UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Overcoming polarization with chatbot news? Investigating the impact of news content containing opposing views on agreement and credibility

Zarouali, B.; Makhortykh, M.; Bastian, M.; Araujo, T.

DOI
10.1177/0267323120940908

Publication date
2021

Document Version
Final published version

Published in
European Journal of Communication

Link to publication

Citation for published version (APA):
Zarouali, B., Makhortykh, M., Bastian, M., & Araujo, T. (2021). Overcoming polarization with chatbot news? Investigating the impact of news content containing opposing views on agreement and credibility. *European Journal of Communication*, *36*(1), 53-68. https://doi.org/10.1177/0267323120940908

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

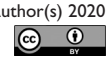
Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:06 Oct 2026

Check for updates

Full Length Article

European Journal of Communication
2021, Vol. 36(1) 53–68
© The Author(s) 2020

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0267323120940908
journals.sagepub.com/home/ejc

Overcoming polarization with chatbot news? Investigating the impact of news content containing opposing views on agreement and credibility

Brahim Zarouali
University of Amsterdam, The Netherlands

Mykola Makhortykh
University of Bern, Switzerland

Mariella Bastian and Theo Araujo
University of Amsterdam, The Netherlands

Abstract

Chatbots are a burgeoning opportunity for news media outlets to disseminate their content in a conversational way, and create an engaging experience around it. Since chatbots are social and interactive technologies, they might be effective tools to lower the threshold of engaging with news content containing opposing views. In an experiment, we test this idea by investigating whether people are more likely to accept a news article containing conflicting views when it is delivered by a chatbot, as compared with the same article on a news website. The results indicated that people agreed more to a counter-attitudinal news article when it was delivered by a news chatbot (compared with the website article). In addition, users also perceived this chatbot article as more credible. The underlying process for this effect was that people attributed human-like characteristics to the chatbot on an implicit level (i.e., perceived mindless anthropomorphism). These results are discussed in the light of their potential contribution to an informed public discourse and a decrease in polarization in our society.

Corresponding author:
Brahim Zarouali, Amsterdam School of Communication Research, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands.
Email: b.zarouali@uva.nl

Keywords
Anthropomorphism, chatbots, credibility, news consumption, polarization

Introduction

In recent years, computational and automatization processes have been installed in many areas of online mass communication, including news media outlets. As such, chatbots have emerged within the domain of journalism (Veglis and Maniou, 2019). Chatbots are conversational agents that are programmed to communicate with people through natural language, and when requested, can automatically provide news content and updates to the user (Zarouali et al., 2018). Such chatbots are typically integrated in private messaging applications, such as Facebook Messenger. In the news industry, chatbots are mainly being used as a conversational channel for disseminating information, where users are being served with conversation-sized chunks of news content (Diakopoulos, 2019). By distributing news through a channel that is characterized by its conversational, one-to-one, social and interactive nature, news organizations aim to improve audience reach and engagement (Jones and Jones, 2019). For instance, BBC actively tries to leverage bot technology using chatbots on various messaging platforms (e.g., Facebook, Telegram, etc.) to increase audience reach and automate their news coverage on social media (BBC News Labs, 2019).

Recently, it has been argued that news chatbots could be promising tools to lower the threshold for people to engage with *opposing* or *conflicting* perspectives in news content, and encourage them to rethink their initial attitudinal positions (the notion of 'biased bots' by Dingler et al., 2018). Indeed, research has shown that people might put considerable efforts in establishing common ground with human-like conversational agents (Corti and Gillespie, 2016). This leads to an important question: are people more likely to accept *opposing news content* when it is provided by chatbots because of their social and human-like nature?

As of yet, this question has not been subjected to empirical testing. This research inquiry can be of great societal relevance, since studies have shown that exposure to attitude-incongruent (or opposing) news content can lead to a decrease in partisan divides and polarization as citizens break out of their information bubbles (e.g., Hart et al., 2015; Mutz, 2002), as well as create a more informed public discourse (Dingler et al., 2018). In this study, we aim to investigate the effects of engaging with news chatbots when they are being programmed to *oppose* someone's political views. In particular, we investigate the extent to which online users are likely to agree with an opposing news story on a highly divisive topic (i.e., migration) when it is provided by a conversational chatbot as compared with the same news story on a news website. In addition, we also focus on the credibility effects that such news stories have on online users: do users evaluate a chatbot-delivered counter-attitudinal news article as more credible than when it is provided on a news website? Finally, we also test whether perceived anthropomorphism, or the attribution of human-like qualities to chatbots, serves as the underlying mechanism in these relationships.

