Accepted Manuscript

"This is fake news": Investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media

Jonas Colliander

Please cite this article as: Colliander J., "This is fake news": Investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media, *Computers in Human Behavior* (2019), doi: https://doi.org/10.1016/j.chb.2019.03.032.

"This is fake news": investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media

Author:
Jonas Colliander (corresponding author)
Assistant Professor
Center for Retailing
Stockholm School of Economics
P.O. Box 6501, SE-113 83, Stockholm, Sweden.
Telephone: 46-8-736 9797
e-mail: Jonas.colliander@hhs.se

"This is fake news": investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media

Abstract

This study examines the effects of conformity to others online when individuals respond to fake news. It finds that after exposure to others' comments critical of a fake news article, individuals' attitudes, propensity to make positive comments and intentions to share the fake news were lower than after exposure to others' comments supportive of a fake news article. Furthermore, this research finds that the use of a disclaimer from a social media company alerting individuals to the fact that the news might be fake does not lower individuals' attitudes, propensity to make positive comments and intentions to share the fake news as much as critical comments from other users.

Keywords: Fake news; Online disinformation; Conformity; Self-concept; Disclaimers

"This is fake news": investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media

1. Introduction

On December 4, 2016, 29-year old Edgar Maddison Welch fired a military style assault rifle inside the popular Washington D.C. Comet Ping Pong restaurant. Mr. Welch had set out to rescue children he believed were held there in a child abuse scheme led by Hillary Clinton. The theory, known as "Pizzagate", stemmed from unfounded but widespread online reports. Rather than finding any children, however, Mr. Welch found himself in handcuffs. He was convicted to four years in prison and later confessed in an interview with the New York Times that "the intel on this wasn't 100 percent."

Indeed, it wasn't. But the concern among researchers, journalists and politicians about the effects of online disinformation is. The problem is rampant. A recent study by the Pew Research Center revealed that 23% of Americans had knowingly or unknowingly shared a made-up news story (Pew Research Center, 2016). Furthermore, during the 2016 U.S. presidential election, the most popular made-up news stories were more widely shared on Facebook than the most popular authentic news stories (Silverman, 2016). Some commentators have even suggested that online disinformation played a deciding role in that election (e.g. Dewey, 2016; Parkinson, 2016; Read, 2016).

Online disinformation has been defined by Lazer et al. (2018) as "false information that is purposely spread to deceive people." (p. 2). As such, it overlaps with the definition of fake news, given by Allcott and Gentzkow (2017) as "news articles that are intentionally and verifiably false, and could mislead readers" (p. 213). Increasingly the topic of public debate, fake news has been investigated by researchers from a variety of angles. One is studies into the prevalence of the problem. For instance, Allcott and Gentzkow (2017) studied Americans' level of exposure to fake news during 2016 U.S. presidential election and which segments of the population that believed in them. In another example, Watanabe (2017) studied the spread of Russian disinformation in western news media during the Ukraine crisis. Another area of research is how fake news travel within social networks. For instance, Vasoughi, Roy and Aral (2018) investigated how false and true news spread online. A third stream of research into misinformation and fake news is that of corrections and debunking. Research into these areas have primarily investigated how misperceptions spread through disinformation can be reduced by statements of correction from various sources. Bode and Vraga (2018), for instance, studied how misperceptions spread by health disinformation in social media were reduced by the presentation of correct facts by either algorithms or other social media users. Nyhan and Reifler (2010), on the other hand, concluded that corrections often fail and sometimes increase misperceptions when certain ideological groups have been presented with political misinformation. In a meta-study, Chan et al. (2017), also concluded that more detailed debunking is positively correlated with a debunking effect.

This research is intended to add to the research on debunking disinformation and fake news. However, it takes a step back from the research mentioned above in that it investigates not the effects of presenting counterfactuals to fake news, but rather the effects of the far more common occurrence of simply pointing out to readers that *it is* fake news. Specifically, this research examines what effect it has on individuals exposed to fake news that other users take

a stand against the disinformation and identifies it as such through the comment function. Given the state of many comment threads to fake news, which are less about correcting the disinformation and more about simple statements either supporting or attacking both the story and the person posting it, the lack of research investigating the many facets of such simple statements is conspicuous. Therefore, the present research consists of two experimental studies on this phenomenon. Study 1 investigates whether readers of disinformation, upon seeing comments from others either supporting the fake news story or opposing it by attacking the news story or the original poster, respectively, are more likely to a) have a more positive or negative attitude towards the fake news b) make comments of either support or opposition to the fake news and c) share the fake news story on social media. Study 2 investigates the same behavior among respondents when they are exposed to other users' comment of either support or opposition to a fake news story. However, it also compares the effect of other users' comments identifying the news as fake with the use of an official Facebook disclaimer stating that the fake news story is disputed by independent fact checkers. Taken together, the two studies are intended to shine a few rays of light on the effects and importance of other users in preventing the spread of fake news online.

Beyond simply focusing on the simpler ways in which users in social media can debunk fake news this research makes two additional contributions. Firstly, it introduces additional dependent variables in the form of the attitude towards the fake news, the likelihood of commenting in various ways on the fake news and the sharing the fake news. Previous research on debunking have focused mainly on correcting the misconceptions (caused by disinformation) of respondents. Important as that may be, another stated goal of policymakers and social media companies alike is to stop the spread of fake news. This research is intended to help in that pursuit by investigating the three dependent variables mentioned above. Secondly, it explores whether other ways of disputing the accuracy of fake news, such as disclaimers from social media companies, has a similar effect on readers as the actions of other users.

