Text Versus Paratext: Understanding Individuals' Accuracy in Assessing Online Information

Sandeep Suntwal
University of Colorado
ssuntwal@uccs.edu

Sue Brown
University of Arizona
suebrown@arizona.edu

Abstract

Fake news has emerged as a significant problem for society. Recent research has shown that shifting attention to accuracy improves the quality of content shared by individuals, thereby helping us mitigate the harmful effects of fake news. However, the parts of a news story that can influence individuals' ability to discern the true state of information presented are understudied. We conducted an online experiment (N=408) to determine how different elements (text and paratext) of a news story influence individuals' ability to detect the true state of the information presented. The participants were presented with the headline (control), main text, graphs/images, and sharing statistics of true and fake news stories and asked to evaluate the story's accuracy based on each of these elements separately. Our findings indicate that individuals were less accurate when identifying fake news from headlines, text, and graphs/images. When asked to evaluate the story based on sharing statistics, they could distinguish fake stories from real news more accurately. Our findings also indicate that heuristics that apply to true news are ineffective for detecting the veracity of fake news.

Keywords: Misinformation, fake news, Paratext, Accuracy

1. Introduction

One of the most significant issues we have in modern society is assuring the accuracy and reliability of the information we depend on in nearly every aspect of our life. Although manipulating information for social, ideological, or financial advantage is not a new issue, the spread of erroneous information at extraordinary speed and scale in the current digital world is a recent phenomenon that can cause enormous harm. Misinformation, disinformation, and malinformation are all ways of describing the information that has been distorted and is a version of misrepresentations of the truth. This overarching phenomenon can be described as "fake news" (Baiyere et al., 2020; Lazer et al., 2018). These misrepresentations damage public confidence in essential democratic institutions, the reliability of scientific findings, and real-world communication. They affect democratic processes, public health, national security, efforts to respond to crises, and civil society overall.

Being able to infer the true state of information in the real world is a highly valuable skill to mitigate the harmful effects of fake news (Allcott & Gentzkow, 2017). Encouraging individuals to share high-quality information can help mitigate the harmful effects of fake news. Moving users' attention to accuracy encourages them to share high-quality content, thus reducing online misinformation (Pennycook et al., 2021). We take a novel approach to understanding what affects an individual's ability to infer the veracity of a news story. Past research has focused on understanding how beliefs and sharing intentions are affected by perceived accuracy (Pennycook et al., 2021) or real accuracy (Allcott & Gentzkow, 2017). These studies have largely utilized headlines as the primary artifact for measuring perceived accuracy (Pennycook et al., 2021) or focused on how accuracy is affected by demographics and how individuals' inference abilities are affected by concordant or discordant news (Allcott & Gentzkow, 2017). In order to extend accuracy-based research, our work focuses on understanding how different elements of the fake news *story,* such as headlines, text, sharing statistics, and graphs/images affect an individual's ability to accurately infer if a story is true or fake.

The primary aim of our study is to determine and better understand which elements of a news story affect the inference abilities of individuals i.e., the accuracy of their perceptions. Understanding the most misleading/confusing elements in a fake news story can help researchers focus on those elements of a fake news story that make fake news harmful overall. Therefore, our research question for this study is: How is the ability to infer the truthfulness of a news story influenced by the text and paratext? We investigate if different paratexts and the text have the same role or influence the accuracy of individuals' perceptions differently. Additionally, we also investigate if this

URI: https://hdl.handle.net/10125/103388
978-0-9981331-6-4

relationship is moderated by the type of information (true or fake) presented to the readers.

Motivated by the theory of paratext (Genette, 1997), we divide the elements of the news story into text: the main content of the news story, and paratext: the peripheral elements of information surrounding the text. The paratexts studied in this work are limited to story headlines, social media sharing statistics, and images/graphs associated with the news story.

We conducted an online experiment where participants evaluated news stories based on the text and paratext. For each story, participants were presented with the text and paratext, one at a time, and asked to report their overall evaluation of the story- if they believed the story was true or fake. We found that paratexts affect individuals' accuracy differently. Readers apply heuristics to infer the accuracy; while some help detect true news correctly, others help infer fake news with higher accuracy.

2. Prior Theory and Research

The fake news artifact has lately attracted a strong interest (Allcott & Gentzkow, 2017; George et al., 2021; Moravec et al., 2022; Pennycook & Rand, 2021; Vosoughi et al., 2018) due to its particular issues, such as targeted news (Moravec et al., 2018). Online fake news uses several channels to spread, including mobile messenger programs (WhatsApp), social media platforms (Twitter, Facebook), and information websites, influencing billions of individuals who use these technologies daily. Multiple nations have also identified the propagation of fake news as a cause for worry (Bill 10, 2019). The effect of fake news across geographies, demographics, technologies, topics, and periods makes it a significant and challenging problem with a high social impact that requires substantial attention (Allcott & Gentzkow, 2017).

Information assets such as fake/phishing websites have also leveraged a variety of presentation elements to fool internet users (Abbasi et al., 2010) and manipulate users' beliefs through paratextual aspects to affect consumers' purchase intentions (Zhang et al., 2016). Fake news and disinformation websites employ similar strategies when combining text and paratext. Fake news websites imitate legitimate websites to convey disinformation, deceiving humans (Tandoc et al., 2018).