Theoretical framework

Computers Are Social Actors and mindless anthropomorphism

Studies following the Computers Are Social Actors (CASA) paradigm (Nass and Steuer, 1993; Reeves and Nass, 1996) have largely established that humans tend to react socially to computers and to technology. This effect has been shown for media in general (including television, e.g., Reeves and Nass, 1996), but especially for computers (Moon and Nass, 1998; Nass et al., 1995), websites (Kim and Sundar, 2012; Sah and Peng, 2015) and, more recently, conversational agents, such as chatbots (Araujo, 2018; Ho et al., 2018).

These social reactions take place whenever an object has 'enough cues to lead the person to categorize it as worthy of social responses' (Nass and Moon, 2000: 83). One of such social reactions is perceptions of anthropomorphism, that is, a general tendency to assign 'the imagined or real behavior of nonhuman agents with humanlike characteristics, motivations, intentions, and emotions' (Epley et al., 2007: 864). Perceptions of anthropomorphism can be triggered, for example, by the presence of anthropomorphic cues in the interface, such as the design and communication style of a website (Kim and Sundar, 2012; Sah and Peng, 2015) or, for chatbots in particular, the usage of informal language (Araujo, 2018) as well as visual (e.g., avatar), identity (e.g., name) and interactivity (e.g., contingency in responses) cues (Go and Sundar, 2019). Because of this cumulation of cues, and the simple fact that the communication with a chatbot takes place in a manner usually associated with interpersonal communication – that is in a 'dialogic fashion, using natural language' (Dale, 2016: 811) – news chatbots, we argue, trigger stronger anthropomorphic perceptions compared with online news websites.

It is important to note that anthropomorphism can be studied both as a *mindful* and as a *mindless* process (Araujo, 2018; Kim and Sundar, 2012). Studies associated with the CASA framework suggest that even though users may have social reactions to technology – that is, exhibiting direct behaviours towards an entity in a way that suggests they are interacting with another human – they often refuse to admit these social responses (Nass and Moon, 2000), therefore, denying a *mindful* attribution of anthropomorphism towards the technology. Research explicitly comparing mindful and mindless attributions of anthropomorphism in relation to chatbots and websites is scarce, yet one of the few examples to have done so (Kim and Sundar, 2012) showed that participants provide lower *mindful anthropomorphism* evaluations to a virtual agent compared with a website, while providing higher (marginally significant) *mindless anthropomorphism* evaluations to the virtual agent. So, when it comes to conversational agents, people may not tend to mindfully anthropomorphize chatbots, but they can do so mindlessly. Aligned with these findings, we propose the following hypotheses:

$H1_a$. A chatbot providing a news story will generate stronger *mindless* anthropomorphism perceptions among users, as compared with a news website.

$H1_b$. A chatbot providing a news story will generate weaker *mindful* anthropomorphism perceptions among users, as compared with a news website.

Agreement with attitude-incongruent news

Partisan online news consumption is often referred to as one of the possible drivers of political polarization (Flaxman et al., 2016; Gentzkow and Shapiro, 2011; Stroud, 2010). The rise of digital news platforms enables more freedom of choice for news readers in terms of which news to consume. At the same time, it also facilitates selective exposure to news as readers tend to consume content that fits their pre-existing attitudes (Messing and Westwood, 2014; Yeo et al., 2015). This tendency together with the growing use of personalized news distribution systems can lead to the formation of isolated online communities ('filter bubbles' (Pariser, 2011) or 'echo chambers' (Sunstein, 2017), where readers primarily interact with attitude-congruent content and have limited exposure to attitude-incongruent stories.

