and Term frequency inverse document frequency (TFIDF) (Aizawa, 2003; Yang & Pederson, 1997) models especially when comparing different documents or, in our context, reasoning statements. This refined approach represents sentences in a vector space and has significant advantages when measuring two document similarities. We primarily used sentence embeddings extracted using BERT (Devlin et al., 2018) to convert all the responses to vector format. We then used these vectors to compute distances between the reasons shared while evaluating credible and false articles in our article set. Since all the vectors in our dataset are expressed in the same space, we used the Euclidean distance metric, which measures the regular distance between two points in space (Hossain & Abufardeh, 2019). Please note that 0 bounds this metric at the lower end, which means the minimal distance between any given two points (or two types of reasoning statements) is 0, which indicates their nearly identical nature.

Figure 1 presents a distance heatmap between each category (Euclidean distance between centroids) in our dataset. The darker the shade of blue in the cell, the higher the Euclidean distance, which means lower similarity (in other words: darker shades indicate more semantic and syntactic dissimilarity between two article categories, while lighter shades indicate more similarity). As one can see, articles with a credible ground truth attracted similar justifications, but those justifications differed to the justifications for the not credible articles. This finding indicates that people employ different kinds of reasoning when interacting with credible versus not credible (false) news.

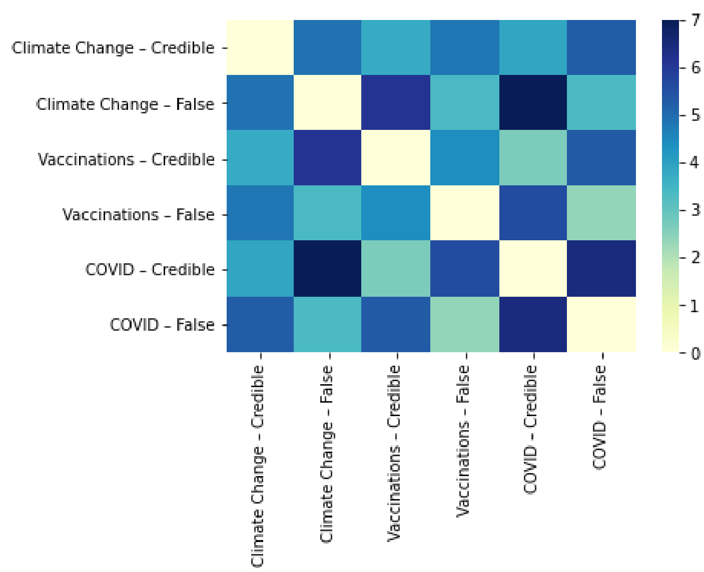

Figure 1. Heatmap of Distances between the Reasoning Statements

The above preliminary analyses highlight the relevance and quality of the data we collected. It shows that respondents did not provide abstract responses but that they related to the task we gave them: to reason about fake and credible news. In Section 4, we analyze the data and look for insights concerning the reasoning that news readers employed.

4 Findings

4.1 Agreement with Ground Truth

Before conducting any further analysis, we first split the data into subsets based on whether the respondents agreed with the article's ground truth. While we did not initially focus on agreement with ground truth, after analyzing the results, we found that it would be beneficial to create subsets based on such agreement. Therefore, we included it in our analysis and elaborate on it in Section 5.

Recall that we asked respondents to indicate the extent to which they believed the article we had them read. We used this scale to gauge agreement with the ground truth. For example, if a person indicated they believed an article that we deemed credible, we marked it as agreement. If they indicated they did not believe an article that we deemed credible, we marked it as disagreement with the ground truth. Of the 1,212 respondents, 534 received credible articles to read, and 75 percent indicated that they believed the article and, thus, the ground truth. The remaining 25 percent either said they did not believe or were unsure about believing the article and, hence, disagreed with the ground truth. Of the 1,212 respondents, 678 respondents received false articles to read, and 69 percent indicated they did not believe it and, thus, agreed with our ground truth. We further examined agreement with the ground truth in the novel news article subset (the COVID-19 articles). Here, the proportion of respondents who agreed with the credible ground truth was 85 percent compared with 68 percent who agreed with the false ground truth. Hence, at the very high level, our study shows that people can identify credible articles slightly better than false articles. We explore the reasoning that the respondents provided next.

4.2 Passive versus Active Tone

We begin our analysis by examining passive versus active tone in the responses, which reflects the two attitudes of detachment and taking responsibility (Yaffe, 2020). Existing literature states that using a passive voice signals detachment and not taking responsibility, whereas active voice highlights that the author takes responsibility (Chan & Maglio, 2020; Yaffe, 2020). Passive forms can fulfill various discursive functions that might easily go unnoticed at a conscious level or introduce interpretational ambiguity (Baratta, 2009; Bolinger, 2014). We leveraged these aspects to highlight how responses in different categories represent the respondents' mindset and confidence in terms of their reasoning. To do so, we used a popular state-of-the-art Python library called Spacy (https://spacy.io/) for natural language processing and classified each response based on whether it predominantly used passive (1) or active (0) voice.

We conducted several tests to examine differences in the responses that used passive versus active voice. First, we note that, regardless of the topic, ground truth, or agreement with the ground truth, responses mostly used active voice. We flagged only 19 percent of the responses as passive. Next, we delve deeper into specific differences between subsets of our data, which we show in Table 4.

