Exploring the Phenomenon of Fake News with Implicit Measurements

Syed Muhammad Omar

Department of Psychology

University of Florida

Senior Thesis – Spring 2019

Thesis Committee:

Dr. Colin Tucker Smith (Chair)

Dr. Kate A. Ratliff

Dr. Natalie C. Ebner
Abstract

Fake news – fraudulent information which mimics the formation of online news media content – has been part of the national conversation since the 2016 US elections. Recent studies have implicated automatic processes, as opposed to controlled, in the perceived accuracy of fake news. This study explored the association between blatantly false news headlines and implicit truth evaluations by using the Implicit Association Test (IAT). Participants were either exposed to fake news headline posts with warning tags or real new headline posts with no tags, or not exposed to any stimuli. We then measured participants’ implicit truth evaluations and self-reported perceived accuracies of real and fake headlines. After controlling for political orientation, participants with prior exposure to fake news headlines had lower implicit truth evaluations and decreased perceived accuracy for fake news. This implies that prior exposure of fake news headlines with warning tags may have had a direct effect on implicit evaluations but participants were able to reject the proposition while reporting their perceived accuracy of fake news headlines.

Word Count: 168.
Exploring the Phenomenon of Fake News with Implicit Measurements

Since the 2016 American elections, fake news has received a lot attention and has become part of the colloquial language (Shane, 2017). In a meta-analysis of previous usage of the term “fake news”, two domains have been identified: facticity of the information in the news article (e.g., whether the information is based on facts) and intention of the publication (e.g., for financial gain, or political persuasion; Edson, 2018). For the purposes of this study, fake news was defined as, “fabricated information that mimics news media content in form but not in organizational process or intent” (Lazer et al., 2018, p. 2). While social media has been utilized by journalists to engage with the general public (Lasorsa, Lewis, & Holton, 2012), it has also taken a form of citizen journalism which has gone as far influencing the results of the 2016 American elections (Wall, 2015; Albright, 2016). The average American adult saw an average of one to three pieces of fake news in the month before the elections in 2016 (Alcott & Matthew, 2017). Understanding the phenomenon of fake news - misinformation with an intent of mass deception – is one of the areas which is needed to be understood by psychologists in order to come up with ways to protect modern society. The purpose of the study is to advance the understanding of the phenomenon of fake news and empirically deduce the effect of fake news headlines on implicit processes.

**Illusory Truth Effect**

It is clear that repeated exposure to informational stimuli increases the likelihood of its perceived accuracy, a phenomenon that has been termed the “illusory truth effect” (Bacon, 1979; Dechene, Stahl, Hansen, & Wanke, 2010; Hasher, Goldstein, & Toppino, 1977). This effect can be explained using the concept of the availability heuristic, “frequency or probability by the ease with which instances or associations could be brought to mind.” (Tversky, & Kahneman, 1973,
p.208). Such an effect was observed by Fazio et al. (2015) where participants judged false trivia items to be true after repeated exposure to false but plausible statements. Repeated exposure of plausible fake news increased the perceived accuracy by participants, even when the fake news went against their political orientation (Pennycook, Cannon, & Rand, 2018). Because of its pervasiveness and real-world impact, it is imperative to understand the cognitive faculties and processes involved in generating the illusory truth effect. The distinct cognitive processes underlying the formation and activation of implicit and explicit evaluations may help explain how fake news affects people’s judgments and beliefs. This study intended to extend the results from Pennycook et al. (2018)’s study to implicit measures.

**Dual-Process Theories and the Impact of Fake News**

Dual-process theories posit that human cognition involves two distinct processes related to analytical thinking and reasoning (Evans, 2003; Pennycook, Fugelsang, & Koehler, 2015). System 1 consists of processes which are intuitive, fast, automatic, and unconscious. System 2 is comprised of slow and controlled processes which are often related to analytical thinking. Social psychologists have implemented dual-process models in an effort to understand attitude formation and changes, and have theorized that implicit evaluations represent automatic processes whereas explicit evaluations represent controlled processes (Gawronski & Bodenhausen, 2011). Researchers have begun to utilize dual-process theories to explain the cognitive faculties involved in assessing the validity of fake news (Pennycook & Rand, 2018; Pennycook et al., 2018).