Selective exposure to attitude-congruent political information is related to multiple factors, varying from the mental discomfort which can arise from exposure to attitude-challenging views to the lower perceived credibility of attitude-incongruent sources (Metzger et al., 2015). Messing and Westwood (2014) suggest that the growing reliance on social endorsement (i.e., the suggestion of specific content made by other users, usually with similar political attitudes) also contributes to users' tendency to selectively expose themselves to certain types of content. Together, these factors can result in the ideological segregation between users with different political attitudes and increase societal polarization. The concerns about polarization effects of the selective exposure lead to the growing interest towards possible mechanisms of countering partisan news consumption. The easiest way to counter-political polarization arising from selective exposure to attitude-congruent information is to expose the readers to the attitude-incongruent information. However, simply showing attitude-challenging news stories to the readers can lead to even higher polarization as the readers would feel confronted with opposing points of view (Babaei et al., 2018; Bail et al., 2018).

As it was noted earlier, a number of studies coming from the CASA paradigm indicate that humans tend to apply social scripts when interacting with computers, in particular when these technologies express cues associated with humans, such as voice or personality (Nass et al., 1995; Nass and Moon, 2000). The linguistic cues, in particular the ability to maintain a dialogue with the user, is an important factor that leads to anthropomorphization of the conversational agents, such as chatbots. By attributing anthropomorphic qualities to the chatbot, users feel like they are interacting with humans, which has implications for how they consume the information provided by the chatbot (e.g. Araujo, 2018; Chaves and Gerosa, 2019).

Zanbaka et al. (2006) found that anthropomorphic cues influence the degree of persuasiveness of the information communicated via the conversational agent and demonstrate that this information is perceived as if it was coming from a human agent. On their turn, Corti and Gillespie (2016) found that when people speak to a conversational agent with human-like features, they were more eager to establish a common ground. In a series of conformity studies, Hertz and Wiese (2016, 2018) found that the extent to which participants agreed with the opinion of a computer agent was comparable with the agreement levels towards human agent opinions. Based on this, we expect that a counter-attitudinal or opposing news story delivered by a chatbot will elicit higher agreement levels as compared with the same news on a website (Dingler et al., 2018). In addition, we

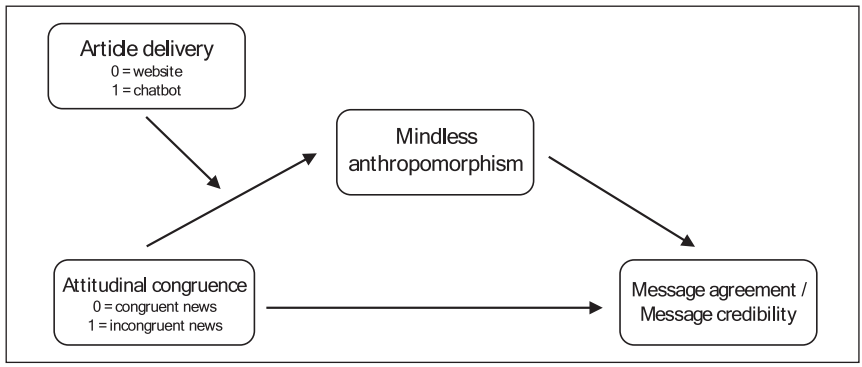


Figure 1. Conceptual moderated mediation model for both dependent variables.

expect mindless anthropomorphism to mediate this relationship (Waytz et al., 2014), meaning that an opposing news story delivered by a chatbot (as compared with a website) will result in a higher mindless anthropomorphism, which in turn should positively influence the agreement level towards the story (see Figure 1 for a conceptual model). In sum, we formulate the following hypotheses:

H2$_a$. Users will agree more with a counter-attitudinal news story provided by a chatbot as compared with the same story on a news website.

H2$_b$. Users will agree more with a counter-attitudinal news story provided by a chatbot as a result of an increase of in mindless anthropomorphism.