Several researchers are aiming to resolve the problem of the spread of fake news (Panenghat et al., 2020). The automated identification of fake news on social media platforms has emerged as a critical method for improving fact verification (Mithun et al., 2021; Shu et al., 2017; Suntwal et al., 2019; Thorne & Vlachos, 2018). However, these methodologies cannot explain why humans spread a specific piece of fake news; more significantly, they do not consider what elements of fake news deceive or convince humans in the first place. By gaining a greater knowledge of human behavior about fake news, it may be possible to improve automated systems.

Studies have also engaged in understanding the influence of information on attitudes and beliefs. Using the theory of engagement (Kahn, 1990), it has been found that readability increased source credibility, but source credibility did not affect active or passive transmission (Maasberg et al., 2018). Additionally, only the originality of content contributed to the active dissemination of knowledge. Maasberg et al. (2018) referred to both individuals and organizations as sources, indicating that consumers form opinions based on a combination of the individual sharing the item and the portal giving the information. However, they did not account for the personal bias of any individuals (e.g., confirmation bias). This was addressed by including confirmation bias in fake news research (Kim & Dennis, 2019). Kim and Dennis (2019) found that story-style headlines were less credible than news-style ones. In addition, they found that source ratings had a considerable beneficial effect on belief. Using reputation theory, they found a strong beneficial influence of confirmation bias on believing. They demonstrated that believing has a favorable impact on activities such as reading, commenting, and sharing.

For true news, what sounds or appears reasonable it is more likely to be categorized or perceived as true. This does not, however, imply that true news always sounds truthful. Occasionally, truth is stranger than fiction. In such circumstances, individuals may incorrectly categorize true news as fake. Typically, for fake news it is not the case that what sounds true is true. Often, fake news attempts to make a falsehood sound convincing (Tandoc et al., 2018). Additionally, the most pervasive fake news articles feature FIBs - Factitious Information Blends - which blend actual and fraudulent aspects to make a narrative appear more credible (George et al., 2021). Therefore, when you trust fake news, you incorrectly categorize it as true. Fake news seeks to entice people to read or access the content. Thus, overall detection accuracy can be lower for true news than fake news.

Because information artifacts such as fake news contain textual content, it is necessary to comprehend the influence of both text and paratext. Comparing text and paratext can help us better comprehend what drives high detection accuracy and what confuses people. We next review the theory of paratexts to understand how paratexts influence behavioral outcomes.

Theory of Paratexts

Paratexts are peripheral instruments and procedures that connect the author, publisher, and reader (Genette 1997). While paratexts can exist independently, text cannot exist without them (Genette 1997, pp. 3). Therefore, it is helpful to study the function of paratexts to understand better the effect of information that is presented to readers. The notion and impact of paratexts are fundamental to comprehending the consumption of digital material (Cronin, 2014). Paratexts may be divided into at least two major categories: spatial and temporal (Genette, 1997) Genette (1997) classifies spatial paratexts as epitext (paratexts that exist outside the book/main text, such as an author interview) and peritext (within the text, e.g., figures, tables). The temporal paratexts are classified as before (appearing before the publication of the book), original (appearing at the time of publication), and delayed (appearing after the book is published). A paratext may fall under one or more of these classifications. Paratexts serve several diverse functions. They promote readability, therefore strengthening the interpretability of text, facilitating the reader's movement across the text, enhancing its navigability, and notifying the reader of the text's commercial value (Birke & Christ, 2013; Cronin, 2014).

Not all paratexts are textual. Paratexts can be illustrative (e.g., symbols, graphics, emojis) or factual (e.g., age, gender, anonymity, and other onymities such as pseudonymity). Although Genette (1997) first introduced the concept of paratexts with physical books as the principal reading device or information artifact in mind, the paratexts discussed in Genette's work translate very easily into the digital realm.

Most original paratext features (such as author information and titles) are accessible on digital and social media platforms. Given the function, scope, and magnitude of digital media, understanding the influence of online paratexts has attracted considerable interest. Research has demonstrated that online information consumption is a multidimensional experience (Hayler, 2016; Hayles, 2004), suggesting that readers do not depend just on text or paratext when making decisions but on both. Multiple disciplines have examined the influence of paratexts. These studies include understanding digital documents (Skare, 2021), determining the relationship between text and viewers/audience on social media (Völcker, 2020), studying the influence of paratextual features of social networking platforms on viral advertising (Seo et al., 2018), determining the role of fan fiction as paratext (Leavenworth, 2015), and understanding videogame walkthroughs (Bergstrom et al., 2018). These fields of study have emphasized paratexts' relevance and influence on various online attitudes.

However, their impact on information processing is understudied, particularly in the arena of online disinformation and other malevolent artifacts. In the field of study on fake news, automated fake news detection methods focus on the text, whereas behavioral methods have concentrated on the specific epitext paratexts. Given the close relationship between text and paratext and the complexity of human information processing, it is important to consider the role of texts and paratexts together to comprehend their influence on several capabilities, such as a reader's ability to detect fake news. Paratexts help decide the trustworthiness and authenticity of a document. However, they are used to different ends based on the context (Leavenworth, 2015, p.57). Conversely, text is scrutinized more carefully as readers judge based on argument quality (Suntwal et al., 2020). Based on this, we propose the following hypotheses:

Hypothesis 1A: *Headline-based accuracy will be higher for true news compared to fake news.*

Hypothesis 1B: *Text-based accuracy will be higher for true news compared to fake news.*

Hypothesis 1C: *Sharing statistics-based accuracy will be higher for true news compared to fake news.*

Hypothesis 1D: *Graph/Images-based accuracy will be higher for true news compared to fake news.*

Human Processing of Information

Individuals process information through two distinct cognitive functions (Kahneman, 2011). Various dual-process theories and models, such as the Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1986), the Heuristics Systematic Model (HSM) (Chaiken, 1980), and System 1–System 2 thinking (Kahneman, 2011), explain how information consumers process information differently. Nonetheless, they concur that there are at least two methods for processing information. The first kind of processing involves carefully processing information depending on its substance or reasoning. This is referred to as the primary route in ELM by the systematic method in HSM and System 2 in System 1 and System 2 thinking. During this phase of data consumption, the reader engages in greater elaboration, motivation, and logical thought. The second processing mode relies on paratexts or cues, which might be derived from the information source or the message itself.

