2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

How Do People Decide Political News Credibility?

Francesca Spezzano
Department of Computer Science
Boise State University
Boise, Idaho (USA)
francescaspezzano@boisestate.edu

Don Winiecki
Department of Organizational
Performance & Workplace Learning
Boise State University
Boise, Idaho (USA)
dwiniecki@boisestate.edu

***Abstract*—In this paper, we share preliminary results from our research work focused on understanding how people assess news items as fake or real and improve their ability to identify fake news. Using existing real/fake news samples and best practices in qualitative, inductive data analysis, we identify factors that appear to impede ability of individuals to identify fake news. Based on this work we suggest one approach to improve human ability to identify fake news, and sketch a process for systematic development of means for supporting people in identifying fake news.**

***Index Terms*—Misinformation, social media, analytic social science, qualitative methods.**

I. Introduction

The issue of jaundiced, biased, or fake news is not new. However, with the global reach of social media and algorithmic media distribution, the issue has grown into a pan-societal problem. Research shows that misleading news distributed online has exploded in volume and popularity in recent years through social network platforms and online news sources. The problem is worsened because people have difficulty discerning Web misinformation from truthful information [1], especially younger generations [2]. Moreover, while 64% of Americans believe that misleading stories cause confusion about basic facts of current events, people frequently share news on social media without even reading what they are forwarding [3]. This suggests that — whether they are reading and critically analyzing news items or not — humans are the most virulent agent in spreading fake news [4]. In turn, we can only conclude that any set of interventions to influence the detection and spread of fake news must also include efforts to assist humans in detecting and responding to fake news.

In this paper, we describe our methods and present results of a preliminary investigation into detection of fake news by humans. Our results show that people mainly rely on their "gut-hunches" when deciding news credibility, independently of the way news is presented to them (e.g., whether only the news headline and image are provided or the full news excerpt). Also, they frequently rely on *their subjective trust* in the source of information when this information is available.

Furthermore, our observations and methods point to a factor commonly-associated with low rates for detecting fake/real news by human readers who did not have any experience with social scientific analysis of textual and visual data. In contrast, when the same fake/real news items were analyzed by individuals with experience in social scientific data analysis, rates of detecting fake/real news were substantively improved.

These preliminary findings allow us to provide specific guidance for improving the ability of human readers to discriminate fake from real news items. At the same time, our methods and in-process discoveries in this investigation point to future research that will allow us to systematically verify and extend our findings, and use them in enhancing the ability to develop means for supporting people in detecting fake news items.

IEEE/ACM ASONAM 2020, December 7-10, 2020
978-1-7281-1056-1/20/$31.00 © 2020 IEEE

II. Related Work

The study of how people assess the credibility of online information requires a multidisciplinary effort as it touches information science, psychology, sociology, communication, and education. Metzger and Flanagin have summarized existing research on the factors that influence people when they make credibility evaluation decisions under different categories: site or source cues; author cues; message cues; receiver characteristics; and social interactions [5]. Furthermore, the dual processing model of credibility assessment states that people tend to use two general strategies, namely "analytic" and "heuristic", that reflect a greater and a lesser degree of cognitive rigor, respectively [6]. The analytical strategy suggests that people are systematically identifying details provided in a news item and what they confidently "know" from other sources in order to produce a reasoned assessment. "Heuristic" implies that people use pre-existing rules that they adapt to the case at hand. Heuristic — or what we might call "gut-hunch" methods — also involve a superficial (i.e., non-systematic) evaluation of the piece of information where user's gut-hunches are often predominant. "Whether the use of heuristic evaluation strategies leads to good or bad credibility decisions" is left as an open question in [5].

In information science, several methods have been proposed to automatically determine whether a piece of news is real or fake [1], [7]. These methods use features extracted from news content, title, associated images, and news propagation patterns in social networks. Studies have highlighted that humans are poor at identifying false information [1], [4], while Horne et al. [8] found that AI assistance with feature-based explanations improves people's accuracy of news perceptions — a human augmented with AI has enhanced ability to identify fake news.