Credibility in partisan political news

Credibility plays a crucial role for the audiences' perception of news in various regards: on the level of the medium, the source and the message (Metzger et al., 2003). This study seeks to explore how credible people perceive the specific content – the message – delivered by chatbots. Following Appelman and Sundar (2016), we understand this form of credibility – message credibility – as 'an individual's judgement of the veracity of the content of communication' (p. 63). Multiple factors can influence the perception of message credibility, including the degree to which information communicated through a specific medium and in a specific story is consistent with users' political perspectives (Flanagin and Metzger, 2014). Kahan et al. (2010) showed in their study about attitudes towards HPV vaccines that respondents attach higher credibility to information if it is attitude consistent. Similarly, Metzger et al. (2015) found that users view attitude-consistent news sources and stories as more credible than attitude-incongruent ones.

While the relationship between attitude-congruent news content and credibility is well-recognized, the effects of technology (in particular, of new technologies of news delivery) through which news are consumed for message credibility are less common subjects of scholarly inquiry. Still, a couple of recent studies offer some relevant empirical evidence in this area. For instance, Graefe et al. (2018) found that subjects

evaluated computer-generated news articles as more credible and higher in journalistic expertise than human-generated articles. In a similar vein, Wölker and Powell (2018) found that news readers tend to evaluate stories produced by automated agents as similar in terms of content credibility to the ones written by journalists.

In the light of these results, we can argue that that users are likely to perceive the conversational agent-based medium for incongruent news consumption as more credible than a website-based one. Thus, we expect that a chatbot delivering an attitude-inconsistent news story will generate higher levels of message credibility as compared with a traditional news website providing the same story. As we already discussed in earlier sections, humans often perceive conversational agents as social entities as shown by the research based on the CASA paradigm. The mindless application of social rules to interactions with conversational agents affects the perceived quality of these interactions, including the information credibility (Edwards et al., 2014; Kim and Sundar, 2012). Therefore, we again expect mindless anthropomorphism to fulfil a mediating role, in that, users will evaluate the chatbot article containing opposing views (compared with the website article) as more mindlessly anthropomorphistic, which in turn will have a positive influence on the credibility of the news article (see Figure 1 for a conceptual model). We formulate the following hypotheses:

H3$_a$. Users will perceive a counter-attitudinal news story provided by a chatbot as more credible compared with the same story a news website.

H3$_b$. Users will perceive a counter-attitudinal news story provided by a chatbot as more credible as a result of an increase of in mindless anthropomorphism.

Methods

Design and participants

We carried out an experiment with a 2×2 between-subjects design. The first factor includes the manipulation of the news article frame: *congruent* or *incongruent* with a participant's own attitude (where incongruent means that it contains opposing views); the second factor includes news dissemination mode: a *chatbot* or a *website*. Based on a meta-analysis on attribute framing effects from the work of Freling et al. (2014), we conducted an a priori power analysis with a significance level of $\alpha = .05$, a medium effect size of $d = .50$ ($f = .25$) and a desired power of $(1–\beta) = .90$. This revealed an estimated sample size of 171 participants. We collected data from an online research panel of 190 young adults aged 18–35 years ($M = 27.82$, $SD = 5.20$; 55% female) in the Netherlands. This age range was chosen because of sample relevance: younger adults are avid users of social media and instant messaging applications, and therefore, most likely to receive chatbot news updates.

Stimulus materials

News story. Two short news articles were carefully developed in the layout of NOS, the public broadcaster of the Netherlands. To manipulate the news framing, we chose a

highly polarized and divisive news topic: migration. The first frame had a pro-immigration stance (positive valence), and the second one contained anti-immigration views (negative valence). In a recent empirical review, Eberl et al. (2018) revealed that three issue frames dominate the news coverage about immigration: the economic, cultural and security frame. In the present study, we chose to focus on the economic consequences of immigration. To keep the two frames constant in terms of their factual content, we adopted an equivalence frame (Druckman, 2001; Tversky and Kahneman, 1981). In this scenario, two logically equivalent alternatives are portrayed in different ways (e.g. 5% unemployment or 95% employment, 97% fat-free or 3% fat). Thus, the information being presented referred to the same facts, but the 'frame' in which it is presented varied. In our case, we used the same fact for both news stories (19% unemployment rate among immigrants), but its interpretation (i.e., whether this rate is high or low, and what are the consequences for the welfare state and social security) was exactly the opposite (see Online Appendix for the news article framings).