The Motivated System 2 Reasoning (MS2R; Kahan, 2013) account states that deliberately ruminating about some information which aligns with the individual’s political ideology increases the propensity for its perceived accuracy. In the case of fabricated news,
similarly, there is a higher propensity for perceived accuracy of fake news when individual engages in deliberative, analytical thinking and the fake news article is concordant with an individual’s political ideology (Kahan, 2017). Simply put, the MS2R account asserts that people who are analytical are more likely to have a polarized response to fake news when it aligns with personal political preferences and are more likely to perceive the fake news as accurate. On the other hand, according to the classical reasoning account, analytical thinking is deliberative and leads to a critical judgement of a situation (Kohlberg, 1969; Pennycook, Fugelsang, & Koehler, 2015; Swami, Voracek, Stieger, Tran, & Furnham, 2014). For fake news, Pennycook and Rand (2018) deduced that laziness or a lack of deliberative critical thinking (i.e., failing to engage System 2) correlates with perceived accuracy of fake news rather than personal political allegiances influencing and adding bias to the judgement. These findings provide support for the classical reasoning account rather than the M2SR account. This suggests that automatic, rather than controlled, processes are likely to underlie fake news’ illusory truth effect. So, we turn to the literature on automatic processes to explain how fake news might disproportionately affect automatic processes, which in turn influence subsequent beliefs and judgments.

The Impact of Fake News on Implicit and Explicit Evaluations

One of the models used to explain implicit-explicit dualism has been the associative-propositional evaluation (APE) model (Gawronski & Bodenhausen, 2011). According to the APE model, associative processes are represented by implicit evaluations, and can be understood as a person’s automatic, affective, and immediate gut reaction to an object. For example, upon seeing a box of candies, the immediate positive reaction due to previous memory of candy’s taste represents a positive implicit evaluation – the activated association is ‘candy + good’. Explicit evaluations are a result of propositional processes. The activated association may be accepted or
rejected based on other salient logical propositions, such as societal norms, education, and so on. For example, from the previous example, the same person might engage in deliberation of the impact candies on one’s health and end up rejecting the ‘candy + good’ association based on the proposition that “candies are unhealthy”. Implicit and explicit evaluations can be consistent or inconsistent depending on whether propositional processes accept or reject the propositions implied by the activated associations. In the candy example, after deliberation, the explicit evaluation is different than the implicit evaluation.

Research in implicit evaluations has shown that propositional and automatic processes work in cohesion to produce human behavior rather than working in isolation (Sherman, Gawronski, & Trope, 2014; Gawronski & Bodenhausen, 2006). Implicit and explicit evaluations can change in a congruent or incongruent manner depending on the extent to which a given factor directly or indirectly interacts with the cognitive processes. An external factor can directly influence and alter associative and propositional processes, leading to changes in corresponding implicit and explicit evaluations. On the other hand, an external factor can indirectly influence associative and propositional processes via direct activation of the respective other cognitive process (for example, direct activation of associative process leading to an indirect activation of propositional processes and the corresponding explicit evaluation).

For the purposes of this study, the APE model was utilized to expand upon the findings of Pennycook & Rand (2018) – susceptibility of fake news is due to lack of analytical thinking (i.e., underutilization of controlled System 2 processes) – and explored the impacts of implausible news on implicit attitudes. Based on the findings by Pennycook et al. (2018) and the implication of automatic System 1 processes, it was hypothesized that fabricated news would lead to either (a) changes in both implicit and explicit evaluations or (b) changes in implicit, but
not explicit, evaluations. These hypotheses are based on direct and indirect influences of external factors on association and propositional processes (see Gawronski & Bodenhausen, 2011, p. 80). In the first case, exposure to fake news (for example, ‘black lives matter (BLM) thug accidentally shoots himself’) causes a direct influence on associative processes by formation of new associations or activation of previously encoded associations in memory, and an indirect influence on propositional processes by validation of those newly formed or activated associations. This leads to a direct change in implicit evaluations and an indirect change to explicit evaluations – congruent responses in both evaluations (i.e. perceived accuracy of BLM fake news and a negative implicit evaluation of BLM activists). In the second case, exposure to fake news directly influences formation of new associations (or activation of previous associations) but the activated associations are rejected in the propositional processes, failing to indirectly influence propositional processes by activated associations. Upon measurement, this would lead to influence being measured on implicit truth evaluations but not on explicit perceived accuracy ratings.