Research focused on creating models of human behavior involved in spreading information in social networks has considered behavior associated with the presence of misinformation [9], [10]. Most of these works are based on minor variations of epidemiological models or opinion dynamics models. Zhou et al. [7] point out a number of fundamental positivistic (i.e., cause-effect or stimulus-response) theories in psychology, philosophy and economics that describe how people interact and share information in social networks. These fundamental theories have come to have a substantial impact in attempting to explain how fake news spreads in social networks and must be taken into account when defining information diffusion models. Understanding how people asses news credibility is one of the first steps towards the development of these models. Our effort follows this pursuit of understanding of how humans assess news credibility, but starts from a more naturalistic (i.e., oriented to human sensemaking instead of attempting to identify stimulus-response links) rather than positivistic position. We chose the former not because we believe the latter to be false or inaccurate, but because a humanistic approach may allow us to identify in greater detail how humans do what they do, rather than just identify what they do.

III. Methods

This study was conducted using an online "Fake News" survey provided via Qualtrics. Through this online survey, participants were asked to determine if news items were real or fake news and then explain the reasoning for their decision. We considered 16 real and 16 fake news from the FakeNewsNet dataset for our study [11]. This dataset contains news articles (title, excerpt and associated image) labeled as real or fake by Politifact and Buzzfeed fact-checking websites. We extracted the article source bias from the MediaBias/FactCheck website which assigns seven degrees of bias: extreme-right, right, right-centered, least-biased, left-centered, left, and extreme-left. For each category of news used in our survey (real or fake), we selected half left-leaning and half right-leaning news.

The participants responded to four different types of questions where we varied information provided: 1. Title, image, and source bias; 2. Title and image; 3. Title and source bias; 4. News excerpt (text only, with no additional indication of source bias). Each participant responded to three questions for each of types 1, 2, and 3 and two questions of type 4.

In addition to asking for an assessment of whether each survey item is an example of real or fake news, one of the principal features of the survey is a question asking respondents the question "Please explain why you said the above news real or fake." The reason for this question was to create data with which we could begin to identify what led respondents to identify a news item as real or fake. We recruited n=17 people (Boise State undergraduate and graduate students) to participate in our survey and collected a total of 187 answers across all the conditions considered.[1]

[1]This research has been approved by Boise State Institutional Review Board (IRB) under protocol number 126-SB19-220.

IV. Preliminary Analysis and Discussion

In order to analyze survey participants' responses to the request for an explanation of their assessment of real or fake, we proceeded as follows.

First, we recruited two upper-division undergraduate students of Sociology to accomplish inductive coding of the (news) items in the "Fake News" survey.[2] These students were given access to the same survey items as were the subjects of the "Fake News" survey.

Each of these students was assigned to independently develop a provisional codebook induced from inspection of items in the survey. Their codebooks were to identify what they saw to be notable elements of the media as used to present content in each question in the "Fake News" survey. After independently producing their codebooks, they met to synchronize them by adapting and/or accommodating codebook items to incorporate similarities across their codebooks, and/or accommodate differences. Following creation of a unified codebook, they applied their codes to survey items to demonstrate how their codebook items were to be used. This process of independent coding, paired-inspection, accommodation of codes, and application, is consistent with best practices for inductive coding of qualitative data [12], [13]. Example codes and definitions in this codebook are shown in Table I.

Next, the resulting codebook was provided to the second author, an ethnographic sociologist who has used qualitative methods and inductive coding processes in research, and taught qualitative data collection and data analysis courses at the graduate level. The second author reviewed media in the "Fake News" survey and refined the codebook; then the resulting codebook was used to code responses provided by survey respondents to the question "Please explain why you said the above news is real or fake" in which they provided accounts for why they had responded as they had to each question in the survey.