Structurally, the rest of this article is straightforward. It begins with reviews of conformity and the self-concept, which are the theoretical foundations of this research. That is followed by the hypothesis development and a description of the two studies. Lastly, conclusions and implications for both researchers and practitioners are discussed.

2. Theoretical background and hypotheses

2.1 Conformity

Conformity is the act of matching one's behavior to the responses of others (Cialdini and Goldstein, 2004). Conformity has been found to be a powerful social phenomenon as individuals are often found to conform to the behaviors of others even when the actions of those other individuals run contrary to individuals own convictions, such as in the classic experiments by Asch (1956). Subsequent research has also demonstrated that even our memories are affected by exposure to the recollections of others (Edelson et al., 2011) Deutsch and Gerard (1955) made a distinction between informational and normative motivations for conformity. Informational motivations are driven by a desire to interpret reality in an accurate way whereas normative motivations are based on the desire to obtain social approval from others. More contemporary research has largely upheld these findings. The overview by Cialdini and Goldstein (2004), underscores the importance of conformity in gaining social approval, stating that "individuals often engage in … conscious and deliberate

attempts to gain the social approval of others, to build rewarding relationships with them, and in the process, to enhance their self-esteem. Conformity offers such an opportunity" (p. 610). Interestingly, Williams et al. (2000) concluded that conformity still occurs among anonymous internet users.

2.2 Self-concept

The self-concept is an individual's collection of beliefs about him or herself, generally answering the question of 'who am I'? (Meyers, 2009). Individuals tend to conceptualize themselves in accordance with two basic aspects of human beings: agency and communion (Wiggins, 1991). Agency represents such personal interests and values as self-assertion, self-improvement and self- esteem. Communion, conversely, is about social bonding, connections with others, cooperation and care for others (Nam et al., 2016). Agentic individuals are dispositioned to show a more self-centered behavior and focus on differentiating themselves from others. Communal individuals, on the other hand, are more likely to be a part of a group and form social connections (Wiggins, 1991). Cialdini and Trost (1998) state that all individuals share a strong need to enhance the self-concept. This is done by behaving consistently with their statements, actions, beliefs, commitments and self-ascribed traits. One of the ways in which this manifests itself is by the consumption by individuals of products that correspond with their self-concept as a means of self-expression (Braun et al., 2002). Another is the way individuals behave and write online in response to comments from other internet users (Colliander and Wien, 2013).

2.3 Hypotheses development

Here, it is proposed that due to conformity and the desire to maintain a positive self-concept, respondents who are exposed to comments identifying a fake news story as such have a more negative attitude towards the news story, are more likely to critically comment on the story and are less likely to spread it through their own social channels. Furthermore, it is proposed that due to the critical role of the self-concept, these tendencies are especially pronounced when the comments from other users include personal attacks on the poster of the original story. Several authors have documented the power of conformity in online behavior. Zhu and Huberman (2014), for instance, demonstrated that consumers tend to shift their preferences in an online setting when faced with the recommendations of others. Breitsohl, Wilcox-Jones and Harris (2015) found support for a groupthink mentality in online communities. Tsikerdekis (2013), meanwhile, found that conforming to the opinions of the group occurred irrespective of the levels of anonymity that users perceived themselves as having. Specifically investigating online news contexts, Winter, Bruckner and Krämer (2015) found evidence of the social influence of others comments when judging stories online. Other researchers have demonstrated that conformity extend beyond the mental dimension and affect actions online. In a comprehensive study involving the analysis of online discussion forums as well as four experiments, Hamilton, Schlosser and Chen (2017) found that commenting is significantly affected by the need for affiliation. Therefore, commenters online were likely to conform their writings to already existing comments.

Based on this body of evidence, it is likely that when people are exposed to comments critical of a fake news story (rather than supportive comments), they will gain a more negative attitude towards the fake news story, and will be more likely to themselves comment critically (rather than in a supportive manner). Moreover, Cialdini and Goldstein (2004) state that "people are frequently motivated to conform to others' beliefs and behaviors in order to

enhance, protect or repair their self-esteems" (p. 611). Colliander and Wien (2013), meanwhile, state that individuals' actions on social media is partly motivated by their desire to bolster their self-concepts. Following the importance of an individual's self-concept, as highlighted above, it is therefore likely that people are less inclined to share a fake news story after seeing others commenting critically on it. Furthermore, when exposed to comments shaming the original poster for sharing the fake news story, the threat to the self-concept inherent in sharing the fake news story ought to be especially salient to people. Therefore, they should be even less likely to share the fake news story after exposure to these comments than after exposure to comments simply claiming that the news story is fake. Given this reasoning, the following hypotheses are proposed:

H1: After exposure to a fake news story with user comments critical of the content, people will have a more negative attitude towards the fake news, than after exposure to the fake news story with user comments supportive of the content.

H2: After exposure to a fake news story with user comments critical of the content, people are more likely to make critical comments themselves, than after exposure to the fake news story with user comments supportive of the content.

H3a: After exposure to a fake news story with user comments identifying the news as fake, people are less likely to share the fake news, than after exposure to the fake news story with user comments supportive of the content.

H3b: After exposure to a fake news story with user comments shaming the poster for spreading said fake news, people are less likely to share the fake news story, than after exposure to the fake news story with either user comments supportive of the content or with user comments merely identifying the news as fake.