To determine whether this manipulation was successful, we conducted a pretest among 26 students (M_{age}=25.65, SD_{age}=2.35). We presented them both frames of the article. Participants were then asked how they evaluate the views in the news article, ranging from 1 (*anti-immigration*) to 10 (*pro-immigration*). They rated the anti-immigration article with a mean score of 1.96 (SD=.92), whereas the pro-immigration story scored significantly higher with 7.96 (SD=1.37) $t(25)=-17.10$, $p<.001$). The manipulation was successful.

Chatbot. The chatbot was developed using the Conversational Agent Research Toolkit (CART; Araujo, 2020), which offers researchers computational tools to design conversational agents for experimental research. CART uses the DialogFlow API as a dialogue management tool, the Microsoft Bot Framework to make the agent available in a web chat, and a connection to a MySQL database to store all the conversation logs (see Online Appendix for a visual representation of the chatbot – Figure 1). The toolkit also allows to fully integrate the chatbot in the online survey platform, in this case, Qualtrics.

After successfully completing the chatbot set-up, we configured the customized dialogue. First, the participant had to say 'Hi' to the chatbot, which in turn greeted the participant ('Good day! I am the NOSbot, the chatbot from NOS. How are you today?'), asked the participant whether they wanted to receive a news update, and finally provided the news article (pro- or anti-immigration article, at random) directly in the chat interface (see Online Appendix for an example of a conversation – Figure 2). As a control check, we included a conversation code at the end of the news article, which the participant needed to enter in the survey tool to continue with the questionnaire.

Procedure

At the start of the study, respondents were presented with an informed consent form. In case of agreement, a screening question was presented to assess their attitude towards immigrants. This allows us to determine whether they are predominantly pro- or anti-immigration (see measures – attitudinal congruence). After this screening question, they started with the first part of the survey, which consisted of demographic questions.

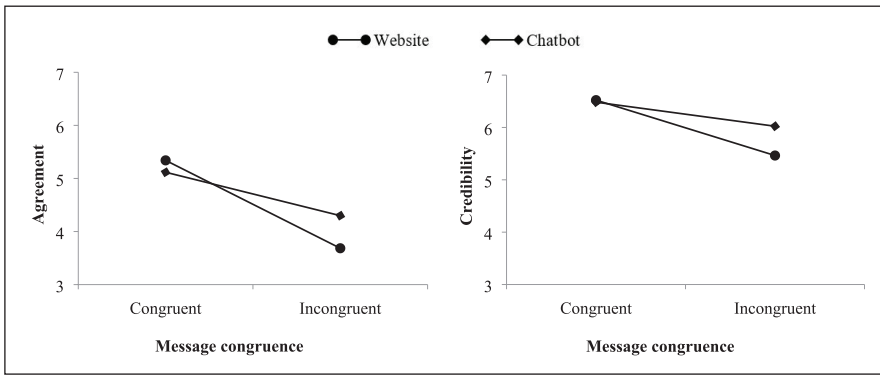


Figure 2. Visualization of the moderation effects in $H2_a$ and $H3_a$.

Then, participants were randomly exposed either to a pro- or anti-immigration news story (and depending on their own attitude on migration, this message was then congruent or incongruent), randomly delivered by either an online news website or the chatbot. A control check was integrated in every condition to ascertain that respondents read the presented news item (i.e. a confirmation code at the end of the chatbot conversation, or a forced minimal reading time of 15 seconds in case of the online article). After the experimental exposure, participants had to answer questions related to the dependent and mediation variable(s). Finally, all participants were thoroughly debriefed. This study protocol was approved by the ethical review board of our institution filed under number 2019-PCJ-10491.

Independent variables

Attitudinal congruence. By presenting the respondents with a pro- or anti-immigration framing, we could assess whether the particular frame was congruent or incongruent with their own attitudes on the issue. To do so, we followed the same procedure as Hameleers and Van Der Meer (2019). At the start of the survey, we asked participants to indicate how strongly they support or oppose that immigrants enter the Netherlands, on a scale from 1 (*strongly oppose*) to 7 (*strongly support*). Participants with a score of 1 through 3 were categorized as the ones with anti-immigration views, and those with a score of 5 through 7 as the ones with pro-immigration views. Participants with the score 4 were excluded from further participation. Based on this screening question, we could classify the article as either *congruent* or *incongruent* to one's own attitudinal stance towards immigration.