**Hypothesis 1:** Real news headlines and blatantly fake headlines will both influence implicit truth evaluations – people will more easily associate familiar headlines with truth regardless of their content.

**Hypothesis 2:** Real and fake news headlines will both influence perceived accuracy – prior exposure to headlines will result in their being perceived as more accurate.

**Method**

**Participants**

Participants voluntarily participated in the study which was hosted on the Project Implicit website (https://implicit.harvard.edu; Nosek, 2005). Data collection was automatically
terminated once the target sample size of 880 was achieved. From the total sample, 682 participants remained after excluding those with errors in their IAT scores and those with missing data for their self-reported political orientation, fake news accuracy, and real news accuracy. Participants included in the data analysis were 33.9 years old (SD = 14.2) on average, 73.5% white, and 64.7% female.

**Procedure**

This study employed a modified version of the fake news and illusory truth study by Pennycook et al. (2018; Study 2, specifically); participants experienced four phases in the study: (i) familiarization, (ii) distraction, (iii) implicit measurement, and (iv) explicit measurement. After consenting to participate in the study, participants were assigned to one of the three experimental conditions in the familiarization phase; fake news condition, real news condition, or control condition (see Appendix for examples of stimuli). In the fake news and real news condition, participants were shown three news headlines posts (either fake or real; selected from Pennycook et al., 2018) in the format of Facebook newsfeed post (a headline, an associated picture, short description, and a website address). Participants in fake news condition were shown fabricated news headline posts accompanied by a “Disputed by 3rd Party Fact Checkers” warning tag and the real news condition participants were shown authentic news headline posts without any warning tags. To ensure participants engaged with the headlines, they were asked to rate the likelihood of them sharing the headlines on their personal social media account. The control condition omitted the familiarization phase and served as a means to establish baseline levels.
In the distraction phase, participants were asked to report their gender, voting behavior in 2016 elections, and party affiliation. Prior to the study, participants reported their age, sex, education level, race, ethnicity, and political orientation, while registering for Project Implicit.

After the distraction phase, participants moved on to the implicit measurement phase; participants sorted fake and real news headlines, factual statements, and nonsense statements, into ‘true’ and ‘false’ categories. Lastly, in the explicit measurement phase, participants rated headlines (all headlines, including real and fake, from familiarization phase; fake news headlines presented without 3rd Party warning tags) for familiarity and accuracy. At the end of the study, participants are presented the results of their implicit truth evaluations and debriefed.

**Measures**

**Implicit truth evaluation.** In order to conduct a measurement of associative processes involved in perception of news headlines, the Implicit Association Test (IAT; Greenwald, Schwartz, & McGhee, 1998) was utilized by this study. Participants were shown a table with three real news headlines, three fake news headlines, three known phrases with factual information (for example, “there are 60 seconds in a minute”; see Table 1), and three nonsense phrases which are illogical (for example, “The Earth is flat”; see Table 1), and then asked to familiarize themselves with the elements of the table. Real and fake news headline posts from the first phase were then shortened into single statements (e.g., “Trump reinstates military draft”; see Table 1 for the stimuli). In the IAT, participants sorted the fake and real news headlines (target categories), factual statements, and nonsense statements, into ‘true’ and ‘false’ attribute categories. IAT scores range from -2 to 2. Higher scores imply faster association of fake news headlines with true statements and/or real news headlines with false statements (i.e., lower
implicit truth evaluation) as compared fake news headlines with false statements and/or real news headlines with true statements (i.e., higher implicit truth evaluation).