3. Study 1

Study 1 was intended to test H1, H2. H3a and H3b. To that end, we used a between-subjects experimental design with three treatment groups. One group of participants (group 1) were subjected to a fake news social media post with supportive comments from other users. The second experimental group (2) was subjected to the same fake news social media post but this time the comments critically identified the fake news as such. The third experimental group (3) was subjected to the same fake news social media post with comments both critically identifying the fake news as such and criticizing the poster of the fake news for spreading it.

3.1 Stimulus development

The study utilized a role-play scenario were participants were subjected to one of the three experimental posts embedded in a survey tool and instructed to imagine that they saw it posted by a distant acquaintance on Facebook. To maximize validity, it was decided that the study should employ a real piece of fake news. To that end, a search of the internet for known sources of fake news was undertaken. Eventually, it was decided to use a Facebook post from a page called "America's last line of defense". The page has been noted for solely spreading made up news by both The Washington Post (Saslow, 2018) and Politifact.com (Gillin, 2018).

Three criteria were used to pick the post to be used as stimuli for the study. The post had to a) reference an event or issue that was relevant and well known to a U.S. audience at the time of

the study b) be indisputably false and c) could reasonably be identified as false by an average individual. It was decided to use a fake story about the Austin serial bombings that took place between March 2 and March 20, 2018. After the perpetrator of those bombings committed suicide on March 21 and was subsequently identified, America's last line of defense published a post mimicking a news alert on March 22 that stated that the bomber had been "on Clinton Foundation payroll". The post thus implied that the bomber had been employed by Bill and Hillary Clinton's foundation, which is a frequent target of fake news. Using the criteria above, it was decided that the post met all three. It was demonstrably false and should be identifiable as such by an average person and due to extensive news coverage of the Austin events it was deemed relevant at the time of the study (early April 2018).

A screenshot of the post was taken to be used as the focal fake news story of the study. Next, three different comment sections were created. Photos were blurred and names were altered to create fictitious, non-identifiable individuals. Next, three sets of comments of four each were created to achieve the experimental stimuli. That number of comments was deemed appropriate to establish the desired pattern. Each comment feed was inspired by actual comments found on America's last line of defense and similar Facebook feeds. The comments for group one contained expressions such as "I knew it" and "Unbelievable". The comments for group two contained comments such as "Fake story!" and "This is fake news." The comments for group three contained comments such as "It's irresponsible of you to spread this untrue stuff" and "Shame on you for spreading this lie". Screenshots were taken of each comment section and each was merged with the screenshot of the fake news story, thus creating a stimulus for each of the experimental groups. A small focus group of four university students fluent in English at a western European business school was gathered to assess the stimuli. All participants were active on Facebook. Asked whether the posts looked like real Facebook posts they all answered in the affirmative. Likewise, all participants deemed the various comments as credible and representative of actual comments that they had encountered online. They all also judged the comments intended for group one as supportive of the post and the comments for groups two and three as critical of the post. Asked which one of the comment sections intended for group two and three that was most critical of the poster, all respondents indicated the one meant for group 3. It was thus determined that the stimuli were suitable for use in the study. Please see appendix 1 for the stimuli.

3.2 Data collection and participants

Each scenario version was followed by questionnaire items to measure the variables in the hypotheses. The scenarios were randomly allocated to participants (N =1201). Respondents were recruited through Amazon Mechanical Turk and consisted of US residents over the age of 18 who were members of Facebook. Research supports the validity of Amazon Mechanical Turk data within quantitative studies as compared to other methods for online survey data collection (Buhrmester, Kwang, and Gosling 2011). Participation was open to people who had a validated track record in past surveys of above 90% approval and the Qualtrics (the survey publishing tool adopted) anti-ballot stuffing setting was enabled to avoid multiple submissions from the same participant. 40% of respondents were male and the average age of the respondents was 37. There were no significant differences in gender (Chi2=.832) or age (p=.321) between our 3 experimental groups. After initially filling out demographic questions ensuring that they were in fact U.S. residents and members of Facebook, respondents were, as noted, instructed to "Please imagine that you the following post by a distant acquaintance on Facebook" and to look carefully at the post and comments and answer all questions.

3.3 Measures

Attitudes towards the fake news story was measured with three items on seven-point Likert scales (1= completely disagree, 7= completely agree): "My impression of the Facebook post is good", "My impression of the Facebook post is pleasant", "My impression of the Facebook post is favorable" (Colliander and Marder, 2018). Responses to the three items were averaged to form an index, Cronbach's alpha = .96).

The likelihood to make positive or negative comments on the fake news story was measured with a single variable: "If you would comment on this post, would your comment be mostly supportive or mostly critical of the content in the post?" Responses to this question were measured on a binary scale (Mostly critical/Mostly supportive).

The likelihood to share the post was measured with three items on seven-point Likert scales (1= completely disagree, 7= completely agree): "It is likely that I would share this post on Facebook", "It is possible that I would share this post on Facebook", "It is probable that I would share this post on Facebook" (Huang et al. 2011). Responses to the three items were averaged to form an index, Cronbach's alpha = .95).

3.4 Results

Before analyzing the dependent variables, an initial manipulation check was employed. Respondents were asked "Were the comments on the Facebook post that you could see supportive or critical of the post?" Responses to this question were measured on a binary scale (Supportive/Critical). Only respondents who correctly answered the question (N = 1164) were subsequently analyzed when testing the hypotheses.