News article dissemination method. The news item was delivered in the format of an online NOS news article (we mimicked the exact same layout of a real NOS new article), or as an item provided by the NOS chatbot (i.e., the update was embedded in the messaging interface of the chatbot). Both conditions were identical in every respect, and only differed in the way they deliver a news item to the intended audience.

Dependent variables

Message credibility. To measure message credibility, we adopted a set of items developed by Appelman and Sundar (2016). These scholars suggested that credibility, in the context of news content, can be measured by asking participants to rate how well the following three adjectives describe the news item: *accurate, authentic* and *believable*, with answer options ranging from 1 *(describes very poorly)* to 7 (*describes very well; M* = 4.13; *SD* = 1.39; Cronbach's $\alpha = .92$)

Message agreement. To assess agreement with the news story, we followed the same protocol as Hameleers and Van Der Meer (2019). Participants were offered three concrete issue positions that were forwarded in the article they were exposed to. Thus, people in the pro-condition received a different set of three items than people from the anti-condition. After reading these positions, they were asked to which extent they agree with these items, ranging from 1 (*totally disagree*) to 7 (*totally agree; M* = 4.54; *SD* = 1.41; Cronbach's $\alpha = .88$).

Anthropomorphism. Mindful anthropomorphism was measured using three items from Kim and Sundar (2012). These items asked participants to evaluate the chatbot when it comes to being *human-like/machine-like, natural/unnatural* and *lifelike/artificial*, along a 7-point semantic differential scale. These items were then averaged (*M* = 4.17, *SD* = 1.47, Cronbach's $\alpha = .93$).

Mindless anthropomorphism was used both as a dependent ($H1_a$) and mediator variable in this study ($H2_b$ and $H3_b$). Following the operationalization of Kim and Sundar (2012), we measured mindless anthropomorphism by asking participants to evaluate the delivery platform (chatbot/website) based on four adjectives: '*likeable*', '*sociable*', '*friendly*' and '*personal*'. The response categories ranged from 1 (*describes very poorly*) to 7 (*describes very well; M* = 4.42; *SD* = 1.33; Cronbach's $\alpha = .91$).

Results

Manipulation and randomization check

The manipulation of the news item's attitudinal stance was successful ($t(188) = 5.03$, $p < .001$). The 97 participants exposed to the pro-immigration story evaluated the views in the story as significantly more in favour of supporting immigration (*M* = 6.76, *SD* = 2.54) than the 93 participants exposed to the anti-immigration news item (*M* = 4.84, *SD* = 2.73). In addition, we conducted a between-condition randomization check at the outset of the analyses. This showed that the conditions did not differ with respect to age ($F(3, 186) = .011$, $p = .99$), gender ($\chi^2(3) = 5.19$, $p = .16$), education ($\chi^2(15) = 6.74$, $p = .96$) and professional status ($\chi^2(15) = 8.87$, $p = .88$).

Main effects

In a first series of analyses, we tested the main effects hypothesized in $H1_a$ and $H1_b$. It was found that the chatbot had significantly lower scores for mindful anthropomorphism

(M=3.81, SD=1.47) compared with a website (M=4.49, SD=1.40; F(1, 188)=10.72, $p<.001$, Cohen's d=.48). When testing mindless anthropomorphism, an opposite pattern was found: the chatbot revealed higher scores (M=4.69, SD=1.06) for mindless anthropomorphism compared with the website (M=4.19.81, SD=1.50; F(1, 188)=6.89, $p<.001$, Cohen's d=.38). These results confirm $H1_a$ and $H1_b$.

Moderation effects

To test the interaction in $H2_a$, we conducted an analysis of variance with message delivery and message congruence as factors, and message agreement as a dependent variable. The two-way interaction was found to be significant (F(1, 186)=5.26, $p<.05$, Cohen's d=.34; Figure 2). We then investigated the simple main effects, and this revealed that online users agreed significantly more with a counter-attitudinal news story when it was provided by a chatbot (M=3.69), as compared with the same story displayed on a news website (M=4.30; F(1, 186)=6.24, $p<.01$, Cohen's d=.36). Thus, $H2_a$ was supported.