Application of the codebook required additions and modifications to codes and code definitions in the codebook in order to account for factors included by respondents to the survey that were not included by our undergraduate research assistants. Importantly, in some cases additions and modifications served to identify what we will describe as "non-empirical" factors or judgments that respondents to the survey used in explaining why they responded to a question as they had. This is subtly but importantly different from empirical factors incorporated into the combined codebook by our research assistants. For example, while the initial codebook included items like "Political/Public People" (indicating the fact that individuals involved in politics or with public recognition are included), respondents sometimes went beyond such facticity to make judgments that only started with such facts.

For example, in several cases, survey respondents reported an explicit trust or distrust in the indicated source for informa-

[2]These students had received coursework in inductive qualitative data analysis and had at least two semesters of experience in faculty-guided research activity involving open-ended and semi-structured interviewing, and inductive coding of interview transcripts.

TABLE I
EXAMPLE CODES AND DEFINITIONS PRODUCED BY SOCIOLOGY STUDENTS.

Code	Code Definition
Clickbait	Vague details; selective capitalization for emphasis; image and title mismatch; exaggerated title; images distorted, or taken out of context and juxtaposed to support title/story; unprofessional.
Frequency	Many; more; or various news sources; mainstream; popular; mainstream; popular.
Plausible	Quotations, looks believable, intuition, seems consistent with things outside this question, nonpartisan, professional or believable title, consistent with other information outside this question.
Political/Public People	Image or text references political person individuals of opposing parties; public person.
Questionable Language	Trigger terms or highly loaded, emotional terms; quotation used/implied that mocks, implicitly accuses, or shames individuals named and/or depicted; hyperbolic and/or absolute assertions.
Questionable Formatting	Use of non-standard capitalization or text features to highlight or emphasize items that appeal to partisan arguments (all caps, selective italics, color, etc.)
Silly	Details make an absurd claim (e.g., American issue inserted into English context; claims unsupported by other details).

tion in a survey question (e.g., "I think it's real IF it's presented by ABC News"; "It's ABC [so] you never know"; "Fox news is typically correct"). In other cases survey respondents invoked information or ideas that rest on data that had to be sourced from outside the survey question itself, or opinions and assumptions without any empirical link to the media included with a question itself (e.g., "Doesn't make sense for an economic speech to focus on people with disabilities"; "The source (Extreme Right) could easily be seen making such a statement about the Washington Post, but I suspect the quote is out of context or even made up").

In other cases respondents admitted to some cynicism relative to news reporting in general and even doubts over the prospect of discriminating real from fake news at all (e.g., "[Contents of this story are] just too good to be true"; "Some sexist and disrespectful people might actually have such opinions. But should the media and agencies use such languages in the headlines? Don't think so.."; "The main problem here is that the whole fake news" versus "real news" is an absolutely terrible metric. Yes there are news sources that are entirely fabricated however that metric doesn't give us any way to directly improve the source. It's meant for complaining not fixing anything"). This leads us to consider that some people simply abandon any idea that real/fake news is even a legitimate typology, thus that they may simply take whatever is presented and treat it as only as a trigger to what they would believe anyway (i.e., confirmation bias).

In this process, it was also identified that respondents tended to include features of photos if provided, or the source attribution, even in cases where the photo or source attribution (for example, the company logo) was nominally only decorative or incidental to the news item (e.g., "I think it's real IF it's presented by ABC News. It is an image of Mrs. Clinton, but she doesn't appear to be ill or unbalanced"; "ABC has a reputation for genuine news and the article title and image caption looks believable"; "ABC news have a good reputation for reliable news").

V. IMPLICATIONS, CONTRIBUTIONS, AND FUTURE WORK

While preliminary, the work described above points to several avenues for further investigation and new contributions.