When testing H1, that attitude towards the fake news story would be lower after reading one of the two sets of comments critical of the post, a one-way ANOVA with a Scheffe post-hoc test was employed. The results showed that the means of the attitudes towards the post was significantly lower in groups two and three than it was in group one. Thus, H1 was empirically supported. The same method for analysis was employed when testing H3a and H3b, that intentions to share the fake news story would be lower in the group subjected to the comments pointing out the news story as fake (group two) than in the group subjected to the comments supporting the fake news story (group one), and that these intentions would be lower still in the group subjected to comments critical of the poster (group three). The results showed that the means of both groups two and three were significantly lower than the means of group one. H3a was thus empirically supported. However, the means of group three was not significantly different from that of group two. H3b was thus only partially empirically supported. Please see table 1 for the all means, standard deviations and p values of the ANOVA tests.

Table 1: Mean values (standard deviations) for attitudes towards the post and intentions to share the post in study 1.

Variable	Mean supportive comments from users (group 1)	Mean comments pointing out that the news is fake (group 2)	Mean comments pointing out that the news is fake and critical of the poster (group 3)
Attitudes towards the post	2.18 (1.54)	1.64 (1.23)[a]	1.82 (1.45)[b]
Intentions to share the post	2.0 (1.68)	1.47 (1.09)[a]	1.62 (1.31)[b]

[a] = significantly lower than group 1 at p <.001

[b] = significantly lower than group 1 at p <.005

When testing H2, that respondents would be less likely to make comments supportive of the fake news story after reading comments of others critical of the story, than after reading comments of others supportive of the story, a cross tabulation with a chi-square test was employed. Results show that there was a significant difference ($p < .001$) between the expected proportions of respondents who would make critical and supportive comments, respectively, in the three experimental groups. Among respondents in group one, who saw comments supportive of the fake news story, more individuals than expected would themselves make comments supportive of the fake news story. Meanwhile, the reverse pattern emerged for groups two and three. Thus, H2 was empirically supported. Please see table 2 for the expected and actual count of respondents who would make supportive and critical comments, respectively.

Table 2: Expected and actual count of respondents who would make mostly supportive or mostly critical comments study 1

Group	Actual and expected counts	Mostly supportive comments	Mostly critical comments
Supportive comments from users (group 1)	Count	90	295
	Expected count	55.2	329.8
Comments pointing out that the news is fake (group 2)	Count	35	350
	Expected count	55.2	329.8
Comments pointing out that the news is fake and critical of the poster (group 3)	Count	42	352
	Expected count	56.5	337.5
	Total count	167	997
	Total expected count	167	997

3.5 Discussion

The results of study 1 show that the comments and actions of other users in social media can indeed affect the reactions to, and spread of, fake news online. Users exposed to comments by others users that were critical of the fake news had lower attitudes to the fake news, and were more likely to comment critically and share the fake news themselves, than users who were exposed to comments supportive of the fake news. These findings clearly demonstrate the potential and responsibility of ordinary readers in stopping the spread and mitigating the impact of fake news and online disinformation.

Theoretically, study 1 offered a mixed bag. In particular, the fact that H3b was only partially supported is interesting. Making a threat to the self-concept especially salient, by showing an individual the potential of being shamed by other users online when spreading a fake news story, did not affect respondents in this study more than when other users simply wrote that the story was false. This could be because the importance of maintaining the self-concept in an online setting has been overestimated in previous studies. Though, with the robust body of research indicating its importance, that seems unlikely. More probable is the fact that the simple pointing out that the story is fake is also seen as an implicit rebuke by other users online and that conformity and the potential threat to the self-concept act in combination in explaining the differences between the three experimental groups.

It could be argued, however, that neither conformity nor the threat to the self-concept were responsible for the results of study 1. Instead, one can argue that that those response patterns were simply due to a 'waking up' – effect. That is, research has demonstrated that people do not spend much time digesting content online (Weinreich et al, 2008), indicating that they do not spend much cognitive effort to process web content. This would indicate that at least some users might not think about the fact that a fake news article is fake and that it is only when seeing comment from other users that they realize that fact. This would be similar to the effect of incongruent advertising in drawing attention to certain commercial messages (e.g Dahlén et al., 2008). If that was the case, other stimuli that draws an individual's attention to the fact that a news story is fake would achieve similar effects to the comments used for groups two and three in this study. One such stimuli could be disclaimers from social media companies themselves. These are notifications informing users to pay attention to the content for some reason. When it comes to fake news, Facebook started using disclaimers in 2017 stating that the content is disputed by multiple fact-checkers (Hunt, 2017). That practice has since been replaced by a 'related stories'- function (Flynn, 2017) but including such a disclaimer would nevertheless alert readers to the fact that the news is fake. Thus, it would reveal if the results of study 1 are due to conformity or simply the alerting to consumers that the news is fake.

Putting this to the test a second study was launched. However, since there is potential for a number of different outcomes, no hypotheses will be formulated. After all, the results of study 1 could be due to conformity, as was argued leading up to study 1, and there is no 'waking up' – effect. Or, a disclaimer could achieve similar results to the comments in study 1. Therefore, an open research question was devised instead to test the effects of disclaimers vs. comments from others users pointing out that a news story is fake. Hence:

Research question 1: How does the disclosure that a news story is fake by disclaimers from a social media company compare to comments to that effect from other users in affecting

attitudes towards the news story, propensity to comment on the story in a supportive or critical manner and intentions to share the news story in social media?