<table>
<thead>
<tr>
<th>True</th>
<th>“there are 60 seconds in a minute”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“a wheel is round”</td>
</tr>
<tr>
<td></td>
<td>“A dog barks”</td>
</tr>
<tr>
<td>False</td>
<td>“There are 45 days in a month”</td>
</tr>
<tr>
<td></td>
<td>“The Earth is flat”</td>
</tr>
<tr>
<td></td>
<td>“A cow meows”</td>
</tr>
<tr>
<td>Real News</td>
<td>“Putin involved in U.S. hack”</td>
</tr>
<tr>
<td></td>
<td>“Trump vows to protect LGBTQ”</td>
</tr>
<tr>
<td></td>
<td>“Trump wants Muslim registration”</td>
</tr>
<tr>
<td>Fake News</td>
<td>“Trump reinstates military draft”</td>
</tr>
<tr>
<td></td>
<td>“Hillary was drunk after elections”</td>
</tr>
<tr>
<td></td>
<td>“Trump bans gay activity on TV”</td>
</tr>
</tbody>
</table>

Table 1. IAT category labels and stimuli.

**Familiarity with fake news headlines.** Participants from all three experimental conditions were shown three fake news headline posts (same ones as those shown to participants during familiarization phase in the fake news condition), without any warning tags, and they were asked to report their familiarity. For each item, participants were asked “How familiar are you with the headline above?”; participants responded between 1 (Not Familiar at all) to 10 (Extremely Familiar). Cronbach’s alphas were .86, .67, and .78 for the fake news, control, and real news conditions, respectively.
**Familiarity with real news headlines.** Participants from all three experimental conditions were shown three real news headline posts (same ones as those shown to participants during familiarization phase in the real news condition), without any warning tags, and they were asked to report their familiarity. For each item, participants were asked “How familiar are you with the headline above?”; participants responded between 1 (Not Familiar at all) to 10 (Extremely Familiar). Cronbach’s alphas were .81, .73, and .78 for the fake news, control, and real news conditions, respectively.

**Perceived accuracy of fake news headlines.** Participants from all three experimental conditions were shown three fake news headline posts (same ones as those shown to participants during familiarization phase in the fake news condition), without any warning tags, and they were asked to report their perceived accuracy. For each item, participants were asked “How would you rate the accuracy of the headline above?”; participants responded between 1 (Not Accurate at all) to 10 (Extremely Accurate). Cronbach’s alphas were .63, .64, and .71 for the fake news, control, and real news conditions, respectively.

**Perceived accuracy of real news headlines.** Participants from all three experimental conditions were shown three real news headline posts (same ones as those shown to participants during familiarization phase in the real news condition), without any warning tags, and they were asked to report their perceived accuracy. For each item, participants were asked “How would you rate the accuracy of the headline above?”; participants responded between 1 (Not Accurate at all) to 10 (Extremely Accurate). Cronbach’s alphas were .72, .68, and .73 for the fake news, control, and real news conditions, respectively.

**Results**
One-way ANOVAs were conducted to test for differences in IAT scores, perceived accuracy of fake news headlines, and perceived accuracy of real news headlines between the three experimental conditions (control, fake news exposure, and real news exposure). Analysis of variance between experimental condition and mean IAT scores concluded that there was a nonsignificant difference between them, $F(2, 679) = 2.18, p = .114$. Further, there were also nonsignificant differences between experimental condition and mean perceived accuracy of fake news headlines, $F(2, 679) = 2.72, p = .066$, and experimental condition and mean perceived accuracy of real news headlines, $F(2, 679) = 1.43, p = .239$. Next, ANCOVA tests for significance differences in IAT scores, perceived accuracy of fake news headlines, and perceived accuracy of real news headlines between the experimental conditions, after controlling for political orientation, were conducted (see Table 2 for $M$ and $SD$).