On one hand, identifying the importance of codes to account for "non-empirical" factors provides a window into aspects of the decision-making of our survey respondents that may have to be investigated more deeply if such decision-making processes are used to develop and/or improve models for human credibility assessment.

Additionally, by analyzing the most used codes we observe several things. First, there is a frequent use of what we will call "gut-hunches" when assessing news credibility or to indicate that an item needed follow-up investigation. This is apparent in the common use of respondents' accounts coded "plausible." This is independent of the way the news was presented to the participants as it appeared at least 27% of the time in each of the four cases considered in the "Fake News" survey. This code indicates subjects' references to quotations, an un-elaborated indication that an item "looked believable," use of intuition to assess an item, reference to knowing or having seen similar stories previously, declaring a news item to be nonpartisan, appearing professional, or including use of a believable title.

Second, when information about the source or source bias was provided, subjects referenced source bias with high frequency in both case 1 (37%, title, image, and source bias provided) and case 3 (31%, title, and source bias provided). The fact that subjects alluded to personal "gut-hunch" with such frequency, and so commonly included references to source bias, indicates that information from outside the story itself, such as personal belief in, and even commitment to, political bias has a strong influence on how people perceive realness or fakeness in news.

Similarly, when multiple media and types of information are provided (i.e., mixtures of photo, story text, and/or source bias), respondents commonly included internal references to each type in their accounts, and told how they saw those bits coming together to contribute to a true or false determination. This shows how our respondents are actively constructing their decision with information provided but also with "glue" from outside the item itself.

On the other hand, in this pilot study we can report a remarkable observation. The ability of our survey respondents to identify real or fake news was almost a coin flip; they were correct on only 53% of the 187 survey items. We associate this

with the observation that in many cases survey respondents made mention of personal opinions and assumptions — but not empirical details — having influenced their decisions. This suggests that an individual's already-existing information-and influence-networks have a substantive impact on what they "see" and "understand" in the survey items. We contrast this with the accuracy of survey responses accomplished when the second author (himself a credentialed social scientist) completed the same survey using codes produced by our undergraduate social science research assistants to guide responses to the survey items. Definitions of these codes in our code dictionary indicated that these codes accounted for objective features of survey items (e.g., terminology, match or mismatch of images and text, source citations, etc.), and analytic social science interpretations of combinations of these factors within particular survey items. The accuracy of real/fake determinations was 68% correct in this application.

To provide more detail on this process, codes like "Questionable formatting," "Questionable language," and "Clickbait" were used to identify features such as use of text colors, bold and italic fonts, and hyperbolic terminology *in patterns* that — when made apparent — led the reader's focus to politically-charged interpretations mostly uncommon in truthful items. Similarly, disagreement between titles, pictures, and story contents suggests editorial intent in attracting individuals with preconceived ideas about the item. It was the *lack of* use of such codes which often cleared the way toward an interpretation that an item was more likely to be factual. This suggests that fake items are those which demonstrate particular patterns in the use of formatting, vocabulary, and semantic interpretations produced through semantic analysis of those objectively-identifiable patterns. While 68% accuracy is hardly a demonstration of remarkable success, it does reflect a substantive improvement in accuracy achieved when using more rigorous social science logic. This reinforces existing research indicating that (a) when left to their own "gut-hunch" means, humans are poor at identifying fake news [1], [4]. It also points to (b) the relative value of replacing "gut-hunch" processing with analytic social science (i.e., focusing on abstractions based on empirically identifiable factors, patterns of those factors, and semantic analysis of those patterns found in news media) as a more promising approach for discriminating real and fake news items.

From this, and with the goal of improving humans' ability to discriminate fake from real news, we suggest two things. First, we advise development of instruction and tools that augment people with information, skills, and tools supporting an empirically-grounded approach to assessment of news. Second, we advise development and systematic delivery of instruction that demonstrates hazards of a non-empirical (or "gut-hunch") approach to assessment of news. In combination, these will aid people in assessing news items, and raise more awareness of the problem in order to decrease the unconscious spread of misinformation [14]–[16].