Table 4. Tests for Differences in the Proportion of Passive versus Active Voice

Test	Statistical Analysis	Comparisons	$P_{passive1}$	$P_{passive2}$	Sig.
Proportion of passive voice	Z test for proportions	False vs. credible	0.226	0.159	0.0018
		Common vs. novel	0.179	0.227	0.0214
		Agree vs. disagree with ground truth	0.196	0.206	ns

Since our analysis generated a classification variable (0 for active and 1 for passive), it focused on proportion tests. We conducted paired proportions tests on three key data splits:

1) The proportion of passive voice that respondents used to reason about false versus credible articles. We found that they used passive voice proportionally more when reasoning about the false articles (significant at the 0.01 level).

2) The proportion of passive voice that respondents used to reason about common versus novel articles. We found that they used passive voice proportionally more when reasoning about the novel articles (significant at the 0.05 level).

3) The proportion of passive voice that the respondents used to reason about when they agreed versus when they disagreed with the ground truth (we note here that respondents did not know our ground truth judgement: we simply presented them with the articles and later computed the agreement). We found no significant difference for this test.

We found further support for these results from conducting a chi square goodness-of-fit test in which we compared the distribution of the proportion of responses to each of four categories—believe credible articles (agree with ground truth), disbelieve credible articles (disagree with ground truth), believe false articles (disagree), and disbelieve false articles (agree)—with the proportion of passive voice use for the same categories. We did not find significantly different distributions.

Returning to the literature, we know that respondents may leverage passive text when they do not want to take responsibility and when the article makes ambiguous claims that make it difficult for readers to interpret and take a clear stance (Baratta, 2009; Bolinger 2014; Yaffe, 2020). Thus, our findings show a greater lack of clarity when respondents interact with false news and when they interact with novel news topics.

4.3 Psycholinguistic Emotions and News Articles

Next, we use the psycholinguistic lexicon LIWC (Pennebaker et al., 2001) to examine how emotional attributes are associated with different news article categories. LIWC generates 69 preset output variables that include linguistic dimensions (e.g., pronouns, articles, and verbs), psychological processes (e.g., affect and cognition), personal concern categories (e.g., work, home, leisure activities), and more. In the psychological processes, specific attributes are associated with different emotional categories. For example, positive and negative emotions, anger, or sadness are all associated with affective processes; family and friends are associated with social processes; insight, certainty, and differentiation are associated with the cognitive process; and so on. Despite this hierarchy, LIWC scores both lower- and higher-level categories based on the dictionary.

We used LIWC2015 to run the analysis on the 1,212 responses that we received. After we received the scores for each response, we used two subsets and two different hypothesis tests to focus on the more interesting insights. Specifically, as we mention above, we split the data based on agreement with the ground truth. After doing so, we ran two different t-tests for each data subset. The first compared the average LIWC scores for credible versus false articles, and the second compared the average LIWC scores for common versus novel articles. We show the results in Tables 5 to 8, respectively. Due to space limitations, we only show the significant categories in each table.

Table 5. LIWC Scores and Comparison Tests for Reasoning on False versus Credible Articles (Respondents who Agreed with the Ground Truth)

Category	Examples	t	*p*-value	μ_{false}	$\mu_{credible}$
Affective processes	happy, cried	2.222	0.027	4.904	4.060
Negative emotion	hurt, ugly, nasty	5.490	0.000	1.983	0.863
Anger	hate, kill, annoyed	3.765	0.000	0.783	0.271
Family	daughter, dad, aunt	-2.171	0.031	0.000	0.050
Insight	think, know	-5.317	0.000	5.149	7.697
Causation	because, effect	-3.263	0.001	3.669	4.895
Differentiation	hasn't, but, else	5.425	0.000	4.413	2.594
Biological processes	eat, blood, pain	5.232	0.000	1.471	0.575
Health	clinic, flu, pill	6.169	0.000	1.291	0.346
Affiliation	ally, friend, social	-2.008	0.045	0.377	0.629
Achievement	win, success, better	2.489	0.013	0.678	0.398
Power	superior, bully	-1.923	0.055	1.198	1.615
Reward	take, prize, benefit	-2.010	0.045	0.460	0.769

Table 5. LIWC Scores and Comparison Tests for Reasoning on False versus Credible Articles (Respondents who Agreed with the Ground Truth)

Past focus	ago, did, talked	-2.180	0.030	2.988	3.735
Relativity	area, bend, exit	-3.973	0.000	6.676	8.716
Motion	arrive, car, go	-3.888	0.000	0.847	1.646
Time	end, until, season	-3.255	0.001	1.616	2.465
Leisure	cook, chat, movie	-2.166	0.031	0.093	0.247
Nonfluencies	er, hm, umm	-2.639	0.009	0.049	0.260

Table 6. LIWC Scores and Comparison Tests for Reasoning on False versus Credible articles (Respondents who Disagreed with the Ground Truth)

Category	Examples	t	*p*-value	μ_{false}	$\mu_{credible}$
Affective processes	happy, cried	1.730	0.085	5.350	4.247
Positive emotion	love, nice, sweet	2.440	0.015	3.782	2.562
Social Processes	mate, talk, they	-1.885	0.061	4.545	5.993
Health	clinic, flu, pill	4.992	0.000	1.392	0.212
Reward	take, prize, benefit	3.592	0.000	1.183	0.397
Present focus	today, is, now	-2.319	0.021	15.271	17.573
Relativity	area, bend, exit	-1.868	0.063	6.888	8.559
Motion	arrive, car, go	-2.105	0.037	0.830	1.590
Time	end, until, season	-2.649	0.009	1.674	2.801
Leisure	cook, chat, movie	-1.690	0.093	0.031	0.132
Assent	agree, OK, yes	1.839	0.067	0.171	0.039