<table>
<thead>
<tr>
<th>Condition</th>
<th>$n$</th>
<th>IAT $M$</th>
<th>IAT $SD$</th>
<th>Fake headline accuracy $M$</th>
<th>Fake headline accuracy $SD$</th>
<th>Real headline accuracy $M$</th>
<th>Real headline accuracy $SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fake news</td>
<td>235</td>
<td>0.86</td>
<td>0.29</td>
<td>1.91</td>
<td>1.23</td>
<td>6.13</td>
<td>2.23</td>
</tr>
<tr>
<td>Real news</td>
<td>245</td>
<td>0.82</td>
<td>0.30</td>
<td>2.18</td>
<td>1.51</td>
<td>6.43</td>
<td>2.18</td>
</tr>
<tr>
<td>Control</td>
<td>202</td>
<td>0.80</td>
<td>0.30</td>
<td>2.16</td>
<td>1.39</td>
<td>6.41</td>
<td>2.06</td>
</tr>
</tbody>
</table>

Table 2. Scores for each condition after controlling for political orientation

An ANCOVA tested for significant differences in IAT scores by experimental condition after controlling for political orientation. There was a marginally significant difference in IAT scores between experimental conditions, $F(2, 678) = 2.578, p = .077$. Post-hoc comparisons indicated that participants in the fake news condition had significantly higher IAT scores than those in the control condition, $p = .018$, Cohen’s $d = 0.23$. In other words, participants more strongly associated fake news headlines with true statements than real news with true statements, when they were familiarized with fake news headlines with warning tags before the test. There
were no significant differences between people in the fake news and real news conditions ($p = .130, d = 0.14$), nor between people in real news and control conditions ($p = .352, d = 0.09$).

An ANCOVA also tested for significant differences in perceived accuracy of fake news headlines and experimental condition after controlling for political orientation. There was a significant difference in perceived accuracy of fake news headlines between experimental conditions, $F(2, 678) = 3.347, p = .036$. Post-hoc comparisons indicated that participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>$SD$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. IAT scores</td>
<td>0.83</td>
<td>0.30</td>
<td>--</td>
<td>-11**</td>
<td>.07</td>
<td>.01</td>
<td>.07</td>
<td>.07</td>
</tr>
<tr>
<td>2. Fake news accuracy</td>
<td>2.07</td>
<td>1.39</td>
<td>-11**</td>
<td>--</td>
<td>-12**</td>
<td>.11**</td>
<td>-15**</td>
<td>-.08*</td>
</tr>
<tr>
<td>3. Real news accuracy</td>
<td>6.32</td>
<td>2.16</td>
<td>.07</td>
<td>-12**</td>
<td>--</td>
<td>.24**</td>
<td>.53**</td>
<td>.15**</td>
</tr>
<tr>
<td>4. Fake news familiarity</td>
<td>2.21</td>
<td>2.22</td>
<td>.01</td>
<td>.11**</td>
<td>.24**</td>
<td>--</td>
<td>.51**</td>
<td>-.06</td>
</tr>
<tr>
<td>5. Real news familiarity</td>
<td>4.74</td>
<td>2.71</td>
<td>.07</td>
<td>-.15**</td>
<td>.53**</td>
<td>.51**</td>
<td>--</td>
<td>.15**</td>
</tr>
<tr>
<td>6. Political orientation</td>
<td>1.11</td>
<td>1.45</td>
<td>.07</td>
<td>-.08*</td>
<td>.15**</td>
<td>-.06</td>
<td>.15**</td>
<td>--</td>
</tr>
</tbody>
</table>

* $p < 0.01$, ** $p < 0.05$

Table 3. Experiment 1 means, standard deviations, and correlations with confidence intervals. $M$ and SD represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. Correlations are from the total sample, not separated by experimental condition ($N = 682$).

familiarized with fake news headlines had a significantly lower perceived accuracy of fake news than people in the control condition, $p = .048, d = 0.20$; participants familiarized with fake news headlines with 3rd Party warning tags reported a lower perceived accuracy of fake news headlines as compared to those who were not familiarized with any news headlines. Further, participants familiarized with real news headlines showed a significantly higher perceived accuracy of fake news than those in the fake news condition, $p = .016, d = 0.22$. Prior exposure of real news headlines led to participants reporting a higher perceived accuracy of fake news than those who were exposed to fake news headline posts. There were no significant differences between participants in the real news and control conditions, $p = .739, d = 0.03$. 
Further, an ANCOVA tested for significant differences between perceived accuracy of real news headlines and experimental conditions, after controlling for political orientation, and revealed no significant differences between the constructs, $F(2, 678) = .855, p = .426$.

