Modeling the Impact of Fake News on Citizens

Stephanie Tulk (stulk@GMU.Edu)
Department of Psychology, George Mason University
Fairfax, VA 22030 USA

Niloofar Bagheri Jebelli (nbagher2@GMU.Edu)
Computational Social Science Program, George Mason University
Fairfax, VA 22030 USA

William G. Kennedy (wkennedy@GMU.Edu)
Center for Social Complexity, George Mason University
Fairfax, VA 22030 USA

Abstract

The impact of "fake news" on the 2016 presidential election became a serious concern after the surprising results. The volume of fake news on social media, which people used as a serious news source, could have significantly affected voters' opinions. It is important to consider how social and cognitive processes were affected by this fake news to estimate the true impact of this computational propaganda technique. We built a cognitive model of a citizen deciding what to believe when encountering election stories on social media, eventually developing an opinion and using motivated reasoning to help determine which stories are true. Modeling 100 citizens, we assemble polls of the agents over the 9 months leading up to the election that replicates the qualitative characteristics of actual polls but leaves many questions outside the purview of cognitive modeling.

Keywords: Cognitive modeling; opinion modeling; motivated reasoning; computational social science.

Introduction

After the 2016 presidential election in the United States, "fake news" became a topic of concern due to a surge of false news articles viewed and shared on social media in the months leading up to the election. Further investigation revealed that engagement with fake news articles on Facebook outnumbered those of real news (Silverman 2016) and the shared fake news was much more likely to be pro-Trump (the Republican candidate) than pro-Clinton (the Democrat) (Allcott & Gentzkow, 2017). It is worth considering that rather than being insignificant noise, the biased fake news may have influenced voters enough to impact the election results. This new misinformation approach to influencing public opinion and its effectiveness make up an interesting cognitive phenomenon.

We investigated the effect of fake news on people's ability to process news information with their cognitive limitations and biases. Our approach was to build a cognitive model of an individual considering the apparent evidence and deciding who to support during 9 months leading up to the decisive election. We exposed the many copies of the model representing the population and its diversity to input representing the variety of reportedly news items. Over the period, we polled our synthesized electorate and compared the modeled population to available survey data. We start with a discussion of the data available for this phenomenon.

Data on News and Fake News

The fake news surrounding both political candidates originated from a variety of sources within and outside the US (Allcott & Gentzkow, 2017). Once these fake articles were published, they required social actors to promote the material to their peers (tweets/re-tweets on twitter or shares on Facebook). Researchers at Indiana University, Bloomington concluded that "bots", or algorithms that pose as social agents on social media, played a key role in the dissemination of fake news by tweeting and retweeting misinformation to promote some "news" items to a wider audience, and tagging popular users to project the appearance of relevance and legitimacy (Shao et al., 2017). With the help of these bots, any fake news article may go viral and possibly reach a large population of voting Americans. Importantly, human users consistently retweeted misinformation from bots, increasing the reach of fake articles.

In the last three months leading up to the election, shares of fake news articles on Facebook outnumbered those of real news (Silverman, 2016), see Table 1.

More specifically, shares of fake news articles were nearly four times more likely to be anti-Clinton/ pro-Trump than anti-Trump/pro-Clinton (Allcott & Gentzkow, 2017). This contrasts the coverage of real news, where overall the amount of pro-Clinton/anti-Trump articles slightly outnumbered pro-Trump/anti-Clinton articles (Patterson, 2016); see Table 2.

By compiling a database of real and fake articles in the three months leading up to the election and testing the average American's recollection of major headlines, Allcott and Gentzkow (2017) estimated that the average adult US citizen actually read and remembered 1.14 fake news articles. This does not account for the exposure to headlines and thumbnails that may operate more like political advertisements. In surveys conducted shortly after the election, those who considered Facebook a major source of news reported believing 83% of fake news headlines they

remembered seeing over the course of the election (Silverman & Singer-Vine, 2016).

Table 1: Engagement for Top 20 Election Stories on Facebook in 2016. Engagement refers to shares, reactions and comments on stories (Silverman, 2016).