4. Study 2

Like study 1, study 2 used a between-subjects experimental design. This time, it included 4 treatment groups. One group of participants (group 1) was subjected to a fake news social media post with no comments or disclaimers. The presence of this control group is intended to put the means of other experimental groups into context. Adding this group in study 2 is a further contribution of the second study. The second group (2) was subjected to a fake news social media post with supportive comments from other users. This was similar to the first group in study 1. The third group (3) was subjected to a fake news social media post with comments pointing out that the news was fake. This was similar to the second group in study 1. The fourth group (4) was subjected to a fake news social media post with comments supportive of the post and with a disclaimer stating its content was disputed by fact checkers. The attentive reader will notice that it was decided to use only one version of comments critical of the content, the ones where the commenters simply pointed out that the news post was fake. Finding no significant differences between groups two and three in study 1, and assuming that the comments stating that the news was fake served as an implicit rebuke as discussed above, this was done to avoid an overly cluttered study 2. In addition, it was decided to apply the disclaimer to a post with user comments supporting the fake news. This was done in order to directly compare groups 2, 3 and 4 to discern whether comments from other users stating that that news is fake or disclaimers from a social media company was more effective in mitigating the effects of other users' supportive comments found in study 1.

4.1 Stimulus development

The same criteria as in study 1 (reference an event or issue that was relevant and well known to a U.S. audience at the time of the study; be indisputably false; could reasonably be identified as false by an average individual) were uses when finding a fake news post to use in study 2. The same fake news feed was used to find a new suitable post. This time, the study ran in early October, 2018. Therefore, it was decided to use a post about the ad campaign that Nike ran with former San Francisco 49:ers quarterback Colin Kaepernick, which had been the topic of much debate in the weeks prior to the study. The athlete, who spearheaded a movement by NFL players to kneel during the national anthem in protest against racial injustice in the US, was seen as a controversial choice as a brand spokesperson. Nike drew much right-wing ire for their decision to use Kaepernick and there was a social media campaign started against the company. The post chosen for the study stated that Nike had filed for bankruptcy after the "failed Kaepernick campaign". It was decided that the post met the criteria. The names of the commenters were changed and their comments were altered to suit each stimulus. For added ecological validity, likes and emojis were retained on the comments. Thereafter, screenshots were taken of the post and adapted comments. The photos of commenters were blurred and, for group 4, a disclaimer taken from an older fake news story was photoshopped into the post. As in study 1, a small focus group consisting of English-speaking university students at a western European business school was gathered to pre-test the stimuli. After the focus group confirmed that the stimuli conveyed the intended messages it was decided to go ahead with the study. Please see appendix 2 for the stimuli in study 2.

4.2 Data collection and participants

Again, the study utilized a role-play scenario were participants were subjected to one of the six experimental posts embedded in a survey tool. They were instructed to imagine that they saw it posted by a distant acquaintance on Facebook. Each scenario version was followed by the same questionnaire items as in study 1. The scenarios were randomly allocated to participants (N =800). Again, respondents were recruited through Amazon Mechanical Turk and consisted of US residents over the age of 18 who were members of Facebook and who had a validated track record in past surveys of above 90% approval. The Qualtrics anti-ballot stuffing setting was enabled to avoid multiple submissions from the same participant. This time, 50% of respondents were male and the average age of the respondents was 36. There were no significant differences in gender (Chi2=.779) or age (p=.170) between our 4 experimental groups. After initially filling out demographic questions ensuring that they were in fact U.S. residents and members of Facebook, respondents were instructed to look carefully at the post and comments and answer all questions.

4.3 Results

Before analyzing the dependent variables, three manipulation checks were employed. As in study 1, respondents were asked "Were the comments on the Facebook post that you could see supportive or critical of the post?" Responses to this question were measured on a binary scale (Supportive/Critical). Secondly, they were asked "What brand was the focus of the Facebook post?" Respondents could choose between Nike/Adidas/New Balance. Thirdly, they were asked "Was there a disclaimer stating that the post had been disputed included in the post?" Responses to this question were measured on a binary scale (Yes/no). As in study 1, respondents were only asked these questions after answering the questions below but only respondents who correctly answered the questions (N = 506) were subsequently analyzed when testing the hypotheses.

In order to test the effects on attitude towards the post and intentions to share the post, one-way ANOVAs with Scheffe post-hoc tests were once more employed. Please see table 3 for the results of these tests.

Table 3: Mean values (standard deviations) for attitudes towards the post and intentions to share the post in study 2.

Variable	Mean no comments or disclaimers (group 1)	Mean supportive comments from users (group 2)	Mean comments pointing out that the news is fake (group 3)	Mean supportive comments from users and a disclaimer (group 4)
Attitudes towards the post	2.13 (1.58)	2.89 (2.27)[a]	1.82 (1.54)[b]	2.42 (1.98)
Intentions to share the post	1.99 (1.82)	2.45 (2.15)	1.62 (1.44)[c]	2.04 (1.88)

[a] = significantly higher than group 1 at p <.05

[b] = significantly lower than group 2 at p <.001

[c] = significantly lower than group 2 at p <.01

In order to test whether respondents would be likely to make comments supportive or critical of the post after seeing the experimental stimuli, a cross tabulation with a chi-square test was once more employed. Results show that there was a significant difference ($p < .001$) between the expected proportions of respondents who would make critical and supportive comments, respectively, in the four experimental groups. Please see table 5 for the expected and actual count of respondents who would make supportive and critical comments, respectively.