Table 7. LIWC Scores and Comparison Tests for Reasoning on Common versus Novel Articles (Respondents who Agreed with the Ground Truth)

Category	Examples	t	*p*-value	μ_{common}	μ_{novel}
Anxiety	worried, fearful	-1.770	0.077	0.167	0.319
Social processes	mate, talk, they	3.836	0.000	5.829	4.221
Family	daughter, dad, aunt	2.169	0.031	0.038	0.000
Male references	boy, his, dad	3.919	0.000	0.272	0.000
Causation	because, effect	4.526	0.000	4.848	3.291
Discrepancy	should, would	-3.708	0.000	0.692	1.359
Tentative	maybe, perhaps	-2.945	0.003	4.095	5.293
Certainty	always, never	-1.899	0.058	2.590	3.182
Biological processes	eat, blood, pain	-3.027	0.003	0.829	1.399
Health	clinic, flu, pill	-2.921	0.004	0.653	1.152
Affiliation	ally, friend, social	2.751	0.006	0.617	0.303
Work	job, majors, xerox	1.766	0.078	2.368	1.859
Money	audit, cash, owe	2.206	0.028	0.238	0.083
Religion	altar, church	2.742	0.006	0.189	0.025

Table 8. LIWC Scores and Comparison Tests for Reasoning on Common versus Novel Articles (Respondents who Disagreed with the Ground Truth)

Category	Examples	t	*p*-value	μ_{common}	μ_{novel}
Affective processes	happy, cried	1.739	0.083	5.300	4.210
Feel	feels, touch	2.333	0.020	0.701	0.278
Biological processes	eat, blood, pain	-3.186	0.002	0.701	1.964
Health	clinic, flu, pill	-3.453	0.001	0.524	1.850
Motion	arrive, car, go	2.281	0.023	1.328	0.661
Religion	altar, church	1.755	0.081	0.110	0.000
Informal language	na	2.021	0.044	0.426	0.170

Each table shows the significant (or borderline significant) LIWC categories that emerged from the analysis along with word examples from the LIWC2015 dictionary. The t scores and p-values come from independent samples' t-tests using the raw LIWC scores for each person in each subset and comparing credible versus false article responses (Table 5) and common versus novel article responses (Table 6). The mean values in the tables are the average values for each response in each category and represent the percentage of total words for that category in the written response. The higher the value, the more prominent the category in the response sentence. To reduce noise in our results, we also ran the LIWC analysis on the original article set. For example, categories such as health and biological processes, motion, drives, and work appeared prominently in the articles themselves, which might explain their prominence in the responses. In our discussion, we focus on the significant and prominent categories (high mean values) in the responses and that did not strongly link to the article text.

A first finding concerns the affective processes. Regardless of whether respondents agreed with the ground truth or not, affective processes appeared significantly in respondents who reasoned about false versus credible news (Tables 5 and 6). But whereas negative emotions and anger appeared more prominently among respondents who reasoned about false news articles and who correctly identified them as false (Table 5), positive emotions appeared more prominently among respondents who reasoned about false articles and did not identify them as such (Table 6). This emotional reasoning might indicate some level of ideological alignment with the provided justification and the use of motivated reasoning by respondents. We expand on this point later in the text.

Next, insight, causation, and differentiation all constitute cognitive processes. While insight and causation had a higher mean for credible article responses in Table 5, differentiation had a higher mean for false articles. Interestingly, none of these cognitive reasoning processes appear in Table 6, which represent the respondents who disagreed with the ground truth. This finding lends support to arguments that people might fall for false news due to a lack of cognitive reasoning.

Social processes, while not greatly prominent in the data, did have a borderline significant difference among respondents who did not agree with the ground truth. These respondents used social processes to support disagreement with credible articles (higher mean for the credible articles in Table 6). Future research could explore this interesting finding with a study design that emphasized social processes more than our study did.

Finally, the temporal focus of responses changes from past to present as we move from Table 5 to Table 6. However, both are more strongly associated with reasoning concerning credible news, so we interpret it as an artifact of the response style when one agrees versus disagrees with a given claim and not necessarily tied to reasoning on fake news per se.

Moving to Tables 7 and 8, we can see that correctly reasoning on novel news tends to be more tentative and cautious, which represents the uncertainty that exists around novel news topics. We also can see that social processes play a stronger role in common news, which indicates reasoning about these topics may rely on prior social experiences or with the opinions of social contacts.

Interestingly, Table 8 shows that affective processes were again significant among those respondents who disagreed with the ground truth but that they expressed negative emotions (albeit not significant in our tests). Taken with our previous insight, this finding shows that emotions play a role in how people read and process news articles. Moreover, if we consider Kahan's (2013) argument for motivated reasoning based on partisanship, we can link negative emotion toward COVID-19 and positive emotion toward articles that object to vaccinations and climate change arguments as being affiliated with a particular opinion group. Hence, this evidence lends support to assertion that specific respondent groups used motivated reasoning.

5 Discussion

With a rich data set that contained news consumers' veracity decisions, we employed multiple linguistic tools to better understand how news readers reason about the articles that they read. We examined the differences in reasoning for credible versus false news and for novel versus common news topics.

Framing our work in the prior literature that has argued for both ideologically motivated and cognitive analytical reasoning, we looked for textual clues in responses to understand what reasoning respondents employed under each context. We found that respondents reasoned differently about false versus credible news in general with more nuanced insights stemming from breaking our sample into those who agreed with the ground truth (i.e., correctly identified false and credible news) versus those who disagreed.