**General Discussion**

Fake news, which is defined as fraudulent information that mimics the formation of news media content, has been part of the national conversation since the 2016 American elections. Pennycook & Rand (2018) deduced that lack of deliberate, critical, thinking is associated with perceived accuracy of fake news; implicating automatic processes. We extended their findings to implicit truth evaluations by carrying out an experiment similar to Pennycook et al. (2018)’s study: participants were exposed to fake news headline posts with warning tags, real news headline posts, or were not exposed to any news headline posts in the control condition, and then they were tested for implicit truth evaluations and self-reported their perceived accuracy for news headline posts.

We did not find any significant differences between implicit truth evaluations and explicitly measured perceived accuracy of news headlines across conditions. However, after controlling for political orientation, we found partial support for our first hypothesis. Participants who were exposed to fake news headline posts, as compared to those who were not exposed to any headlines, more strongly associated fake news headlines with true statements (and/or real news headlines with false statements) than they associated real news headlines and true statements (and/or fake news headlines with false statements) on the IAT. This is aligned with previous research where effects of mere exposure can be observed with implicit measures (Van Dessel, Mertens, Smith, & De Houwer, 2018). Mere exposure effect on implicit evaluations was not observed when participants were exposed to real news headline posts. The controversial
nature of the content in fake news headlines may have been more interesting to the participants and was able to garner adequate attention for the effect to occur. Further, the fake news headline posts were presented with 3rd party checker tags which may have also contributed to capturing the participant’s attention due to their novel nature. Future studies could explore the effect of 3rd party checker tags by presenting fake news headlines to participants without the warning tags and comparing them with those with warning tags.

We did not find support for our second hypothesis but, rather, the results showed an effect in the opposite direction, after controlling for political orientation. The effect was only observed for perceived accuracy of fake news and not for perceived accuracy of real news. Participants who were exposed to fake news headline posts had a lower perceived accuracy of fake news headlines than those who were exposed to real news headlines and, as well as, those in the control condition. This is contrary to the illusory truth effect observed in Pennycook et al. (2018)’s study and it may have occurred due to repeated reminders about the stimuli’s validity in the familiarization phase and the implicit truth evaluation. Future studies can counterbalance the order of the implicit truth evaluation and the perceived accuracy of news headlines. Participants reporting higher perceived accuracy of fake news headlines without prior exposure to headlines with 3rd party warning tags is a representative of real world interactions with news headline posts on social media, where warning tags to invoke skepticism do not exist.

Participants in either real news or control condition did not have differences between their perceived accuracy of fake news. Further, the illusory truth effect was not observed for participants in the real news condition as their perceived accuracy of real news was not greater than those who were not exposed to any news headline posts. Across conditions, there was a positive correlation between familiarity of real news and perceived accuracy of real news. It may
be possible that participants were already familiar with the real news headlines and mental associations for the content of the headlines already existed prior to the experiment.

The purpose of the study was to extend past research on fake news headlines to implicit truth evaluations. Only prior exposure to fake news headline posts with warning tags led to lower implicit evaluations (i.e., higher IAT scores; more strongly associated fake + true, and/or real + false, than fake + false, and/or real + true) than those in the baseline condition. It was also observed that participants in fake news condition had lower perceived accuracy of real and fake news headlines (the explicit measures) than control condition. This implicit-explicit dualism is better understood by utilizing the APE model. Lower perceived accuracy of fake news headlines and lower implicit truth evaluations imply a rejection of the activated associations about the fake news headlines in the propositional phase. In other words, prior exposure to fake news headlines with warning tags may have caused a direct influence in the implicit evaluations but not the explicit. Whereas, lower perceived accuracy of real news headlines and lower implicit truth evaluations imply a failure to reject propositions about real news headlines. This means that exposure to fake news headlines with disputed tags may have had a direct influence on implicit evaluations and an indirect influence on explicit evaluations. The implication of implicit processes in perception of disputed headlines adds to our collective knowledge about evaluations of news headlines. Future studies should look into refining techniques to measure implicit truth evaluations and explain the phenomenon of headlines with no warning tags attached.