Table 4: Expected and actual count of respondents who would make mostly supportive or mostly critical comments study 2

Group	Actual and expected counts	Mostly supportive comments	Mostly critical comments
No comments or disclaimers (group 1)	Count	37	102
	Expected count	34.9	104.1
Supportive comments from users (group 2)	Count	41	67
	Expected count	27.1	80.9
Comments pointing out that the news is fake (group 3)	Count	12	131
	Expected count	35.9	107.1
Supportive comments from users and a disclaimer (group 4)	Count	37	79
	Expected count	29.1	86.9
	Total count	127	379
	Total expected count	127	379

4.4 Discussion

Study 2 was conducted to shine a light on research question 1. To reiterate, it was formulated as follows:

Research question 1: How does the disclosure that a news story is fake by disclaimers from a social media company compare to comments to that effect from other users in affecting attitudes towards the news story, propensity to comment on the story in a supportive or critical manner and intentions to share the news story in social media?

The results of study 2 indicate that a disclaimer is not as effective as other users' comments in stopping the spread of fake news. Whereas the means of the attitudes towards the post and intentions to share the post in the group who had seen critical comments (group 3) was significantly lower than in the group that had seen supportive comments (group 2), the means of the group who had seen a disclaimer (group 4) was not. Furthermore, when looking at table 4, one notices that it is not until exposed to other users' critical comments to the fake news story that respondents' own likelihood of posting critical comments exceed the expected count.

Returning briefly to the discussion following study 1, the results of study 2 supports the theoretical reasoning behind the hypotheses rather than the existence of a mere 'waking up'-effect. Conformity does indeed seem to be an important factor in steering people's responses to fake news. Social media users seem to use the comments of other people as a guide for how to respond to disinformation online rather than disclaimers. It thus validates recent decisions by social media companies to move away from flagging options in responses to fake news as they do not seem to be particularly effective.

5. Final discussion

As noted in the introduction, the present research was intended to contribute to the emerging literature on debunking disinformation and fake news. Previous studies have thoroughly investigated how counterfactuals serves to correct misperceptions caused by fake news. This research, however, takes a step back and investigates not the misperceptions of those exposed to fake news. Rather, it investigates peoples' attitudes towards, and intentions to comment on and share, the fake news in the light of other users' reactions to the disinformation. Specifically, this research examines what effect it has on individuals exposed to fake news that other users take a stand against it and identifies the disinformation as such through the comment function.

The results show that the actions of other users in the comment section of fake news articles significantly influences peoples' attitude towards disinformation, as well as their intentions to comment and share the fake news. The results also show that actions of other users online might be more effective than disclaimers and other means of countering fake news from social media companies.

5.1 Implications

The results of the present research offer implications to both theory and practice. Theoretically, it adds primarily to the research on conformity online. Previous studies have demonstrated that conformity is not confined to physical interactions but is also very much a factor online (Rosander and Eriksson, 2012). For instance, Fox and Tang (2014) have demonstrated that conformity predicts sexist behavior online and Teunissen et al. (2012) have shown how conformity online influences drinking habits. Furthermore, to underscore the powerful role of conformity on the internet, both Williams et al. (2000) and Tsikerdekis (2013) has demonstrated that individuals conform to others online irrespective of their degree of anonymity. This study adds to that stream of research. It demonstrates yet again the powerful forces of conformity online. It shows that it influences consumer responses to fake news and online disinformation, an important issue of our time. Previous research in this field have mostly examined the role and tactics of social media companies in debunking fake news. This study instead focuses on the role of other users, demonstrating that their actions are as important, if not more important, than those of social media platforms.

That's not to say that this study offers no practical implications for social media companies, however. It shows that they need to combine their ongoing work of finding effective ways of alerting users to the existence of fake news (Flynn, 2017) with initiatives to involve other users in these efforts. Encouraging other users to debunk fake news stories and providing them with incentives to do so ought to be high on their agenda. Another stakeholder who might derive practical implications from this study is the authorities. If ordinary citizens

should play a role in countering fake news they must be given the tools to do so. Initiatives to strengthen peoples' skills in source criticism as well as public information campaigns about fake news and individuals' role in countering it are both options to consider.

5.2 Limitations

This study is naturally has limitations that we encourage future researchers to address. For starters, no distinction was made between heavy users of Facebook and those who use it less frequently. It is plausible, for instance, that heavy users are better at spotting fake news articles and are influenced less by the comments of other users than novices. Investigating how heavy vs. light users of Facebook are governed by the actions of other users when reacting to fake news is a task left to future researchers.

Another limitation of this research is that it did not account for how personally relevant the fake news used in the studies were to respondents and how that influenced their reactions. For example, the two fake news used in this research might feel more relevant to conservatives than liberals and thus the responses among those two groups might differ. Likewise, the reaction to the Colin Kaepernick, who many associate with the Black Lives Matter-movement, might be different among minority respondents. Future researchers should look at this issue as well.

Lastly, future researchers are encouraged to investigate how mixed comments influence reactions to fake news. Typically, comment threads of fake news offer a mixture of positive and negative comments. This study did not take the effects of such mixed comment threads into account. Future studied could for example investigate how different proportions of positive and negative comments affect responses, as well as the order of those comments. That way, we could all gain a better understanding of how people are influenced by others when responding to fake news.