At the high analysis level, looking at how respondents used passive tone in reasoning about false versus credible news, we found that, regardless of agreement with the ground truth, people used more passive tone when reasoning about false news than when reasoning about credible news. Therefore, we propose that:

> **P1:** Compared to reasoning about credible news, reasoning about false news involves more passive voice, which reflects greater ambiguity and a more distant attitude.

Examining our results further and separating the data based on agreement with the ground truth, we found that respondents who correctly identified articles' ground truth relied on several cognitive processes. We found as much for both common and novel topics, although respondents displayed a more tentative approach to reasoning about novel news. While they expressed some affect, emotions played a lesser part in their reasoning and were mostly attached to false, rather than credible, news. Respondents who did not agree with the article's ground truth, on the other hand, relied more heavily on affective and social processes, which indicates that political affiliation might have driven the specific emotions they expressed. Cognitive processes did not significantly differ for respondents who disagreed with the article's ground truth neither in the credible/false analysis nor in the common/novel analysis. Taken together, these findings explain why the literature provides evidence both ways (for cognitive and motivated reasoning) and also suggests that emotional and social reasoning have a dominant role in why people disagree with ground truth, whereas cognitive reasoning is associated with people correctly judging news articles. This finding has important implications for designing tools to help mitigate fake news from spreading since one should target different reasoning among users. Thus, we propose that:

> **P2:** Agreement with articles' ground truth is more strongly associated with cognitive reasoning and, to a lesser extent with emotional and social reasoning, whereas disagreement with articles' ground truth is more strongly associated with emotional and social reasoning.

In this study, we also explored news topic novelty as a potential contingency for reasoning style. Past research has tied novelty to information overload and uncertainty, and, in this work, we found support for the latter via analyzing passive versus active voice in the reasoning that respondents provided. Specifically, we found that people use more passive voice to reason about novel versus common news topics. Thus, we propose that:

> **P3:** Compared to reasoning about common news, reasoning on novel news involves more passive voice, which reflects greater ambiguity and a more distant attitude.

While we did not find strong support to propose a direct effect of novelty on reasoning styles through our linguistic analysis, we do see a potential moderation effect in that news topic novelty changes the magnitude of specific emotions and social process. Thus, we propose that:

> **P4:** A news topic's novelty does not change the dynamic between cognitive, emotional, and social reasoning but it weakens their effect likely due to uncertainty.

5.1 Implications for HCI and Interventions Design

With this paper, we open a dialogue in the information systems, psychology, and journalism research communities about the role that reasoning plays in decisions about news' veracity and how those reasonings can inform automated interventions. Specifically, we focus on whether one can design tailored support for decision making about news' veracity. Below, we return to discussing existing tools that can help one detect and mitigate fake news. Given our findings, we also discuss HCI and design implications related to developing automated intervention tools and technologies for cognitive self-reflection. Note that we focus on fake news but do not make any distinctions here between misinformation or disinformation.

Broadly, the literature has suggested three approaches to automatic or semi-automatic interventions: 1) reducing the visibility of unreliable information (filtering, de-recommending, banning producers, etc.), 2) placing a warning label on unreliable information as one shows it to a consumer, and 3) educating information consumers to better evaluate misleading information. Our findings shed light on potential improvements in these three approaches.

First, current approaches treat all consumers as the same and view one type of intervention as suitable for everyone (e.g., Twitter's warning label system as of 2021). However, our findings demonstrate that this approach may not be the most effective one. Furthermore, the past literature on cognitive biases (e.g., see Table 9) has clearly demonstrated that different people will interpret interventions differently. Eliciting consumers' different biases or other philosophical or ideological traits in future studies may provide more granular results than what we can conclude. For example, personal outlooks on specific topics, related past experiences, and ideological leaning may all play a role in the effectiveness of interventions and the reasoning behind belief decisions. In future studies, one could measure these individual differences through both implicit and explicit survey questions. For instance, one could use policy-based political questions or questions related to sociopolitical events to validate self-reported political leanings (e.g., see Horne et al. 2019) in order to capture both explicit and implicit measures of a single individual trait.

Second, warning label systems that one tailors to the consumer may increase the effectiveness of automated and semi-automated interventions. Our study highlights that past shared content plays a critical role here. Using past reasoning to learn about biases (such as the ones in Table 9) may improve intervention designs. For example, using consumers' previous agreement or disagreement with fact-checked information (using previous shares as a proxy) may approximate their typical reasoning and reveal any confirmation bias or valence effect. As such, it could allow one to use different types of warning labels. Again, as we state above, other individual traits besides interactions with past content may provide further insights into effective warning label design. Such traits could include ideological positions, personality traits, philosophical outlooks, or policy positions. One could theoretically indirectly measure these traits through a news reader's online activity. However, eliciting those traits opens these technical solutions to many ethical issues that one should carefully consider and may open these solutions to an array of privacy issues that may do more harm than good. Finally, while a fact check may work for someone using cognitive reasoning, that same fact check might not impact a person who has mainly relies on emotional and social processes. Although one would need to carefully implement and explicitly test the specific change in warning label design using this approach, other works support this notion. Specifically, Horne et al. (2020) and Pennycook et al. (2020) both suggest that warning labels besides explicit fact checking can be effective.

One can broadly apply a similar logic to various debunking and prebunking methods. Debunking and prebunking provide corrections to misinformation with help from technology. However, a correction may leave an inconsistency in one's memory or may cause the incorrect information to be more fluent. One needs to understand how the individual reasons before delivering a correction to avoid negative side effects. For example, providing a cognitive correction to a predominantly emotional thinker might lead to resistance towards receiving the advice. Alternatively, providing a correction for someone with overconfidence in their beliefs might result in lower impact. Again, knowing how users might receive information represents a powerful design tool for AI interventions.