System 1 thinking is spontaneous and constant, producing involuntary judgments outside our conscious awareness (Kahneman, 2011). System 1 thinking is "fast thinking," during which simple heuristics generate perceptions and behaviors in less than one second (Kahneman, 2011). System 1 is the instinctive judgment system (Achtziger & Alós-Ferrer, 2014); when people experience an instinct or "gut response," system 1 dominates. When we use system 1 to conclude, we make decisions based on the emotion and salience of the available information, as opposed to a more nuanced, well-researched approach (Bellini-Leite, 2013; Kahneman, 2011). System 1 is affected by prior memories as well (Moravec et al., 2022). System 2 is more thoughtful and involves deliberation and thoughtfulness (Kahneman, 2011). Thus, while processing information based on the paratexts, users will apply system 1 thinking, and while processing text, they will use system 2 thinking. Therefore, we propose:

Hypothesis 2A: *Headline-based accuracy will be lower than text-based accuracy for true news compared to fake news.*

Headlines provide a higher context compared to sharing statistics. Inference based on the headline is more context and information-driven than sharing statistics. When considering the sharing statistics, individuals look at the popularity and lack thereof. The sharing statistic can be driven due to several factors. However, we know that fake news is shared more than true news on social media (Vosoughi et al., 2018). Previous research has shown that when participants saw a rumor-related tweet with a big number of retweets, likes, and responses, they tended to view it as credible and compelling (Kim, 2018). Therefore, individuals are more likely to think viral content to be true, thus leading to an incorrect determination about the accuracy of the information presented to them. Therefore, we propose:

Hypothesis 2B: *Headline-based accuracy will be higher than sharing-based accuracy for true news compared to fake news.*

Graphs are valuable for informing readers succinctly. However, graphs can be used to mislead individuals in several ways. Simple techniques such as changing scales can manipulate the effects' impact by making small effects look big and vice versa. Graphs can also omit important context about what they represent (Stoffers & Hackett, 2017). Additionally, manipulated images have been used to deceive and influence public opinions (Shen et al., 2019). Headlines are relatively easier to discern through easier heuristics such as loud claims, unreasonable reading times, or punctuations and hype. Therefore, it is harder to determine the accuracy of a story using graphs compared to headlines. Based on this, we propose:

Hypothesis 2C: *Headline-based accuracy will be higher than graph/image-based accuracy for true news compared to fake news*

3. Research Method

Participants

We recruited 51 undergraduates (Juniors and Seniors) from a business course at a large U.S. university. All were between 18 and 24 years old, and around 55% were female. Regardless of demographic background, students use a comparable variety of sources to form decisions (Rosenzweig et al., 2019). We also collected socio-political demographic indicators; see Table 1.

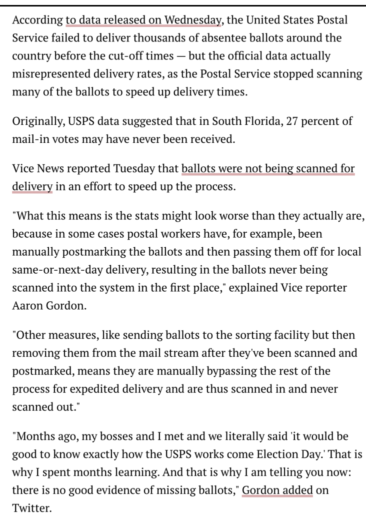
Description	Categories	Percentage
Social media visit frequency	Not everyday	2%
	Once a day	8%
	2-5 times a day	24%
	5-10 times a day	35%
	10+ times	31%
Social media share frequency	Never	23%
	Every few months	31%
	Every few weeks	18%
	Weekly	18%
	Daily	8%
	Multiple times a day	2%
Political affiliation	Democrat	20%
	Independent	23%
	Libertarian	4%
	Republican	33%
	No Preference	20%

Table 1. Participant Demographics

Task

The participants viewed the text (main content of the article) and paratexts (headline, sharing statistics, graph, or image) of two news articles and reported if they found the story accurate based on the element of the article presented to them. Two news stories were left-leaning, and two were right-leaning. Half of the stories were validated to be accurate at the time of the study, while the other half were verified to be fake. We used publicly available datasets (Pennycook et al., 2021) and fact-checking platforms such as PolitiFact and Snopes to select true and fake news articles for our experiment. The same datasets and platforms were used to assign the partisanship labels to the news stories. We eliminated the source name to reduce any news source-specific effect (e.g., some sources are well known and trusted by some readers, while others are not). Similar techniques: using a neutral/unknown name for the source for controlling source effect have been used in existing studies (Moravec et al., 2022). We did not inform our participants explicitly if the article was true or fake, similar to the real world. The text and paratext were developed to prevent large disparities in the type and degree of feelings elicited by those appealing to different ends of the political spectrum. Headlines were presented as they would appear in a news feed. Sharing statistics were presented in Twitter's format. The text was presented how it would appear in a news article. Sharing statistics for all news articles were similar. The graphs/images used in the study were sourced from the articles or a popular tweet that had tweeted the image concerning the article. Example news article with text and paratext (headlines, sharing statistics, and images) is presented in Table 2.