A similar result was found in testing $H3_a$, where an analysis of variance revealed a significant interaction effect between message delivery and congruence on message credibility (F(1, 186)=4,32, $p<.05$, Cohen's d=.31; Figure 2). More precisely, an analysis of the simple main effects showed that users perceived the incongruent news story provided by a chatbot (M=6.24) as more credible as compared with the same story provided by an online news website (M=5.46; F(1, 186)=8.66, $p<.01$, Cohen's d=.43). $H3_a$ was, therefore, confirmed.

Moderated mediation analyses

To test the mediating role of mindless anthropomorphism in these interaction effects, we used PROCESS to estimate conditional indirect effects (Model 7 – 10,000 bootstrap intervals – BC 95% confidence intervals; Hayes, 2013). With agreement as a dependent variable, results yielded a significant index of a moderated mediation model (b=.20, SE=.13; BC 95% CI (.017–.523)). More precisely, users that received an incongruent news item through the chatbot (and not those in the website condition) had an increase in perceived mindless anthropomorphism, which then positively influenced their agreement with that news item (b=.22, SE=.10; BC 95% CI (.060–.469)).

The same analyses were repeated for credibility as a dependent variable. The confidence interval of overall moderated mediation model did not include zero, indicating a significant model (b=.35, SE=.18; BC 95% CI (.047–.769)). The conditional indirect effect showed that *only* users that were provided an incongruent news item by a chatbot generated an increase in mindless anthropomorphism, which in turn lead to higher levels of message credibility (b=.38, SE=.13; BC 95% CI (.161–.676)). Altogether, these results support $H2_b$ and $H3_b$.

General discussion

This study investigated the extent to which people are more likely to accept opposing news content when it is provided by a chatbot when compared with a news website, as well as the extent to which the mode of delivery (chatbot vs news article) influences

credibility perceptions. In this respect, we found that participants tend to agree more to a news article containing opposing views on migration when it was delivered by a news chatbot, compared with an online website. At the same time, users also perceived these counter-attitudinal stories as more credible when provided by a chatbot. As an explaining mechanism, we found an influence of attributions of human-like characteristics on an implicit level (perceived mindless anthropomorphism) to the chatbot, which might increase agreement and credibility of a news story containing conflicting views. These results suggest that news chatbots can be interesting technology tools that help users to see both sides of an argument, to consider middle ground for discussion, and to reflect on their views. Hence, chatbots may contribute to a more informed public discourse and depolarization in our fragmented societies (Dingler et al., 2018).

From a theoretical perspective, our study contributes to the body of research using the CASA paradigm, and to how chatbots are perceived as human-like in a *mindful* or *mindless* way (Araujo, 2018; Kim and Sundar, 2012). In our results, we found that when people had to evaluate the chatbot in a mindful way (by directly asking them to evaluate the chatbot based on items such as 'human-like vs machine-like'), perceived anthropomorphism was lower compared with a website. However, when participants had to evaluate the chatbot in a mindless way (by indirectly asking them to rate the chatbot based on personality characteristics items, such as 'sociable', 'likeable' and 'friendly'), we found that they reported higher levels of perceived anthropomorphism compared with the website. This might indicate that on a conscious and thoughtful level, people somewhat deny that they treat chatbots as human-like (when thinking about it intentionally, they refuse to admit that they react socially to chatbots), whereas on a more implicit, less thoughtful level, they do attribute social and human-like features to chatbots to a much greater extent. Therefore, this supports the idea that people may not mindfully anthropomorphize technology, but they can do so mindlessly (Kim and Sundar, 2012).

Furthermore, this study takes work on CASA a step further by revealing that this mindless attribution of human-like characteristics to a chatbot even leads to agreeing more with counter-attitudinal information that is being presented by the chatbot. Also, the chatbot information is perceived as more credible than the website information. To our knowledge, this is the first study to show that chatbots, by virtue of their human-like and social qualities, are persuading people to a greater extent in agreeing to messages that are opposed to their own views and beliefs, and at the same time, make these messages more credible.