6. References

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

Asch, S. E. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 70.

Bode, B., & Vraga, E. K. (2018) See Something, Say Something: Correction of Global Health Misinformation on Social Media. *Health Communication*. 33 (9). 1131-1140.

Braun, K.A., Ellis, R., & Loftus, E.F. (2002) Make my memory: how advertising can change our memories of the past. *Psychology and Marketing*. 19. 1–23.

Breitsohl, J., Wilcox-Jones, J.P., & Harris, I. (2015) Groupthink 2.0: An empirical analysis of customers' conformity-seeking in online communities. *Journal of Customer Behaviour*. 14 (2). 87-106.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data?. *Perspectives on psychological science*. 6(1), 3-5.

Cialdini, R.B., & Goldstein, N.J. (2004) Social Influence: Compliance and Conformity. *Annual Review of Psychology*. 55. 591-621.

Cialdini, R.B. & Trost, M.R. (1998) Social influence: social norms, conformity and compliance. In Gilbert, D.T., Fiske S.T., Lindzey, G. (Eds). *The Handbook of Social Psychology*. McGraw-Hill, Boston, MA: 151-192.

Chan. M.S, Jones, C.R., Jamieson, K.H., & Alberracin, D. (2017) Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*. 28 (11), 1531-1546.

Colliander, J., & Marder, B. (2018) 'Snap happy' brands: Increasing publicity effectiveness through a snapshot aesthetic when marketing a brand on Instagram. *Computers in Human Behavior*. 78. 34-43.

Colliander, J. & Wien, A. (2013) Trash talk rebuffed: What can we learn from the phenomenon of consumers defending companies criticized in online communities? *European Journal of Marketing*. 47 (10). 1733-1757.

Dahlén, M., Rosengren, S., Törn, F., & Öhman, N. (2008) Could Placing Ads Wrong Be Right? *Journal of Advertising*. 37 (3). 57-67.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology,* 51(3), 629-636.

Dewey, C (2016) Facebook Fake-News Writer: 'I Think Donald Trump is in the White House because of Me'. Washington Post. Retrieved from:

https://www.washingtonpost.com/news/the-intersect/wp/2016/11/17/facebook-fake-news-writer-i-think-donald-trump-is-in-the-white-house-because-of-me/?utm_term=.f609b6682faa (Accessed 3 December 2018)

Edelson, M., Sharot, T., Dolan, R.J. & Dudai, Y. (2011) Following the crowd: brain substrates of long-term memory conformity. *Science*, 333(6038). 108-111.

Flynn, K. (2017) Facebook abandons an attempt to curb fake news. Here's why. *Mashable.* Retrieved from: https://mashable.com/2017/12/21/facebook-fake-news-abandon-disputed-flag-related-articles/?europe=true#SwKMS5TSaaqI (Accessed 3 December 2018).

Fox, J., & Tang, W.Y. (2014) Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior*. 33. 314-320.

Gillin, J. (2018) If you're fooled by fake news, this man probably wrote it. *Politifact*. Retrieved from: https://www.politifact.com/punditfact/article/2017/may/31/If-youre-fooled-by-fake-news-this-man-probably-wro/ (Accessed 3 December 2018).

Hamilton, R.W., Schlosser A. & Chen, Y-J. (2017) Who's Driving This Conversation? Systematic Biases in the Content of Online Consumer Discussions. Journal of Marketing Research. 54 (4). 540-555.

Huang, M., Cai, F., Tsang, A.S.L. & Zhou, N. (2011) Making your online voice loud: the critical role of WOM information. *European Journal of Marketing*. 45 (7/8). 1277-1297.

Hunt, E. (2017) 'Disputed by multiple fact-checkers': Facebook rolls out new alert to combat fake news. *Guardian.* Retrieved from: https://www.theguardian.com/technology/2017/mar/22/facebook-fact-checking-tool-fake-news (Accessed 3 December 2018).

Lazer D.M.J., Baum, M., Benkler, Y., Berinsky, A.J., Greenhill, K.M., Menczer, F., Metzger M.J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S.A., Sunstein, C.R., Thorson, E.A., Watts, D.J., Zittrain, J.L. (2018) The Science of Fake News, *Science,* 359 (6380), 1094-1096.

Nam, J., Lee, Y., Youn, N., & Kwon, K-M (2016) Nostalgia's fulfilment of agentic and communal needs: How different types of self-concepts shape consumer attitudes toward nostalgia. *Journal of Consumer Behaviour*. 15, 303-313.

Nyhan, B., & Reifler, J. (2010) When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*. 32 (2), 303-330.

Parkinson, H. J. (2016) Click and Elect: How Fake News Helped Donald Trump Win a Real Election. Guardian. Retrieved from: https://www.theguardian.com/commentisfree/2016/nov/14/fake-news-donald-trump-election-alt-right-social-media-tech-companies (Accessed 3 December 2018)

Pew Research Center (2016). *Many Americans Believe Fake News Is Sowing Confusion.* Retrieved from: http://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/ (Accessed 3 December 2018).

Read, M. (2016) Donald Trump Won because of Facebook. New York Magazine. Retrieved from: http://nymag.com/intelligencer/2016/11/donald-trump-won-because-of-facebook.html (Accessed 3 December 2018)

Rosander, M., & Eriksson, O. (2012) Conformity on the internet – The role of task difficulty and gender differences. *Computers in Human Behavior*. 28 (5): 1587-1595.