Furthermore, while news item novelty had only a small effect on reasoning, predicting when a situation is novel or during events with high uncertainty (e.g., crisis events) may allow warning labels in general to be more effective. We leave it to future work to clearly define and understand what constitutes "novel enough" to reap these intervention benefits.

Third, increasing news literacy through technologies for self-reflection can be a foundation to educate users about detecting fake news. Indeed, our results show the prevalence of passive voice in reasoning about false news and novel news topics, which indicates uncertainty and insufficient knowledge to enable accurate assessment. By using similar methods as we used to characterize how readers reason about news articles, we believe that online platforms can provide individuals with additional capabilities for cognitive self-reflection. These capabilities could add to intervention approaches that focus on educating news consumers. For example, if one elicited consumers' reasoning, one could use automated approaches, such as the ones we used in this study, to determine reasoning type (cognitive or motivated). Then, similar to shifting one's attention to certain information's accuracy (Pennycook et al., 2020), interventions can shift consumer attention to their own reasoning, which can allow them to self-reflect and potentially increase their future ability to make correct veracity decisions. Again, just as with the previously discussed method, one should explicitly test this approach, which we leave for future work.

Table 9. Theories of Cognitive Biases

Bias	Definition (emphasis added)
Overconfidence effect (Dunning et al., 1990)	"When an **individual's subjective confidence** in their judgements is higher than the objective accuracy of the judgements, it will be considered as overconfidence effect."
Social desirability bias (Fisher, 1993)	"Respondents are often unwilling or unable to report accurately on sensitive topics for ego-defensive or impression management reasons. The result is data that are systematically **biased toward respondents' perceptions of what is 'correct' or socially acceptable** and is considered as social desirability bias."
Selective exposure (Freedman & Sears, 1965)	"People are exposed to disproportionate amounts of supportive information, and **it is extremely difficult to alter the beliefs of people holding clear opinions**." During this process, individuals voluntarily avoid contradictory information that leads to selective exposure bias.
Valence effect (Frijda et al., 1986)	Valence effect emerges when "**individuals tend to overestimate the likelihood of good things** happening rather than negative things."
Contrast effect (Hovland et al., 1957)	Contrast effect is an "unconscious bias that happens when **two entities are judged** based on their differences **in comparison to other** instead of evaluating them individually."
Prospect theory (Kahneman & Tversky, 2013)	Prospect theory focuses on "determining how **individuals assess their loss and gain** perspectives in relation **to their specific situation** rather than in absolute terms".
Confirmation bias (Nickerson, 1998)	Confirmation bias refers to the "**inclination to look for, interpret or favor information** in such a way that **aligns with one's prior beliefs and values**".

5.2 Limitations and Future Work

As with all studies, this study has several limitations. First, the data set that we considered in this research pertained specifically to the chosen news context. Different news topics and context could change the reasonings that news consumers use. We also did not conduct specific manipulation tests to check whether respondents perceived COVID as a novel news topic. However, based on the results, we believe that these articles provided a sufficient basis to investigate different linguistic cues that reflect consumer reasoning and that our findings have some general implications. Second, we may have had a low number of data points (consumer responses) compared to the usual data set sizes that researchers have used in quantitative, natural language processing analyses. However, given the relative uniformity among the number of responses (which the neural sentence similarity heatmap in Figure 1 shows), we argue that we obtained valid insights that do not suffer from any visible biases even given the limited number of responses. Furthermore, we obtained a relatively large number of data points compared to usual human or crowd-sourced experiments.

In this study, we opted for an exploratory study given the lack of direct theoretical guidance on this topic. Future work can start with the propositions that we put forth in this paper and design confirmatory, experimental studies to test them further. Finally, in addition to the future directions that we mention above, we would like to delve deeper into differences across demographics and pre-existing beliefs to investigate individuals' reasoning and sense-making strategies about news articles. To do so robustly, we would like to experiment across wider news contexts and events.

Lastly, our linguistic-based analysis both in the data-quality assessment and in the psycholinguistic assessment may have some alternative explanations. First, while we found in our data-quality assessment that open-ended responses related to the general task, the differences between groups may not have arose due to the type of article that the respondents assessed but rather a more general difference between the words they used when justifying a true or false position. However, this data-quality assessment only constitutes an additional step to ensure robustness. Given the steps that we took in designing our study and using standard AMT quality controls, we have confidence in the data's quality. Second, while researchers have studied LIWC for psycholinguistic analysis, latent factors that we did not measure in our analysis may explain the significant differences between groups. Yet, we did control for LIWC's most obvious shortcoming in our analysis; namely, topic-specific word usage. Specifically, we measured the LIWC categories on the news articles themselves to find groups of words that participants may have used simply due to the topic that they assess (e.g., health and biological processes). Given this control, we have confidence in our statistical results, although some alternative for the found differences may exist as we could not establish a causal connection.

6 Conclusion

In this paper, we highlight the distinct linguistic cues latently present in the syntax, language, and semantics extracted from the news readers' responses on labeling false versus credible news articles. Some interesting insights from this research include: 1) reasoning for false news and for novel news topics involves a more passive tone that indicates uncertainty, 2) news readers employ both cognitive and motivated reasoning processes, and topic novelty moderates (weakens) the effect of cognitive and emotional reasoning, and 3) social processes deserve further investigation as they appear in some reasoning that respondents employed.

Using these insights, we also highlight the different design implications across three broadly used approaches to information veracity interventions. We hope that our approach and findings can influence efforts to design new technological interventions that leverage consumer reasoning to tailor more effective interventions and serve as a stepping-stone for continued work in this area.