	Feb-April	May-July	Aug-Nov 8
Mainstream News	12 mil	9 mil	7.3 mil
Fake News	3 mil	3 mil	8.7 mil

The ability to induce belief in misinformation is an important effect to investigate further, although it is difficult to quantify. Shao and colleagues (2017) reported a weak correlation between activity of bots claiming to be residents of individual states and differences between actual and predicted vote margins in the 2016 election for these states. While they were careful to note that this is in no way conclusive evidence that bots actually impacted the election results, it provides an opportunity to explore the subtle and unseen impact of bot-generated attitude change.

Cognitive Model Foundations

We started with the standard ACT-R cognitive architecture (Anderson, 2009) as a general architecture supporting the modeling an individual's cognition while performing the task of deciding how to interpret news items and developing a political opinion. We begin by considering the social factors that are likely to play a key role in developing an individual's political identity. After an individual holds firmly to a belief system, biased cognitive processes will lead to strengthened political identity and polarization over time (Lord et al., 1979). Our model specifically deals with how opinions are developed with repeated exposure, and strengthened as a result of motivated reasoning and cognitive biases (cognitive dissonance, belief bias and propaganda effect). We wanted to examine what experiences lead to the acceptance of fake news as factual information, as well as radicalization that can occur once an individual has a strong enough opinion.

Even when exposure to individual fake stories does not result in the adoption of an explicit belief, the accumulated effect can impact individuals' social cognitive processes in a number of ways. Viral news (real or fake) not only reaches a wide audience; it carries some social influence by virtue of the importance people place on observing and conforming to the majority opinions (Kiesler & Kiesler, 1969). If something seems to have a viral audience, without knowing how many bots shared it, people think that there is some group consensus that the information is valid.

In the absence of unbiased fact-checking and simply overwhelming news coverage, individuals are left on their own to determine the validity of any claims they encounter. "Epistemic vigilance" refers to the cognitive mechanisms that exist to constantly monitor the potential that we are being misled by another party (Sperber et al., 2010). Sperber et al.

Table 2: Positive and Negative Coverage for 2016 Presidential Candidates (Allcott & Gentzkow, 2017; Patterson, 2016).

	Negative Real	Positive Real	Negative Fake	Real Coverage
D	64	36	79.8	46.5
R	77	23	20.2	53.5

to preserve the value of communication by ensuring that information recipient can detect and punish dishonesty from the communicator.

Unfortunately, our mechanisms for epistemic vigilance seem to be tested by our frequent use of social media. In the cyber world, we lose valuable information such as paralinguistic cues that people use to evaluate honesty in face-to-face communication (Littlepage & Pineault, 1978), and the overall volume of information in online platforms overwhelms our attentional capacities, making it difficult to filter information by quality (Qiu et al., 2017). The delivery of computational propaganda disguised as peer-disseminated information exploits our handicapped epistemic vigilance mechanisms and threatens the overall usefulness of modern internet-based news disseminating platforms. Without adequate fact-checking or epistemic vigilance mechanisms, misinformation will be perceived and interpreted with biased cognitive processes which significantly contribute to the bipartisan divide in the US. Our model decided validity using the following cognitive mechanisms:

(1) "Motivated reasoning" refers to the phenomena where human decision making is impacted by emotional factors such as "cognitive dissonance", or the feeling of mental discomfort produced when experiencing information that seems to disconfirm an individual's beliefs (Festinger, 1957). Researchers have shown that liberals and conservatives experience cognitive dissonance in the same ways, preferring to listen to arguments supporting or opposing politicians or partisan topics that are in line with their own partisan identity to avoid discomfort (Frimer, Skitka & Motyl, 2017).

(2) The "belief bias" is the tendency for individuals to accept an argument as true if they think it is believable and in line with their identity rather than assessing the quality of the argument (Evans, Newstead & Byrne, 1993). When interpreting partisan arguments, for example, the belief bias would lead individuals to more easily believe information that is in line with their partisan identity completely independent of the quality of information. When surveyed, democrats and republicans were both about 15% more likely to believe

headlines of articles that appeared on social media in the three months leading up to the 2016 election when they fit with their partisan opinions (Allcott & Gentzkow, 2017), and this pattern could be stronger for people further isolated in partisan echo chambers.