<table>
<tr><th>Information element</th><th>Image</th></tr>
<tr><td>Headline</td><td>


</td></tr>
<tr><td>Text</td><td>


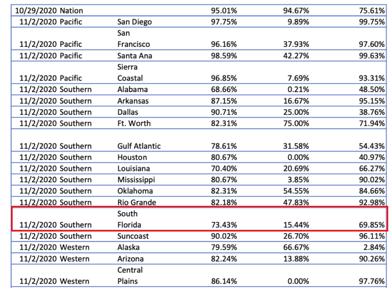
According to data released on Wednesday, the United States Postal Service failed to deliver thousands of absentee ballots around the country before the cut-off times — but the official data actually misrepresented delivery rates, as the Postal Service stopped scanning many of the ballots to speed up delivery times.

Originally, USPS data suggested that in South Florida, 27 percent of mail-in votes may have never been received.

Vice News reported Tuesday that ballots were not being scanned for delivery in an effort to speed up the process.

"What this means is the stats might look worse than they actually are, because in some cases postal workers have, for example, been manually postmarking the ballots and then passing them off for local same-or-next-day delivery, resulting in the ballots never being scanned into the system in the first place," explained Vice reporter Aaron Gordon.

"Other measures, like sending ballots to the sorting facility but then removing them from the mail stream after they've been scanned and postmarked, means they are manually bypassing the rest of the process for expedited delivery and are thus scanned in and never scanned out."

"Months ago, my bosses and I met and we literally said 'it would be good to know exactly how the USPS works come Election Day.' That is why I spent months learning. And that is why I am telling you now: there is no good evidence of missing ballots," Gordon added on Twitter.
</td></tr>
<tr><td>Graph/Image</td><td>

<table>
<tr><td>10/29/2020</td><td>Nation</td><td></td><td>95.01%</td><td>94.67%</td><td>75.61%</td></tr>
<tr><td>11/2/2020</td><td>Pacific</td><td>San Diego</td><td>97.75%</td><td>9.89%</td><td>99.75%</td></tr>
<tr><td>11/2/2020</td><td>Pacific</td><td>San Francisco</td><td>96.16%</td><td>37.93%</td><td>97.60%</td></tr>
<tr><td>11/2/2020</td><td>Pacific</td><td>Santa Ana</td><td>98.59%</td><td>42.27%</td><td>99.63%</td></tr>
<tr><td>11/2/2020</td><td>Pacific</td><td>Sierra Coastal</td><td>96.85%</td><td>7.69%</td><td>93.31%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Alabama</td><td>68.66%</td><td>0.21%</td><td>48.50%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Arkansas</td><td>87.15%</td><td>16.67%</td><td>95.15%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Dallas</td><td>90.71%</td><td>25.00%</td><td>38.76%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Ft. Worth</td><td>82.31%</td><td>75.00%</td><td>71.94%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Gulf Atlantic</td><td>78.61%</td><td>31.58%</td><td>54.43%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Houston</td><td>80.67%</td><td>0.00%</td><td>40.97%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Louisiana</td><td>70.40%</td><td>20.69%</td><td>66.27%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Mississippi</td><td>80.67%</td><td>3.85%</td><td>90.02%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Oklahoma</td><td>82.31%</td><td>54.55%</td><td>84.66%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Rio Grande</td><td>82.18%</td><td>47.83%</td><td>92.98%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>South Florida</td><td>73.43%</td><td>15.44%</td><td>69.85%</td></tr>
<tr><td>11/2/2020</td><td>Southern</td><td>Suncoast</td><td>90.02%</td><td>26.70%</td><td>96.11%</td></tr>
<tr><td>11/2/2020</td><td>Western</td><td>Alaska</td><td>79.59%</td><td>66.67%</td><td>2.84%</td></tr>
<tr><td>11/2/2020</td><td>Western</td><td>Arizona</td><td>82.24%</td><td>13.88%</td><td>90.26%</td></tr>
<tr><td>11/2/2020</td><td>Western</td><td>Central Plains</td><td>86.14%</td><td>0.00%</td><td>97.76%</td></tr>
</table>
</td></tr>
<tr><td>Sharing Statistics</td><td>


</td></tr>
</table>

Table 2. Example news story with text and paratext elements

Treatments

We conducted a within-subject, repeated-measures experiment where the participants viewed one left-leaning and one right-leaning story randomly. The headline paratext was our control. Each participant was first presented with the headline (control) and then presented with the text, sharing statistics and graph/image in random order. For each text and paratext element, the participants responded to the same set of questions. To test that the participants thought through their responses, we also asked them a follow-up question asking them to explain their choice.

The control condition (headline) is based on how news articles appear on the news feed of various news aggregators (e.g., apple, google). This is unlikely to influence the later treatments because users are accustomed to such interfaces. The treatment conditions were presented randomly to simulate online browsing as some individuals prefer to read the text first while others can examine pictures and other peripheral data. Figure 3 represents the overall flow of our experiment.