Acknowledgments

We thank the editors of the Special Issue on Fake News and Deception for their correspondence. We also thank the anonymous reviewers for their helpful comments on the work presented in this paper. This research was supported in part by NSF grant OIA-1937143 Convergence Accelerator Phase I (RAISE): Credible Open Knowledge Network.

References

Abidin, T. F., Hasanuddin, M., & Mutiawani, V. (2017). N-grams based features for Indonesian tweets classification problems. In *Proceedings of the International Conference on Electrical Engineering and Informatics.*

Aizawa, A. (2003). An information-theoretic perspective of tf-idf measures. *Information Processing & Management*, *39*(1), 45-65.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), 211-36.

Aral, S., & Dhillon, P. S. (2022). What (exactly) is novelty in networks? Unpacking the vision advantages of brokers, bridges, and weak ties. *Management Science*, 1-24.

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, 149*(8), 1608-1613.

Bak-Coleman, J. B., Alfano, M., Barfuss, W., Bergstrom, C. T., Centeno, M. A., Couzin, I. D., Donges, J. F., Galesic, M., Gersick, A. S., Jacquet, J., & Kao, A. B., Moran, R. E., Romanczuk, P., Rubenstein, D. I., Tombak, K. J., van Bavel, J. J., & Weber, E. U. (2021). Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, *118*(27).

Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting factuality of reporting and bias of news media sources. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.1810.01765

Banko, M., & Vanderwende, L. (2004). Using n-grams to understand the nature of summaries. In *Proceedings of the HLT-NAACL Conference.*

Baratta, A. M. (2009). Revealing stance through passive voice. *Journal of pragmatics*, 41(7), 1406-1421.

Bates, J. A., & Lanza, B. A. (2013). Conducting psychology student research via the Mechanical Turk crowdsourcing service. *North American Journal of Psychology*, *15*(2), 385-394.

Bond, G. D., Holman, R. D., Eggert, J. A. L., Speller, L. F., Garcia, O. N., Mejia, S. C., & Rustige, R. (2017). “Lyin’ Ted”, “crooked Hillary”, and “deceptive Donald”: Language of lies in the 2016 US presidential debates. *Applied Cognitive Psychology*, *31*(6), 668-677.

Bolinger, D. (2014). *Language—the loaded weapon: The use and abuse of language today*. Routledge.

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, *114*(28), 7313-7318.

Braun, J. A., & Eklund, J. L. (2019). Fake news, real money: Ad tech platforms, profit-driven hoaxes, and the business of journalism. *Digital Journalism, 7*(1), 1-21.

Brashier, N. M., & Schacter, D. L. (2020). Aging in an era of fake news. *Current Directions in Psychological Science*, *29*(3), 316-323.

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C., & Dredze, M. (2018). Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *American Journal of Public Health*, *108*(10), 1378-1384.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2016). Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality data? In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 133-139). American Psychological Association

Carlebach, M., Cheruvu, R., Walker, B., Magalhaes, C. I., & Jaume, S. (2020). News aggregation with diverse viewpoint identification using neural embeddings and semantic understanding models. In *Proceedings of the 7th Workshop on Argument Mining*.

Ciampaglia, G. L., Shiralkar, P., Rocha, L. M., Bollen, J., Menczer, F., & Flammini, A. (2015). Computational fact checking from knowledge networks. *PloS One*, *10*(6), e0128193.

Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological science*, *28*(11), 1531-1546.

Chan, E. Y., & Maglio, S. J. (2020). The voice of cognition: Active and passive voice influence distance and construal. *Personality and Social Psychology Bulletin*, *46*(4), 547-558.

Chakraborty, A., Paranjape, B., Kakarla, S., & Ganguly, N. (2016). Stop clickbait: Detecting and preventing clickbaits in online news media. In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.*

Clarke, J., Chen, H., Du, D., & Hu, Y. J. (2020). Fake news, investor attention, and market reaction. *Information Systems Research*, *32*(1), 35-52.

Colman, A. M. (2015). *A dictionary of psychology*. Oxford University Press.

Conroy, N. K., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1-4.

Craft, S., Ashley, S., & Maksl, A. (2017). News media literacy and conspiracy theory endorsement. *Communication and the Public*, *2*(4), 388-401.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.1810.04805

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, *58*(4), 568-581.

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*. Retrieved from https://misinforeview.hks.harvard.edu/article/pausing-reduce-false-news/

Feng, S., Banerjee, R., & Choi, Y. (2012). Syntactic stylometry for deception detection. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20(2), 303-315.

Freedman, J. L., & Sears, D. O. (1965). Selective exposure. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (vol. 2, pp. 57-97). Academic Press.

Frijda, N. H. (1986). *The emotions*. Cambridge University Press.

Gruppi, M., Horne, B. D., & Adalı, S. (2021a). NELA-GT-2020: A large multi-labelled news dataset for the study of misinformation in news articles. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.2102.04567

Gruppi, M., Horne, B. D., & Adalı, S. (2021b). Tell me who your friends are: Using content sharing behavior for news source veracity detection. In *Proceedings of the International AAAI Conference on Web and Social Media*.

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, *117*(27), 15536-15545.

Hassan, N., Zhang, G., Arslan, F., Caraballo, J., Jimenez, D., Gawsane, S., Hasan, S., Joseph, M., Kulkarni, A., Nayak, A. K., & Sable, V. (2017). Claimbuster: The first-ever end-to-end fact-checking system. *Proceedings of the VLDB Endowment*, *10*(12), 1945-1948.