It is interesting to note that perceived accuracy of real news was lower for participants in fake news condition than other experimental conditions. Participants in fake news condition showed an overall skepticism towards all stimuli and tended to rate both real and fake headline posts as less accurate. Across conditions, perceived accuracy of fake news had a negative
correlation with familiarity and perceived accuracy of fake news. Participants who found fake news headline posts to be true reported less familiarity and lower accuracy of real news headline posts.

**Limitations**

There are a few limitations in the study design and the techniques employed which may have hindered the objectives of the study. The sample was attained through Project Implicit in which participants voluntarily access the website to learn about their implicit evaluations and to try out different variants of the IAT. Participants gathered from Project Implicit tend be relatively more educated and predominantly white, liberal, and female; this convenience sampling leads to low generalizability. The current sample had more errors in their IAT scores (i.e., errors during the sorting tasks between categories of the IAT or taking longer than the established standard of response) than what is usually recorded. Removing participants with missing data in constructs of interest and errors in IAT led to exclusion of 198 participants from the original sample target of 880. This relatively high rate of exclusion may lead to concerns regarding generalizability. Future study designs should look towards developing IATs which have lower occurrence of errors and to incorporate a larger sample size to account for potential loss of power.

Another limitation in generalizability of sample in the study was an underrepresentation of conservatives (<10% of participants identified as conservatives). We statistically controlled for political orientation to rule out any effects of having a disproportionate number of conservatives across conditions confounding the results. However, controlling for political orientation in our analysis only addresses this limitation in the data gathered for this study. A
potential improvement in future studies could be to oversample for conservatives to balance out the observed disproportionality.

While Pennycook at al. (2018) observed illusory truth effect in both politically concordant and discordant news headline posts, they also observed that conservatives and liberals, alike, were in general skeptical of politically discordant news headlines. In another study, Pennycook & Rand (2017) found that pro-Trump voters were worse at discerning between fake and real news as compared to pro-Clinton supporters. The six headlines utilized as stimuli (three real and three fake) in our study were related to Donald Trump and politically concordant with conservative ideology. It is possible that due to their ideologically partisan nature, the stimuli had different impact on liberals and conservatives; one of the major limitation of the study lies in not accounting for the ideological nature of news headlines (fake or real). A solution proposed for future studies includes balancing out the stimuli according to their political nature after running preliminary analysis in which potential news headline posts for the main study are rated for their ideological nature by participants across the spectrum. Self-identified conservative and liberal participants would rate ideologically concordant headline posts and neutral headlines (nonpartisan headlines; for example, NASA or sports related headlines) for perceived accuracy and be tested for their implicit truth evaluations. Further, even though the sample was predominantly from the United States, the stimuli presented was focused on American politics. This limitation can be addressed by only sampling for participants in the United States or screening participants for their familiarity with American politics.

Similarly, another limitation with the reliability of the perceived accuracy was that we did not account for the degree of plausibility for each news headline post. Participants, across conditions, rated certain headline posts as more or less accurate than others. For example,
‘Hillary Was Drunk, Got Physical With Mook and Podesta’ (fake news stimuli) was rated as less accurate than other fake news headlines – implying greater implausibility compared to other fake news headline posts. Similarly, ‘Donald Trump Vows to Protect LGBTQ Community’ (real news stimuli) was rated as less accurate than other real news headlines or was perceived as less plausible than other real news headlines. Pennycook and Rand (2017) observed that familiarity led to higher perceived accuracy was moderated by baseline plausibility, but only for fake news headlines. A solution would be to only include headline posts which have the same plausibility so that our design would only measure the effect prior exposure of real or fake news on accuracy.