Participants completed the socio-demographic questionnaire after completing the experiment. An attention check question was presented during the treatment phase to the participants for each article they evaluated. This question appeared at a random position for each participant. Each participant answered the attention check questions correctly. Similar to Pennycook et al. (2021), we asked the participants to report the importance they placed on sharing accurate information online (*How important is it to you that the only time you share news articles on social media is when they are accurate?*) and if they had verified the stories online *(For any of the stories presented to* *you did you verify the story on the internet?).* We also measured our participants' socio-political demographic indicators such as social media usage frequency, frequency of sharing information online, and political affiliation.

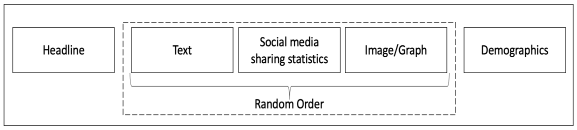


Figure 3. Experiment sequence and flow

Dependent Variable

Participants reported the perceived accuracy of each element (text and paratext) of the article through a binary variable (*To the best of your knowledge, based on the {text/paratext element}, is the news story accurate? Yes/No*). Figure 2 represents an example of participants reporting the perceived accuracy of the story based on the headline paratext. We constructed our dependent variable y_{ise} for a person i, information element e (headline/text/sharing/graph), and story s as follows:

$$y_{ise} = \begin{cases} 1, & if\ i\ correctly\ identifies\ s\ as\ true\ or\ fake \\ & 0, otherwise. \end{cases}$$

We use y_{ise} as the dependent variable and a vector E_i indicating text/paratext elements, T_s indicates if the story s was true or fake in logistic regression:

$$logit(y_{ise}) = \beta_1 E_i * \beta_2 T_s + \beta_0 + \epsilon_{is}$$

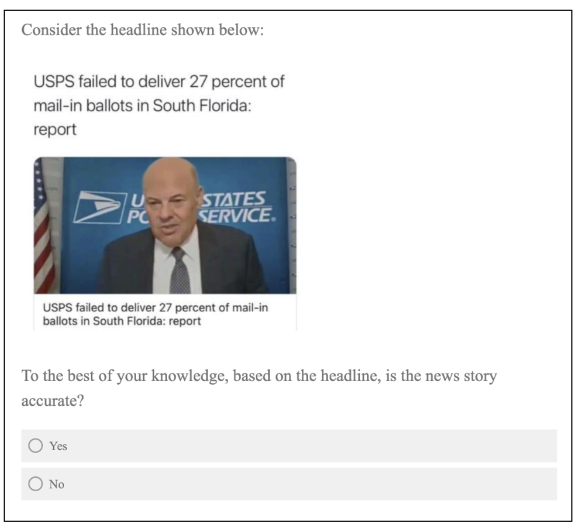


Figure 2. Example news headline with the headline paratext accuracy question.

4. Results

Table 3 reports the results of the model testing. We calculated the interclass coefficient (ICC) value for the model based on the latent variable approach (Goldstein, 2011; Goldstein et al., 2002) and the logistic distribution's variance $(\boldsymbol{\pi^2/3})$ using the expression:

$$\boldsymbol{ICC = \frac{\tau^2}{\tau^3 + \frac{\pi^2}{3}}}$$

and determined there was no necessity for multilevel logistic regression. Thus, we treated each response as an independent data point.

We report the response ratio, individual regression coefficient and significance, standard error for the coefficients, Wald chi-square statistics, odds ratio, the 95 percent confidence interval of odds ratio, model chi-square statistic, Cox and Snell R square, and Nagelkerke pseudo-R^2.

The model fits the data better than the base model. The better fit is evidenced by the significant chi-square test for the difference between the –2LL ratios (the base model includes only the intercept). The Wald statistic $(\boldsymbol{\beta/SE_\beta})$ evaluates the unique contribution of each predictor while holding the other predictors constant. Each significant predictor in our models satisfied the ($p<.05$) standard for statistical significance. The odds ratio indicates the change in odds caused by a change of one unit in the independent variable. We report the confidence interval of 95 percent for this ratio as well. Our analysis coded the correct inference with a 1 and the incorrect inference with a 0. Suppose both limits of the 95 percent confidence interval for the odds ratio are greater than 1. In that case, it is possible to conclude that a one-unit increase in the predictor variable will increase the odds of correctly inferring. Next, suppose both limits of the interval are less than 1. A one-unit increase in the predictor variable will likely increase the odds of incorrect inference, i.e., mistaking a true story for a fake one and vice versa.

The response ratio indicates the proportion of responses that correctly inferred the story from the information element (text/paratext) versus those that did not. The results suggest that more participants could correctly infer the story from the text/paratext element information than those who could not. The –2 log-likelihood ratio (–2LL) determines if the model's predictors impact the dependent variable's prediction. Greater -2LL ratio values indicate a model with poor fit. Two –2LL ratios were computed: one for the base model with no predictors and another for the model with predictors. We expect the –2LL ratio to decrease as we add more predictors to the model. For this purpose, we also performed a model chi-square test, which is the difference between the –2LL ratios of the two logistic models. A significant chi-square statistic indicates that the model with predictors differs significantly from the base model. We reported the R square value (Nagelkerke, CoxSnell), indicating the model's fit. Table 3 presents the overall results based on our model, and Table 4 presents the odds for the text and each paratext.