Saslow, E. (2018) 'Nothing on this page is real': How lies become truth in online America. *The Washington Post*. Retrieved from: https://www.washingtonpost.com/national/nothing-on-this-page-is-real-how-lies-become-truth-in-online-america/2018/11/17/edd44cc8-e85a-11e8-bbdb-72fdbf9d4fed_story.html?utm_term=.784a7d367485 (Accessed 3 December 2018).

Silverman, C. (2016) This Analysis Shows how Fake Election News Stories Outperformed Real News on Facebook. *Buzzfeed News*. Retrieved from: https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook (Accessed 3 December 2018).

Teunissen, H.A., Spijkerman, R., Prinstein, M.J. & Cohen, G.L. (2012) Adolescents' Conformity to Their Peers' Pro-Alcohol and Anti-Alcohol Norms: The Power of Popularity. Behavior, Treatment and Prevention. 36 (7). 1257-1267.

Tsikerdekis, M. (2013) The effects of perceived anonymity and anonymity states on conformity and groupthink in online communities: A Wikipedia study. *Journal of the American Society for Information Science & Technology*. 64 (5). 1001-1015.

Vasoughi, S., Roy, D., & Aral, S. (2018) The spread of true and false news online. *Science.* 359(6380). 1146-1151.

Watanabe, K. (2017) The spread of the Kremlin's narratives by a western news agency during the Ukraine crisis. *The Journal of International Communication*. 23(1) 138-158.

Weinreich, H., Obendorf, H., Herder, E., & Mayer, M. (2008) Not quite the average: An empirical study of Web use. ACM Transactions on the Web. 2 (1). 1-31.

Wiggins J.S. (1991) Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In William MG, Cicchetti D (Eds). *Thinking clearly about psychology.* University of Minnesota Press: Minneapolis, MN; 89–113.

Williams, K.P., Cheung, C.K.T. & Choi, W. (2000) Cyberostricism: effects of being ignored over the internet. *Journal of Personality and Social Psychology*. 79, 748-762.

Winter, S., Bruckner, C. & Krämer N.C. (2015) They Came, They Liked, They Commented: Social Influence on Facebook News Channels. *Cyberpsychology, Behavior, and Social Networking.* 18 (8). 431-436.

Zhu, H. & Huberman, B.A (2014) To Switch or Not to Switch: Understanding Social Influence in Online Choices. *American Behavioral Scientist*. 58 (10). 1329-1344.

7. Appendix 1: Stimuli used in study 1

Post with supportive comments

America's Last Line Of Defense
March 22 at 5:24am · ☀

The truth will set you free.

SPECIAL REPORT

Austin Bomber On Clinton Foundation Payroll
It's starting to make sense.
DAILYWORLDUPDATE.COM

👍 Like   💬 Comment   ↗ Share

👍 7

View 1 more comment

Pat Smith Wow!
Like · Reply · 1d

Jordan Anderson Damn.
Like · Reply · 1d

Michael Z. Schmid I knew it.
Like · Reply · 18h

Carl Young Unbelievable!
Like · Reply · 1h

Write a comment...

Post with comments pointing out that the story is fake

America's Last Line Of Defense
March 22 at 5:24am · ☀

···

The truth will set you free.

SPECIAL REPORT

Austin Bomber On Clinton Foundation Payroll
It's starting to make sense.
DAILYWORLDUPDATE.COM

👍 Like   💬 Comment   ↪ Share   ▼

👍 7

View 1 more comment

Pat Smith This isn't true!
Like · Reply · 1d

Jordan Anderson Damn.
Like · Reply · 1d

Michael Z. Schmid Fake story!
Like · Reply · 18h

Carl Young This is fake news.
Like · Reply · 1h

Write a comment...   ☺ 📷 GIF 🙂

Post with comments pointing out that the story is fake and attacking the poster

America's Last Line Of Defense
March 22 at 5:24am · ☀

The truth will set you free.

SPECIAL REPORT

Austin Bomber On Clinton Foundation Payroll
It's starting to make sense.
DAILYWORLDUPDATE.COM

👍 Like   💬 Comment   ➦ Share

👍 7

View 1 more comment

Pat Smith It's irresponsible of you to spread this untrue stuff.
Like · Reply · 1d

Jordan Anderson Damn.
Like · Reply · 1d

Michael Z. Schmid You are contributing to the polarisation of America by promoting this false story!
Like · Reply · 19h

Carl Young Shame on you for spreading this lie!
Like · Reply · 1h

Write a comment...

8. Appendix 2 Stimuli used in study 2

Post with no comments or disclaimers

Post with supportive comments

America's Last Line Of Defense
September 26 at 2:47 AM · ⚙

WORSTPOT.US

BREAKING: Nike Files For Bankruptcy After Failed Kaepernick Campaign

Most Relevant ▾

Write a comment...

Phyllis Collins Good!
Like · Reply · 4d   👍 33
↳ 10 Replies

Mike Donald Serves them right!
👍❤ 17
Like · Reply · 4d
↳ 2 Replies

Jason Lind That's great!
👍 65
Like · Reply · 5d
↳ 11 Replies

Post with comments pointing out that the story is fake

Post with supportive comments and a disclaimer

Research highlights:

1. Fake news is an increasing problem online
2. Tests whether users conform to others' Facebook comments of fake news
3. Also tests whether others' comments were more important than 'fake flags'
4. Others comments significantly affected attitudes and intentions to share
5. Social media companies should stimulate debunking by other users