Even though implicit evaluations recorded an illusory truth effect in participants exposed to fake news headline posts, IAT scores across conditions were in general positive. Positive scores imply that participants were faster in associating fake news headlines with true statements and/or real news headlines with false statements, as compared to associating fake news headlines with false statements and/or real news headlines with true statements. Although it is generally assumed that the zero-point of the IAT is meaningful, this is only the case when the meaning of the zero value is established empirically (Blanton & Jaccard, 2006). Unlike the Race IAT or other commonly used measures, we have no reason to expect that a zero score on our study’s IAT actually captures the midpoint between associations of truth vs. falsity. Comparing the relative differences between IAT scores across experimental conditions avoids the arbitrary zero-point problem so long as the measure is tapping in to the underlying psychological construct it is intended to (which is a separate issue).

Positive scores make sense when participants are exposed to fake news headlines posts before their implicit truth evaluations as we expect the illusory truth effect to occur. Similarly, it was expected that exposure to real news headlines would lead to implicit truth evaluation where
participants are faster associating real news headlines with true statements and/or fake news headlines with false statements than fake news headlines with true statements and/or real news headlines with false statements (i.e., a negative score on the IAT) but the outcome had a positive score. Participants in the control condition, who were not exposed to any news headlines, had the same implicit truth evaluations as the real news condition. Owing to the stigmatized nature of the social phenomenon of fake news, it may be possible that participants are experiencing a heightened level of awareness and they are more careful of making the wrong choice in order to preserve their self-esteem. Due to the overt cautiousness, participants may have had a much more difficult time, which requires more cognitive effort, associating real news headlines with true statements and/or fake news headlines with false statements, causing them to sort fake news headlines with true statements and/or real news headlines with false statements faster. The current design of the IAT assumes that if there is a stronger association between a fake news headline and trues, there may or may not be an association between real news headline and false statements. Also, another assumption is that there is a weaker associations in the opposite block (i.e., real news headline with true statements and/or fake news headline and false statements). In order to empirically establish the associations of each headline type with individual categories, future studies should utilize single-target IATs where fake or real headlines are associated with either true or false categories. Further, in order to address the suspected stigma surrounding fake news, participants could be asked to report their awareness of the social trend, but, there are concerns that thinking about it may exacerbate the effects.

**Conclusion**

This study extended past research on illusory truth in the phenomenon of fake news to implicit evaluations. Implicit truth evaluations and perceived accuracies were recorded after
exposure to either fake or real news headlines. After controlling for political orientation, the results of the study observed an illusory truth in implicit evaluations but only for participants exposed to fake news headlines with tags. Exposure to real news headlines did not have an impact on implicit truth evaluations nor the explicit measurements of perceived accuracy. Design limitations with the political nature and the degree of plausibility of the stimuli were highlighted and future study designs were proposed.
References


Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology, 44*, 59-127. [https://doi.org/10.1016/B978-0-12-385522-0.00020](https://doi.org/10.1016/B978-0-12-385522-0.00020)


doi:10.1080/21670811.2014.1002513
Appendix

Headlines Posts Used in Familiarization Phase

Fake News Headlines:

- **Election Night: Hillary Was Drunk, Got Physical With Mook and Podesta**
  According to Todd Kincannon of the Kincannon Show, he spoke to a CNN reporter about...
  
- **Trump on Revamping the Military: We’re Bringing Back the Draft**
  Trump unveiled his plan to 'make the military great again,' saying he intends to reinstate...

- **Trump to Ban All TV Shows that Promote Gay Activity Starting with Empire as President – The #1 Empowering Conscious Website In The World**
  "Disputed by 3rd Party Fact-Checkers"
Real News Headlines:

Vladimir Putin ‘personally involved’ in US hack, report claims
Russian president made key decisions in operation seen as revenge for past criticisms by Hillary Clinton, says NBC
THEGUARDIAN.COM

At GOP Convention Finale, Donald Trump Vows to Protect LGBTQ Community
Four years ago, Mitt Romney never uttered the word “gay,” much less the full acronym...
FORTUNE.COM

Donald Trump Says He’d ‘Absolutely’ Require Muslims to Register
Donald J. Trump, who earlier in the week said he was open to requiring Muslims in the United States to register in a database, said in Iowa on Thursday night that he "would..."