The odds results are presented in Table 4. Hypotheses 1A, 1B, and 1D were supported. Hypothesis 1C was not supported. It was statistically significant in the opposite direction of the proposed hypothesis. We observe that the ability to infer or detect fake news decreases based on headline, text, and graph/image and increases for the sharing statistics paratext. The odds of correctly inferring fake news based on headlines, text, and graph/image is 0.31, 0.06, and 0.11 compared to true news. For example, there is a one-third chance of detecting a fake news story as fake when compared to determining whether a true news story is true. The odds for inferring fake news based on just the sharing statistics of the story was 8.6. This means that sharing statistics could help infer fake news with higher accuracy.

The true story's headline was the control condition for our study. The odds ratio for this condition is set to 1. The results indicate that hypothesis 2A was not supported. Compared to true story headlines, the accuracy of detecting stories with fake news text was significant ($p<.05$) in the opposite direction of the proposed hypothesis. Hypothesis 2B was supported. The two-way interaction between sharing statistics and news type was significant ($p<.001$, ***). Overall, these interactions inform us that it is harder to infer fake news based on text or graph/images than headlines. Hypothesis 2C was not supported ($p=.12$, NS); there were no statistically significant differences between inferring a true news headline and fake news graphs/images. The main effect of sharing statistics was statistically significant. Overall, the probability of accurately inferring fake news was 31% compared to true news.

Predictor	$\boldsymbol{\beta}$ $(\boldsymbol{SE_\beta})$	***Wald*** $\boldsymbol{\chi^2}$	***Odds Ratio***	**95% CI**	***P***
Intercept	1.29 (0.3)	14.36	3.63	0.66, 2.0	**.001**
Text	0.54 (0.5)	1.06	1.72	−0.48 ,1.63	0.30
Sharing	−2.26 (0.4)	23.81	0.10	−3.21 , −1.38	**.001**
Graph	−0.21 (0.4)	0.21	0.80	−1.15 , 0.70	0.64
Fake News (FN)	−1.17 (0.4)	7.02	0.31	−2.06 , -0.32	**0.01**
Text: FN	−1.63 (0.6)	5.80	0.19	−2.99 , 0.32	**0.01**
Sharing: FN	3.32 (0.6)	27.45	27.76	2.10, 4.59	**.001**
Graph: FN	−0.97 (0.6)	2.34	0.37	−2.22 , 0.26	0.12
N	408				
Response Ratio	229/179 (1/0)				
$\boldsymbol{-2LL}$	-103.62***				
Nagelkerke $\boldsymbol{R^2}$	0.30				
CoxSnell $\boldsymbol{R^2}$	0.22				

Table 3. Overall Model Results

	True News	*Fake News*
Headline	1.0	0.31*
Text	1.72	0.11***
Sharing	0.10	0.86*
Graph	0.80	0.09**

Table 4. Interaction Odds Ratios

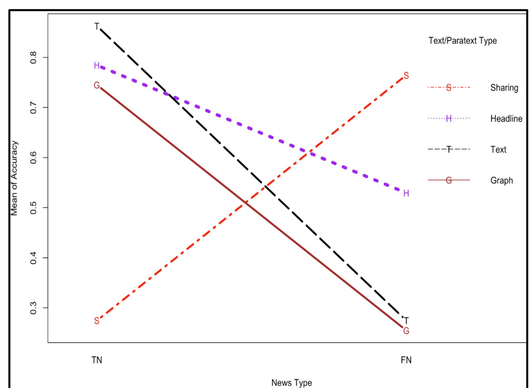


Figure 3. Accuracy plot for the text and paratext

5. Discussion

Our quantitative analysis yielded some unexpected outcomes. Except for the sharing statistics paratext, all other information elements of this study showed that they make it difficult for individuals to infer or detect true stories from fake stories. Figure 3 presents the overall plot for text and each paratext of our study. The x-axis on the plot represents the two levels (true and fake) for the news type factor, and the y-axis represents the mean accuracy.

The headlines paratext was our control as it was the first piece of information that users see in a news feed. Based on just the headlines alone, participants could infer the truthfulness of 65.6% accuracy (identify true news as true and fake news as fake). Participants identified true news stories with 78.43% accuracy and fake news stories with 52.9% accuracy. This indicates that fake news stories confuse individuals more than true stories based on headlines alone. Some participants provided textual feedback indicating why they believed the fake story was true based on the headline; they said "there was not a strong use of buzz words" in the headline, leading them to believe that the story was true. Past studies have indicated that fake news packs more in the headline and less in the body (Horne & Adali, 2017). Our findings show that individuals and institutions that create and spread fake news may understand this and are probably evolving. Fake news peddlers no longer rely on just packing the headline with buzz words and loud claims.

When participants analyzed the stories based on the text, their ability to detect truthfulness dropped to 56.8%. The participants correctly identified true news based on the text with 86.28% accuracy and fake news with only 27.4% accuracy. This finding is counterintuitive because we expected readers to judge the information more critically, using system 2 processing, and reach a correct conclusion based on the details provided in the text. However, we observe that individuals were not able to determine the truthfulness of the fake stories compared to how they inferred the true news stories. This indicates that fake stories may sound very believable and provide convincing arguments, thus making it hard for individuals to identify fake news based on the text. Over the years, fake news propagators have likely invested in the quality of content to make fake news more believable. Therefore, our findings show that if a fake news peddler can make the user click on the headline and read the story, they can deceive the reader with even higher odds. For example, a participant who correctly identified a fake story as fake reported "*it says "2 minute" read and makes major allegations*" *(sic)* as their reason for finding the story fake; for the same story, the participant reported "*Words seem aligned with presidency reputation*" and reported the story to be true based on the text. While individuals have developed heuristics to detect fake news better based on the headlines, fake news peddlers have developed better arguments to convince the readers.