(3) The "propaganda effect" refers to the tendency for people to rate information as more believable after they have been exposed to it previously (Begg et al., 1992). This effect works implicitly when the individual does not consciously remember the exposure, therefore highly partisan propaganda can leave an impression even if it is not consciously believed.

Cognitive Model Details

We use the standard, off-the-shelf ACT-R, version 7.3, without modifications to the architecture. The non-default parameters used turned on sub-symbolic computation (:esc T) and rule utility learning (:ul T) with a relatively low level of noise (:egs 0.25). We used the default for memory retrieval threshold.

Our model was based on the "choice" model of the ACT-R tutorial, which reads input presented on the screen and compares that information to memory to make a decision. Our model uses 12 rules to read and process the news inputs. A block diagram of the model is shown in Figure 1.

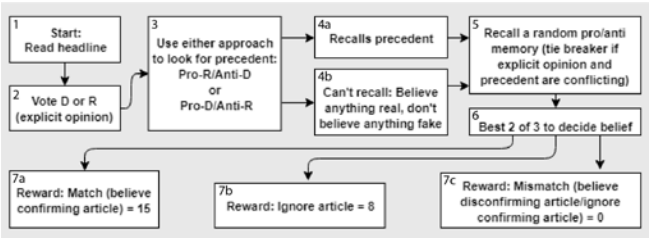

Figure 1. Citizen Model Block Diagram

The model begins a cycle by reading input in the form of a news item (fake or real, positive or negative, R or D), which the model perceives as a "chunk". Each news item is eventually either believed or disbelieved based on three pieces of information derived from proceeding productions: (1) The model expresses its explicit opinion in the form of an R or D vote based on the data it has collected since the first cycle. (2) Next, the model looks for a memory of a previous article that matches the current news item in order to find a belief precedent, deciding whether it wants to search for something that was previously believed or not believed given a specific candidate. If there is a retrieval error, the default precedent is that real news is believable and fake news is not. (3) A random exposure of positive or negative statements about a candidate is recalled in order to make a decision when the explicit opinion and belief precedent conflict. Initially, the modeled citizen had a random preference but it forms stronger opinions of the candidates over repeated exposure to election stories.

Rule utility learning was a part of the sub-symbolic representation of procedural knowledge of ACT-R. These rules became strengthened through use. Utility is considered to be a measure of the rule's value (Anderson, 2009). Models use rules with the highest utility. We used the same mechanism in our model to reward the results based on its utility. When the model ran, the production with the highest utility fired. The production firing had a reward value assigned to it. A reward value is propagated backwards through previous rule firings and depreciated by time.

To develop motivated reasoning, the modeled citizen experienced the highest reward when it was able to decide to believe an article that confirmed its explicit opinion (the vote; either pro-R/anti-D or pro-D/anti-R), resulting in a "match", and a medium reward if it could "ignore" disconfirming information. This was only possible if the citizen was able to recall a similar type of article that fit with the motivation (article precedent; belief bias). If the explicit opinion and the article precedent were contradictory, the citizen would make a "gut decision" based on some implicitly recalled information (propaganda effect). If the citizen was still unable to confirm their belief or ignore disconfirming information, it had to accept that the disconfirming information was true (attending a dissonant belief, resulting in 0 reward), resulting in a "mismatch". As the citizen was continuously exposed to news, it could attempt to reduce cognitive dissonance either by changing its explicit opinion about the candidates or by learning to strongly favor productions that were more likely to lead to matches and ignores.

Experiment

The experiment consisted of running a cognitive model representing an American citizen who was exposed to 1,000 news items over the last 9 months of the 2016 campaign cycle. Of those, 900 were randomly presented at the rates reported in Tables 1 and 2, with three phases of real and fake news ratios. To ensure that simulated citizens had some standardized experiences over time, an additional 100 real items that were presented at time points corresponding to some of the most influential real campaign stories. We recorded data over 100 runs of the model. The number of runs was arbitrary but intended to produce enough data to be useful.