Horne, B. D., Adali, S., & Sikdar, S. (2017). Identifying the social signals that drive online discussions: A case study of reddit communities. In *Proceedings of the International Conference on Computer Communication and Networks.*

Horne, B. D., Dron, W., Khedr, S., & Adali, S. (2018). Assessing the news landscape: A multi-module toolkit for evaluating the credibility of news. In *Proceedings of the International World Wide Web Conference.*

Horne, B. D., Nevo, D., O'Donovan, J., Cho, J. H., & Adalı, S. (2019). Rating reliability and bias in news articles: Does ai assistance help everyone? In *Proceedings of the International AAAI Conference on Web and Social Media*.

Horne, B. D., Nevo, D., Adali, S., Manikonda, L., & Arrington, C. (2020). Tailoring heuristics and timing AI interventions for supporting news veracity assessments. *Computers in Human Behavior Reports*, *2*.

Hossain, M., & Abufardeh, S. (2019). A New method of calculating squared Euclidean distance (SED) using pTree technology and its performance analysis. In *Proceedings of 34th International Conference on Computers and Their Applications.*

Hovland, C. I., Harvey, O. J., & Sherif, M. (1957). Assimilation and contrast effects in reactions to communication and attitude change. *The Journal of Abnormal and Social Psychology*, *55*(2), 244-252.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295-1306.

Jin, Z., Cao, J., Jiang, Y. G., & Zhang, Y. (2014). News credibility evaluation on microblog with a hierarchical propagation model. In *Proceedings of the IEEE International Conference on Data Mining*.

Jones-Jang, S. M., Mortensen, T., & Liu, J. (2021). Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American Behavioral Scientist*, *65*(2), 371-388.

Jordan, K. N., Pennebaker, J. W., & Ehrig, C. (2018). The 2016 US presidential candidates and how people tweeted about them. *SAGE Open*, *8*(3).

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision making*, *8*(4), 407-24.

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In L. C. MacLean & W. T. Ziemba (Eds.), *Handbook of the fundamentals of financial decision making: Part I* (pp. 99-127).

Karkali, M., Rousseau, F., Ntoulas, A., & Vazirgiannis, M. (2013). Efficient online novelty detection in news streams. In *Proceedings of the International Conference on Web Information Systems Engineering*.

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. (2020). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods*, *8*(4), 614-629

Kiely, E., & Robertson, L. (2016). How to spot fake news. *FactCheck.org.* Retrieved from https://www.factcheck.org/2016/11/how-to-spot-fake-news/

Kim, J., Tabibian, B., Oh, A., Schölkopf, B., & Gomez-Rodriguez, M. (2018). Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In *Proceedings of the 11th ACM International Conference on Web Search and Data Mining*.

Kogan, S., Moskowitz, T. J., & Niessner, M. (2019). *Fake news: Evidence from financial markets*. Retrieved from https://doi.org/10.2139/ssrn.3237763

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human learning and memory*, 6(2), 107-118.

Lazer, D., Ruck, D. J., Quintana, A., Shugars, S., Joseph, K., Grinberg, N., Gallagher, R. J., Horgan, L., Gitomer, A., Bajak, A., & Baum, M., Ognyanova, K., Qu, H., Hobbs, W. R., McCabe, S., & Green, H. (2021). *The COVID states project #18: Fake news on Twitter*. Retrieved from https://osf.io/vzb9t/

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.

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.

Maksl, A., Craft, S., Ashley, S., & Miller, D. (2017). The usefulness of a news media literacy measure in evaluating a news literacy curriculum. *Journalism & Mass Communication Educator*, *72*(2), 228-241.

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data & Society Research Institute.

Moravec, P., Minas, R., & Dennis, A. R. (2018). Fake news on social media: People believe what they want to believe when it makes no sense at all. *SSRN.* Retrieved from https://ssrn.com/abstract=3269541

Myers, D., & DeWall, C. N. (2018). *Psychology*. Macmillan.

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

Nørregaard, J., Horne, B. D., & Adalı, S. (2019). Nela-gt-2018: A large multi-labelled news dataset for the study of misinformation in news articles. In *Proceedings of the International AAAI Conference on Web and Social Media*.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*(2), 303-330.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC*. Lawrence Erlbaum Associates.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making*, *10*(6), 549-563.

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. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality*, *88*(2), 185-200.

Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*, 66(11), 4944-4957.

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, *25*(5), 388-402.

Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017). Automatic detection of fake news. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.1708.07104

Popat, K., Mukherjee, S., Strötgen, J., & Weikum, G. (2016). Credibility assessment of textual claims on the web. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management*.

Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2017). A stylometric inquiry into hyperpartisan and fake news. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.1702.05638

Řehůřek, R., & Sojka, P. (2011). *Gensim—statistical semantics in python*. Retrieved from https://www.fi.muni.cz/usr/sojka/posters/rehurek-sojka-scipy2011.pdf

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

Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news detection. In *Proceedings of the ACM on Conference on Information and Knowledge Management*.

Sathe, A., Ather, S., Le, T. M., Perry, N., & Park, J. (2020). Automated fact-checking of claims from Wikipedia. In *Proceedings of the 12th Language Resources and Evaluation Conference*.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of low-credibility content by social bots. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.1707.07592

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.

Spezzano, F., Shrestha, A., Fails, J. A., & Stone, B. W. (2021). That's fake news! Reliability of news when provided title, image, source bias & full article. In *Proceedings of the ACM on Human-Computer Interaction*.

Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R. M. (2014). Rumors, false flags, and digital vigilantes: Misinformation on Twitter after the 2013 Boston Marathon bombing. In *iConference 2014 Proceedings.*

Starbird, K., Arif, A., Wilson, T., Van Koevering, K., Yefimova, K., & Scarnecchia, D. (2018). Ecosystem or echo-system? exPloring content sharing across alternative media domains. In *Proceedings of the International AAAI Conference on Web and Social Media*.

Sundar, S. S., Kalyanaraman, S., & Brown, J. (2003). Explicating web site interactivity: Impression formation effects in political campaign sites. *Communication Research*, *30*(1), 30-59.

Talwar, S., Dhir, A., Kaur, P., Zafar, N., & Alrasheedy, M. (2019). Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. *Journal of Retailing and Consumer Services*, *51*, 72-82.

Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining “fake news”: A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137-153.

Thaler, M. (2019). The “fake news” effect: An experiment on motivated reasoning and trust in news. Retrieved from https://scholar.harvard.edu/files/mthaler/files/mthaler_fake-news-effect.pdf

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, *81*(6), 973-988.

Vakulenko, S., Nixon, L., & Lupu, M. (2017). Character-based neural embeddings for tweet clustering. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.1703.05123

Vicario, M. D., Quattrociocchi, W., Scala, A., & Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. *ACM Transactions on the Web, 13*(2), 1-22.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-1151.

Wang, W. Y. (2017). “Liar, liar pants on fire”: A new benchmark dataset for fake news detection. *arXiv.* Retrieved from https://doi.org/10.48550/arXiv.1705.00648

Wardle, C. (2017). Fake news. It’s complicated. *First Draft*. Retrieved from https://firstdraftnews.org/articles/fake-news-complicated/

Wressnegger, C., Schwenk, G., Arp, D., & Rieck, K. (2013). A close look on n-grams in intrusion detection: anomaly detection vs. classification. In In *Proceedings of the ACM Workshop on Artificial Intelligence and Security*.

Yaffe, P. (2020). Active voice, active writing. *Ubiquity*.

Yang, Y., & Pedersen, J. O. (1997). A comparative study on feature selection in text categorization. In *Proceedings of International Conference on Machine Learning*.

Yannakoudakis, E. J., Tsomokos, I., & Hutton, P. J. (1990). N-grams and their implication to natural language understanding. *Pattern Recognition*, *23*(5), 509-528.

Zak, R., Raff, E., & Nicholas, C. (2017). What can N-grams learn for malware detection? In *Proceedings of the International Conference on Malicious and Unwanted Software*.

Zhang, Y., Jin, R., & Zhou, Z. H. (2010). Understanding bag-of-words model: A statistical framework. *International Journal of Machine Learning and Cybernetics*, 1(1-4), 43-52.

Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019). Fake news: Fundamental theories, detection strategies and challenges. In *Proceedings of the 12th ACM International Conference on Web Search and Data Mining*.

Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2020). Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice*, *1*(2), 1-25.

About the Authors

Lydia Manikonda is an Assistant Professor in the Lally School of Management at Rensselaer Polytechnic Institute. She received her PhD in Computer Science from Arizona State University, Tempe, Arizona. Her research focuses on multi-disciplinary problems related to user behavior and cognition by leveraging social media data. Her work has been published in conference venues such as ACM CHI Conference on Human Factors in Computing Systems (CHI), International Conference on Web and Social Media (ICWSM), and in journals such as *The Journal of Artificial Intelligence*, *Computers in Human Behavior*, *Human Computation*, and *International Journal of Geographical Information Science*.

Dorit Nevo is an associate professor of Information Systems at the Lally School of Management at Rensselaer Polytechnic Institute. She received her PhD in Management Information Systems from the University of British Columbia. Dorit's research focuses on human interactions with information technology, technology support for managing knowledge, and crowdsourcing. Her published work includes articles in *MIS Quarterly*, *Information Systems Research*, *Journal of Management Information Systems*, *Communications of the ACM*, *Sloan Management Review*, and *The Wall Street Journal*.

Benjamin Horne is an Assistant professor in the School of Information Sciences at The University of Tennessee Knoxville. He received his PhD in Computer Science from Rensselaer Polytechnic Institute in Troy, New York, where he received the Robert McNaughton Prize for outstanding graduate in Computer Science. He is a highly interdisciplinary, computational social scientist whose research focuses on safety in media spaces. Broadly, this research includes analyzing disinformation, propaganda, conspiracy theories, and the like in both social media and news media. His work has been published in conference venues such as International Conference on Web and Social Media (ICWSM), and in journals such as *ACM Transactions of Intelligent Systems Technology*, *Computers in Human Behavior*, and *Communications of the ACM*. Additionally, his work has been covered in news media, such as Business Insider, Mashable, IEEE Spectrum, and YLE.

Clare Arrington is a doctoral student in the Department of Computer Science at Rensselaer Polytechnic Institute. Her research interests lie in social and linguistic analyses through natural language processing, with a focus on understanding differences in word senses across communities.

Sibel Adali is a professor at Rensselaer Polytechnic Institute in the Department of Computer Science. She obtained her PhD from the University of Maryland. Her work concentrates on cross-cutting problems related to trust, information processing and retrieval, and social networks. She is the author of the book "Modeling Trust Context in Networks", which was published by Springer in 2013. She has worked as the ARL-lead Collaborative Technology Alliance (CTA) wide Trust Coordinator and the Social and Cognitive Networks Academic Research Center (SCNARC) Associate Director. At Rensselaer, she served as the Associate Head and Graduate Program Director of the Computer Science Department. She currently serves as the Associate Dean of Science for Research and Graduate Studies.

Transactions on Human – Computer Interaction