Based on graph/image paratext, participants inferred the correctness of the overall story with 50% accuracy. Participants could correctly infer a story's truthfulness with 74.5% accuracy when the overall story was true. When the story was fake, the accuracy dropped to 25.4%. Again, this finding shows that fake news peddlers use such paratexts to confuse the readers. An individual who incorrectly inferred a fake news story as true reported, "*The data is from credible sources and also has credible dates.*" *(sic)*; this shows that fake news articles are using more credible information to convince the readers by taking care of the subtle details that readers pay attention to. Users have developed their own heuristics to judge news stories, those heuristics may work for true news stories (74.5% accuracy), but if they apply the same heuristics to true news then they fail to recognize fake news (25.4%). For the same story, another individual who incorrectly reported the fake news story as true after verifying the content on the original website on their own reported, "*It was probably a true screenshot at the time, but the website is now probably able to load and is complete*." This shows individuals may form beliefs that are hard to change even if counterevidence is provided, linking our finding to the existing literature on confirmation bias (Nickerson, 1998) and fake news (Kim & Dennis, 2019; Moravec et al., 2018; Suntwal et al., 2020).

The sharing statistics were the final paratext of our study. Here, we observed that individuals were able to determine true news with *lower* accuracy (27.4%) compared to fake news (76.4%). While it is difficult to judge the truthfulness of the story based on the sharing statistics alone, it is important to understand the influence of this paratext. The average likes and retweets for all the stories that were presented to the users were 280 and 230, respectively. We observed several heuristics that individuals used to arrive at their decisions. These heuristics were largely around three themes: a) Fake tweets get more likes than real ones, b) If something is true, there would be motivation for it to be spread c) likes or retweets cannot determine the accuracy of the story. Some readers reported the story was true because of the high number of likes and shares, and for the same story, users reported that the story was fake because the number of likes and shares was low. Virality is a key focus in the fake news literature (Vosoughi et al., 2018). However, our findings indicate that the virality of information is subjective and causes people to inaccurately infer a true story as fake.

Our findings contribute to the literature in several ways. First, our findings show that there is a statistically significant difference between text and different paratexts and how they differ in helping infer real-world true and fake information. This suggests that beyond headlines, there is a need to investigate other paratexts as well. Dual-process theories argue that spending time and evaluating is a part of the system 2/systematic model/central route evaluation. However, when readers evaluate fake news content critically, the fake news content manages to convince and confuse them into believing that it is true. This indicates that there is a need to better understand how information processing occurs for fake news and true news content. There is a need for better understanding items in the perceived argument quality construct and evaluate if certain items can inform us about the specific areas of perceived argument quality that are similar and different for true news and fake news. Understanding the differences will aid several areas of research such as electronic word of mouth (eWom) literature as well where the argument quality is used to rate user reviews.

6. Limitations and Future Directions

Our story investigates the role of text and paratext in spreading information. There are a few limitations of our study. First, we focused on only three paratexts. Several other paratexts such as source names, comments from other readers, and more on social media platforms should be investigated to further understand the impact of paratexts on individuals' ability to infer the truthfulness of the information presented to them on social media platforms. Second, we focused on one set of sharing statistics. Future research may consider large and very large sharing statistics and study if the sharing and like statistics behave similarly. Third, we used the sharing statistics from one platform (Twitter). Some users may find everything on the platform as fake or true if they dislike or like the platform itself. Future studies should further investigate the platform effect of the statistics to understand the interaction effect of platform and sharing statistics on fake news detection accuracy (Krafft & Donovan, 2020). In the future, we aim to study how learning styles (e.g., visual, reading, hearing) affect participants' reading patterns and ability to infer the true state of information presented. Future studies can also investigate government mistrust, bias, and other factors that influence individuals' political ideology.

7. Conclusion

We conducted an online experiment to understand how text and paratext differ in their impact to infer the state of information presented to the reader. Our readers were presented with true and fake news stories and asked to evaluate whether the story was accurate based on the text or paratext. Participants identified true with as true news with high accuracy based on the text, headline, and graph/image presented to them. Sharing statistics paratext helped in identifying fake news with high accuracy. Our study shows that some paratexts aid in detecting true news correctly while others help identify fake news correctly. Heuristics that apply to true news don't apply as efficiently to fake news and vice versa, necessitating a need to understand how these elements affect people's abilities to judge. The critical evaluation of different news based on the text shows that fake news is evolving and still manages to confuse readers.

8. References

Abbasi, A., Zhang, Z., Zimbra, D., & Chen, H. (2010). Detecting fake websites: The contribution of statistical learning theory. *MIS Quarterly*, 435–461.

Achtziger, A., & Alós-Ferrer, C. (2014). Fast or rational? A response-times study of Bayesian updating. *Management Science*, *60*(4), 923–938. https://doi.org/10.1287/mnsc.2013.1793

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236.

Baiyere, A., Avital, M., Dennis, A., Gibbs, J., & Te'eni, D. (2020). *Fake News: What Is It and Why Does It Matter?*

Bellini-Leite, S. (2013). The Embodied Embedded Character of System 1 Processing. *Mens Sana Monographs*, *11*(1), 239. https://doi.org/10.4103/0973-1229.109345

Bergstrom, K., Flynn-Jones, E., Jenson, J., & Hebert, C. (2018). Videogame walkthroughs in educational settings: Challenges, successes, and suggestions for future use. *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Birke, D., & Christ, B. (2013). Paratext and Digitized Narrative: Mapping the Field. *Narrative*, *21*(1), 65–87.