Over the course of the experiment, the modeled citizen decided to believe or disbelieve each article, storing a memory of the event and slowly learning what candidate to support and what productions led to confirmation of that support and the least cognitive dissonance. Once the modeled citizen had a strong enough opinion about either candidate, it could partially or altogether stop believing real news that contradicted its current belief or believe fake news that confirmed its belief. This allowed us to investigate how motivated reasoning can affect truth-seeking behavior over

an election, possibly resulting in the adoption of radical opinions that are immune to the influence of facts.

In addition to the experiment described up until this point (which we will refer to in later sections as the "Troll" Condition), we also ran a version of the experiment in which the rate of fake news was kept constant over the 9 months (Feb-April in Table 1) and was equally disparaging of each candidate (deemed the "No Troll" condition), as well as a "No Troll" version that also had equal coverage of each candidate in real news ("No Troll/Equal Coverage"). These additional versions were added as control conditions against which we could compare the Troll results.

Finally, we ran 5 additional versions of the Troll condition to compare the rate of belief in real and fake news over time when the initial rate of belief in real ranged from 100% to 50% and fake ranged from 0% to 50%, in 10% intervals.

Experimental Results and Available Data

The experimental results are presented in two forms: as if our modeled citizens were regularly polled for their opinions and how the internal states of the model were shaped over time.

Polling trends were created by plotting the average explicit opinions over time for the 100 modeled citizens. Our model was able to reproduce polling trends by setting reward parameters and noise associated with learning such that a small amount of individual runs waffled between candidates on a significant amount of runs, but many explicitly supported the same candidate for the majority of trials. See Figure 2. See Figure 3 for comparison. The polling results for all 3 model conditions are compared to the actual electorate polling in Table 3.

The first internal factor we looked at was belief in real and fake news over time. Belief in fake news overtime was shown to increase. This is plotted as Figure 4.

Interestingly, the baseline model never disbelieved real news even though this was a possible (yet unlikely) outcome. To explore this further, the initial base rate of belief in real and fake news (retrieval failure, see box 4b in Figure 1) was varied from 100% to 50%. Results are shown in Figure 5.

While we explicitly set rewards to represent motivated reasoning, our model was still forced to observe disconfirming information (mismatch), thereby initially maintaining a solid representation of reality. However, the competition for attentional resources in the Troll condition resulted in decreased absorption of true information in the later months; see Figure 6.

Discussion

The issue of "fake news" had been a source of humor, but it now appears that fake news can affect the public's understanding enough to possibly change the outcome of a presidential election. The data on the frequency and type of fake news items circulated by the social media prior to the 2016 election was enough to cause our cognitive model of a US citizen to change the outcome of an election when averaged over 100 runs.

Our model shows that with the help of motivated reasoning, repeated exposure to large amounts of fake news results in competition for attentional resources that reduces the rate of

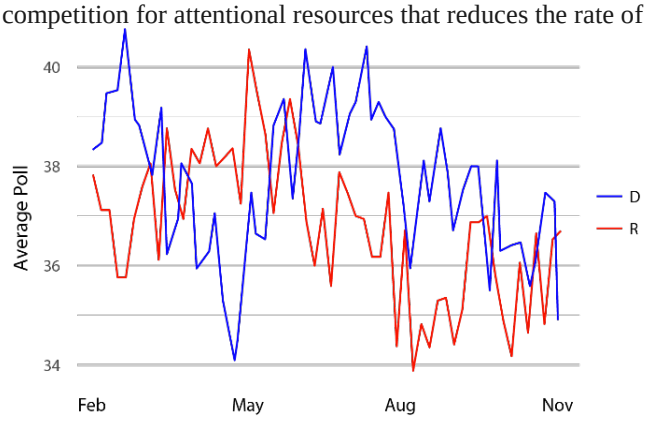

Figure 2. Simulated Polling Results Prior to the Election (averaged over 100 model runs) in the Troll Condition (Error bars are not shown due to variability of the mean).