In addition, this pilot study highlights several things that must be more deliberately pursued in future work. First, creation of a taxonomy of empirically identifiable factors in news sources. Second, determine individual and combined predictive power those factors have in detecting fake news. Third, identify the direction (real/fake) and strength of influence of factors that lead individuals to make "gut-hunch" judgments about news items. Fourth and finally, identify the reliability of readers'/viewers' non-empirically-grounded judgements about news items. Accomplishing these four things will provide firm ground for taking next steps in improving means for supporting people in detecting real or fake news and developing behavioral models of human credibility assessment. The latter will contribute to a better understanding of misinformation spread in social networks.

Acknowledgements: We would like to thank Alyse Palsulich (Treasure Valley Math and Science Center High School) for contributing to initial stages of this research, and Allyssa Hernandez and Ramona Hinrichs (Sociology department, Boise State University) for helping with the inductive coding process. The first author has been partially supported by the National Science Foundation under Award no. 1943370.

REFERENCES

[1] S. Kumar and N. Shah, "False information on web and social media: A survey," *CoRR*, vol. abs/1804.08559, 2018.
[2] S. Wineburg, S. McGrew, J. Breakstone, and T. Ortega, "Evaluating information: The cornerstone of civic online reasoning," *Stanford Digital Repository. Retrieved January*, vol. 8, p. 2018, 2016.
[3] M. Gabielkov, A. Ramachandran, A. Chaintreau, and A. Legout, "Social clicks: What and who gets read on twitter?" *ACM SIGMETRICS Performance Evaluation Review*, vol. 44, no. 1, pp. 179–192, 2016.
[4] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.
[5] M. J. Metzger and A. J. Flanagin, "Psychological approaches to credibility assessment online," *The handbook of the psychology of communication technology*, vol. 32, pp. 445–466, 2015.
[6] M. J. Metzger, "Making sense of credibility on the web: Models for evaluating online information and recommendations for future research," *Journal of the American Society for Information Science and Technology*, vol. 58, no. 13, pp. 2078–2091, 2007.
[7] X. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Comput. Surv.*, vol. 53, no. 5, 2020.
[8] B. D. Horne, D. Nevo, J. O'Donovan, J. Cho, and S. Adali, "Rating reliability and bias in news articles: Does AI assistance help everyone?" in *ICWSM*, 2019, pp. 247–256.
[9] F. Jin, E. Dougherty, P. Saraf, Y. Cao, and N. Ramakrishnan, "Epidemiological modeling of news and rumors on twitter," in *Proceedings of the 7th Workshop on Social Network Mining and Analysis*, 2013, p. 8.
[10] J.-H. Cho, S. Rager, J. O'Donovan, S. Adali, and B. D. Horne, "Uncertainty-based false information propagation in social networks," *ACM Transactions on Social Computing*, vol. 2, no. 2, p. 5, 2019.
[11] K. Shu, S. Wang, and H. Liu, "Beyond news contents: The role of social context for fake news detection," in *WSDM*. ACM, 2019, pp. 312–320.
[12] J. Saldaña, *The coding manual for qualitative researchers*. Sage, 2015.
[13] M. LeCompte and J. J. Schensul, "Analysis from the bottom up: the item level of analysis," *Analyzing and Interpreting Ethnographic Data*, pp. 67–83, 1999.
[14] P. America. (2018) Faking news: Fraudulent news and the fight for truth. [Online]. Available: https://pen.org/faking-news/
[15] J. Kahne and B. Bowyer, "Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation," *American Educational Research Journal*, vol. 54, no. 1, pp. 3–34, 2017.
[16] S. Van der Linden, A. Leiserowitz, S. Rosenthal, and E. Maibach, "Inoculating the public against misinformation about climate change," *Global Challenges*, vol. 1, no. 2, p. 1600008, 2017.