

# Analyzing Biases in Perception of Truth in News Stories and Their Implications for Fact Checking

Mahmoudreza Babaei  
MPI-SWS, Germany

Abhijnan Chakraborty  
MPI-SWS, Germany

Juhi Kulshrestha  
GESIS, Germany

Elissa M. Redmiles  
University of Maryland, US

Meeyoung Cha  
KAIST, South Korea

Krishna P. Gummadi  
MPI-SWS, Germany

## ABSTRACT

Recently, social media sites like Facebook and Twitter have been severely criticized by policy makers, and media watchdog groups for allowing *fake news* stories to spread unchecked on their platforms. In response, these sites are encouraging their users to report any news story they encounter on the site, which they perceive as fake. Stories that are reported as fake by a large number of users are prioritized for *fact checking* by (human) experts at fact checking organizations like Snopes and PolitiFact. *Thus, social media sites today are relying on their users' perceptions of the truthfulness of news stories to select stories to fact check.*

However, few studies have focused on understanding how users perceive truth in news stories, or how biases in their perceptions might affect current strategies to detect and label fake news stories. To this end, we present an in-depth analysis on users' perceptions of truth in news stories. Specifically, we analyze users' truth perception biases for 150 stories fact checked by Snopes. Based on their ground truth and the truth value perceived by users, we can classify the stories into four categories – (i) **C1**: false stories perceived as false by most users, (ii) **C2**: true stories perceived as false by most users, (iii) **C3**: false stories perceived as true by most users, and (iv) **C4**: true stories perceived as true by most users.

The stories that are likely to be reported (flagged) for fact checking are from the two classes C1 and C2 that have the lowest perceived truth levels. We argue that there is little to be gained by fact checking stories from C1 whose truth value is correctly perceived by most users. Although stories in C2 reveal the cynicism of users about true stories, social media sites presently do not explicitly mark them as true to resolve the confusion.

On the contrary, stories in C3 are false stories, yet perceived as true by most users. Arguably, these stories are *more damaging* than C1 because the truth values of the the story in former situation is incorrectly perceived while truth values of the latter is correctly perceived. Nevertheless, the stories in C1 is likely to be fact checked with greater priority than the stories in C3! In fact, in today's social media sites, the higher the gullibility of users towards believing a false story, the less likely it is to be reported for fact checking.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

FAT\* '19, January 29–31, 2019, Atlanta, GA, USA

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6125-5/19/01.

<https://doi.org/10.1145/3287560.3287581>

In summary, we make the following contributions in this work.

**1. Methodological:** We develop a novel method for assessing users' truth perceptions of news stories. We design a test for users to *rapidly* assess (*i.e.*, at the rate of a few seconds per story) how truthful or untruthful the claims in a news story are. We then conduct our truth perception tests on-line and gather truth perceptions of 100 US-based Amazon Mechanical Turk workers for each story.

**2. Empirical:** Our exploratory analysis of users' truth perceptions reveal several interesting insights. For instance, (i) for many stories, the collective wisdom of the crowd (average truth rating) differs significantly from the actual truth of the story, *i.e.*, wisdom of crowds is inaccurate, (ii) across different stories, we find evidence for both false positive perception bias (*i.e.*, a gullible user perceiving the story to be more true than it is in reality) and false negative perception bias (*i.e.*, a cynical user perceiving a story to be more false than it is in reality), and (iii) users' political ideologies influence their truth perceptions for the most controversial stories, it is frequently the result of users' political ideologies influencing their truth perceptions.

**3. Practical:** Based on our observations, we call for prioritizing stories to fact check in order to achieve the following three important goals: (i) Remove false news stories from circulation, (ii) Correct the misperception of the users, and (iii) Decrease the disagreement between different users' perceptions of truth.

Finally, we provide strategies which utilize users' truth perceptions (and predictive analysis of their biases) to achieve the three goals stated above while prioritizing stories for fact checking. The full paper is available at: <https://bit.ly/2T7raFO>

**Acknowledgments:** This research was supported in part by an ERC Advanced Grant for the project "Foundations for Fair Social Computing", funded under the European Union's Horizon 2020 Framework Programme (grant agreement no. 789373). E. Redmiles is supported by NSF Graduate Research Fellowship (Grant No. DGE 1322106) and a fellowship from Facebook.

## CCS CONCEPTS

• **Human-centered computing** Social networking sites;

## KEYWORDS

Truth Perception Bias, False News, Fact Checking

### ACM Reference Format:

Mahmoudreza Babaei, Abhijnan Chakraborty, Juhi Kulshrestha, Elissa M. Redmiles, Meeyoung Cha, and Krishna P. Gummadi. 2019. Analyzing Biases in Perception of Truth in News Stories and Their Implications for Fact Checking. In *FAT\* '19: Conference on Fairness, Accountability, and Transparency (FAT\* '19)*, January 29–31, 2019, Atlanta, GA, USA. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3287560.3287581>