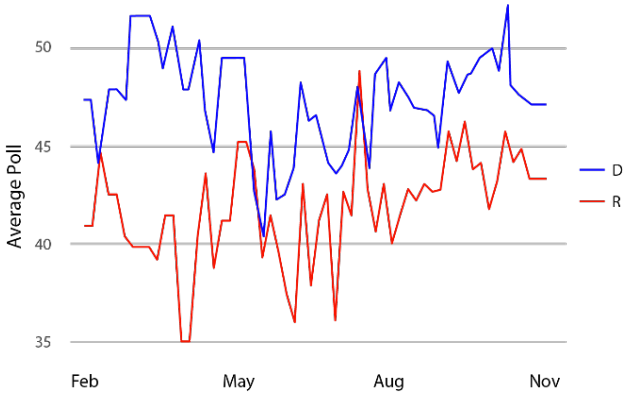

Figure 3. Electorate Opinion Polling prior to Election (Election G., 2016).

Table 3: True and Model Polling Results for Troll, No Troll, and No Troll/Equal Coverage. * indicates winner.

		Real Poll	Troll Poll	No Troll Poll	No Troll/Eq Cov Poll
Overall Average	D	**46.3 ***	**37.9 ***	*38.1 **	**41.3 ***
	R	41.7	36.8	*38.1 **	34.6
Last 5 Days	D	**46.3 ***	**36.5 ***	**38.9 ***	**42.7 ***
	R	42.7	35.8	38.6	33.9
Election Day	D	**48.2 ***	34.9	*39.5 **	**45.3 ***
	R	46.1	**36.7 ***	*39.5 **	34.5

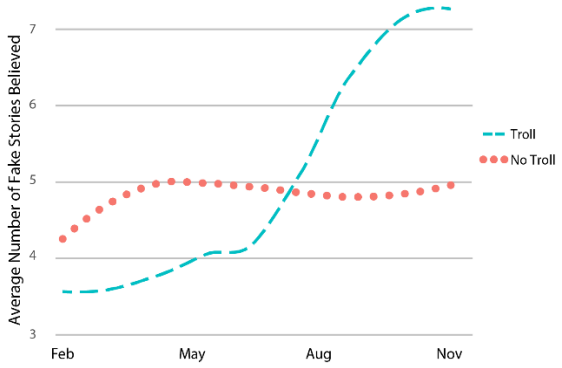

Figure 4. Modeled Citizen's Belief in Fake News in Troll and No Troll Conditions (averaged over 100 runs). Belief in fake news increases over time in the Troll condition.

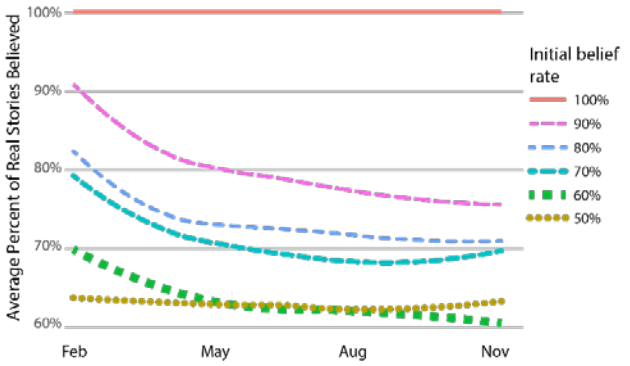

Figure 5. Average Percent of Real Stories Believed over 9 Months Beginning with Different Base Rates of Belief. Results are from the Troll Condition.