Chaiken, S. (1980). Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752–766. https://psycnet.apa.org/record/1981-28035-001

Cronin, B. (2014). *Foreword: The penumbral world of the paratext. Examining paratextual theory and its applications in digital culture*.

Genette, G. (1997). *Paratexts: Thresholds of interpretation*.

George, J., Gerhart, N., & Torres, R. (2021). Uncovering the Truth about Fake News: A Research Model Grounded in Multi-Disciplinary Literature. *Journal of Management Information Systems*, *38*(4), 1067–1094.

Goldstein, H. (2011). *Multilevel statistical models*.

Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning Variation in Multilevel Models. *Understanding Statistics*, *1*(4), 223–231.

Hayler, M. (2016). Matter matters: the effects of materiality and the move from page to screen. *Research Methods for Reading Digital Data in the Digital Humanities*, 14–35.

Hayles, N. (2004). Print is flat, code is deep: The importance of media-specific analysis. *Poetics Today*, *25*(1).

Horne, B. D., & 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. In *aaai.org*.
Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, *33*(4), 692–724.
Kahneman, D. (2011). *Thinking, Fast and Slow*.
Kim, A., & Dennis, A. (2019). Says Who? The Effects of Presentation Format and Source Rating on Fake News in Social Media. *MIS Quarterly*, *43*(3), 1025–1039. https://doi.org/10.25300/MISQ/2019/15188
Kim, J. W. (2018). Rumor has it: The effects of virality metrics on rumor believability and transmission on Twitter. *New Media & Society*, *20*(12), 4807–4825. https://doi.org/10.1177/1461444818784945
Krafft, P. M., & Donovan, J. (2020). Disinformation by design: The use of evidence collages and platform filtering in a media manipulation campaign. *Political Communication*, *37*(2), 194–214. https://doi.org/10.1080/10584609.2019.1686094
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
Leavenworth, M. (2015). The paratext of fan fiction. *Narrative*, *23*(1), 40–60.
Maasberg, M., Ayaburi, E., & Liu, C. Z. (2018). Exploring the Propagation of Fake Cyber News : An Experimental Approach. *Proceedings of the 51st Hawaii International Conference on System Sciences*.
Mithun, M. P., Suntwal, S., & Surdeanu, M. (2021). Data and Model Distillation as a Solution for Domain-transferable Fact Verification. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4546–4552.
Moravec, Kim, A., & Dennis, A. (2022). Do you really know if it's true? How asking users to rate stories affects belief in fake news on social media. *Information Systems Research*, *2019-Janua*, 6602–6611. https://doi.org/10.1287/isre.2021.1090
Moravec, P., Minas, R., & Dennis, A. (2018). Fake news on social media: People believe what they want to believe when it makes no sense at all. *Kelley School of Business Research Paper*.
Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, *2*(2), 175–220.
Panenghat, M. P., Suntwal, S., Rafique, F., Sharp, R., & Surdeanu, M. (2020). Towards the necessity for debiasing natural language inference datasets. *European Language Resources Association (ELRA)*, 11–16.
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., & Rand, D. G. (2021). The Psychology of Fake News. *Trends in Cognitive Sciences*, *25*(5), 388–402.
Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In *Communication and Persuasion* (pp. 1–24). Springer New York. https://doi.org/10.1007/978-1-4612-4964-1_1
Rosenzweig, J., Thill, M., & Lambert, F. (2019). Student constructions of authority in the Framework era: A bibliometric pilot study using a faceted taxonomy. *College & Research Libraries*, 401–420. https://crl.acrl.org/index.php/crl/article/view/17400
Seo, Y., Li, X., Choi, Y., & Yoon, S. (2018). Narrative transportation and paratextual features of social media in viral advertising. *Journal of Advertising*, *47*(1), 83–95. https://doi.org/10.1080/00913367.2017.1405752
Shen, C., Kasra, M., Pan, W., Bassett, G. A., Malloch, Y., & O'brien, J. F. (2019). Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online. *New Media & Society*, *21*(2), 438–463.
Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36.
Skare, R. (2021). The paratext of digital documents. *Journal of Documentation*, *77*(2), 449–460.
Stoffers, C., & Hackett, J. (2017). *Fake News, Fake Data*. Scholastic.
Suntwal, S., Brown, S. A., & Patton, M. W. (2020). How does Information Spread? A Study of True and Fake News. *Proceedings of the 53rd Hawaii International Conference on System Sciences.*
Suntwal, S., Paul, M., Sharp, R., & Surdeanu, M. (2019). On the Importance of Delexicalization for Fact Verification. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3404–3409.
Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining "Fake News": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153.
Thorne, J., & Vlachos, A. (2018). *Automated Fact Checking: Task formulations, methods and future directions*. http://arxiv.org/abs/1806.07687
Völcker, M. (2020). Paratexts on a social network site and their relevance in the production of meaning-Results of a qualitative investigation of Twitter-Feeds. *PLoS ONE*, *15*(9 September). https://doi.org/10.1371/JOURNAL.PONE.0238765
Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151.
Zhang, D., Zhou, L., Kehoe, J. L., Isil, &., & Kilic, Y. (2016). What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews. *Journal of Management Information Systems*, *33*(2), 456–481.