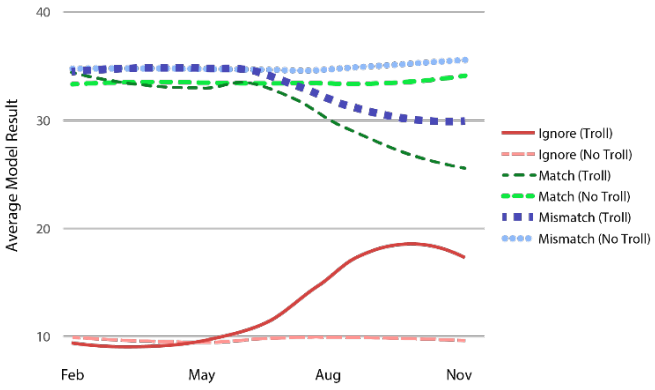

Figure 6. Model Results for Troll and No Troll Conditions. The increase in fake news results in a decline in the ability to process real information.

absorption of true information and increases the amount of fake news that is believed. This makes some sense, and the message is that a person's capability to process truth and update an opinion is hampered by the influx of fake news. Additionally, while there were more real anti-R/pro-D stories overall, the adoption of biases that were explicit (increased belief in fake news) and implicit (propaganda effect) against the heavily trolled candidate seemed to drive down the candidate's popularity in the Troll condition. Still, the impact of more real coverage for the R candidate also seemed to create more popularity for the R candidate in a type of "No press is bad press" fashion (see Table 3). While it would be encouraging to believe that real people never start to doubt true information, it is likely that people do not begin to develop an opinion with 100% truth detecting accuracy, and therefore some immunity to the truth can occur over time. The most dramatic impact occurs for those individuals who have the weakest discriminatory power before developing political bias (see Figure 5).

While our model was able to produce polling results that fit relatively well with the true polls (see Figure 2 vs. 3), there were a few limitations. We modeled our citizen to process about 10 news items daily, spread evenly over 24 hours per day, for the 9 months prior to the election. Our model begins with no prior political identity or opinion of either candidate. The average real citizen would likely have had some political identity before the 2016 election cycle, which tends to lead people to surround themselves with like-minded individuals, which would have affected their true rate of exposure to partisan stories (real and fake). Additionally, our model only understood these stories as simplified chunks, not paying attention to the language of a headline, the user who posted it or the source that published it, which are all pieces of information that would enter into the consideration of validity. Future work will seek to address these processes.

This work is an example of the type of modeling possible in the field of computational social science, where models of individual agents reacting to their environment and other agents can demonstrate possible macro-level results from relatively simple micro-level agents. Combining cognitive modeling and computational social science improves the credibility of results.

Acknowledgments

This work was supported in part by the Center for Social Complexity at George Mason University. The opinions, findings and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the sponsors.

References

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

Anderson, J. R. (2009). *How can the human mind occur in the physical universe?*. Oxford University Press.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, *121*(4), 446.

Election, G. Trump vs. Clinton. *RealClear Politics. URL: http://www.realclearpolitics.com/epolls/2016/president/us/general_election_trump_vs_clinton-5491. html*.

Evans, J. S. B., Newstead, S. E., & Byrne, R. M. (1993). *Human reasoning: The psychology of deduction*. Psychology Press.

Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Stanford university press.

Frimer, J. A., Skitka, L. J., & Motyl, M. (2017). Liberals and conservatives are similarly motivated to avoid exposure to one another's opinions. *Journal of Experimental Social Psychology*, *72*, 1-12.

Kiesler, C. A., & Kiesler, S. B. (1969). *Conformity*. Addison Wesley Publishing Company.

Littlepage, G., & Pineault, T. (1978). Verbal, facial, and paralinguistic cues to the detection of truth and lying. *Personality and Social Psychology Bulletin*, *4*(3), 461-464.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology*, *37*(11), 2098.

Patterson, T. E. (2016). News coverage of the 2016 general election: How the press failed the voters.

Qiu, X., Oliveira, D. F., Shirazi, A. S., Flammini, A., & Menczer, F. (2017). Limited individual attention and online virality of low-quality information. *Nature Human Behaviour*, *1*(7), 0132.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. *arXiv preprint arXiv:1707.07592*.

Silverman, C., & Singer-Vine, J. (2016). Most Americans who see fake news believe it, new survey says. *BuzzFeed News*.

Silverman, C. (2016). This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed News*.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*(4), 359-393.