This study also contributes to the literature on selective exposure and political polarization. In our current media environment that is characterized by choice overload and an increased fragmentation (Andrejevic, 2013; Van Aelst et al., 2017), political polarization may be augmented by selective news exposure and avoidance of political information, resulting in confirmation biases (Hameleers and Van Der Meer, 2019; Stroud, 2010). A potential way to overcome this is to expose people to attitude-incongruent news content, which can lead to a decrease in partisan divides and polarization as citizens break out of their isolated information bubbles. In this respect, our findings that opposing news articles delivered through chatbots may lead to higher agreement and credibility levels may provide an interesting contribution. We argue that conversational agents may be considered as potential tools for decreasing societal polarization by offering people disconfirming news content, and potentially increasing their receptiveness towards attitude-incongruent

news. Instead of automatically dismissing news content based on the simple fact that it opposes one's own beliefs, chatbots might succeed – by virtue of their perceived social features – in increasing a person's willingness to – at least – consider arguments for both sides of the issue, and reflect more upon disconfirming information. This highlights a potential to decrease selective exposure and make citizens better informed about different social and political phenomena.

This study has practical implications for news outlets. The concept of 'media diversity' is anchored in our ideas of a well-functioning democratic society. Having a plurality of voices and ideas in media content is essential for citizens to encounter different opinions and beliefs, self-reflect on their own viewpoints, and make informed choices about a variety of issues (Helberger, 2018; Valcke et al., 2015). In line with the present results, news chatbots could offer a persuasive and engaging channel for exercising the journalistic goal of offering citizens a wide range of opinions and perspectives, including counter-attitudinal ones. A relevant example of how this can be implemented in practice is the 'Angry Uncle Bot', a chatbot project from Smart Politics for *The New York Times*, launched in November 2019 (Tamerius, 2019). This chatbot allows people to hold a conversation with their (virtual) angry uncle that holds opposite political views. Based on this simulated chat, people are forced out of their comfort zone to try and understand why others hold different beliefs. The objective of the chatbot is to move people towards a state of mutual understanding with their politically opposite interlocutor. This example illustrates how media organizations can use chatbots as 'critical media technologies' that intend to invite citizens to reflect on their own views and acquire new ones, which could then contribute to a more informed public discourse (Dingler et al., 2018).

Finally, we discuss the limitations of the present study that could be addressed in future work. First, we used a chatbot that was simply programmed to deliver a news update to the end user. However, news media organizations are experimenting with a new generation of chatbots that move beyond this automated content delivery (Ford and Hutchinson, 2019). For instance, chatbots can personalize news content to each individual interlocutor, or allow users to ask real-time questions about complex articles and subsequently offer them additional information or clarifications (Diakopoulos, 2019; Veglis and Maniou, 2019). Although the present study is an important first step, we highly encourage future research to investigate the interplay between online users and these more 'intelligent' news chatbots.

Second, our study included a highly polarized political issue, that is, migration. Although we argue that the effects found in this study are particularly important in cases when chatbots are used for delivering highly disputed content that raises tensions in the respective societies, we also think that it might be interesting (for the sake of empirical robustness) to test news chatbot effects in relation to other divisive topics (e.g., climate), or alternatively, topics that are not divisive at all.

Third, as a single-shot experiment, this study included only one interaction with the chatbot. However, it is very likely that people's interactions with chatbots will change over time, after repeated engagements (i.e., because of relationship dynamics; Guzman and Lewis, 2019). Nowadays, it is possible to conduct longitudinal studies with chatbots to account for these possible changes over time (see CART;

Araujo, 2020). Therefore, it would be interesting to see whether the current results will strengthen, weaken or hold when participants are being delivered chatbot news updates more than once.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was funded by the Research Priority Area 'Personalised Communication' from the University of Amsterdam.

ORCID iDs

Brahim Zarouali https://orcid.org/0000-0003-1042-4635

Theo Araujo https://orcid.org/0000-0002-4633-9339

Supplemental material

Supplemental material for this article is available online.

References

Andrejevic M (2013) *Infoglut: How Too Much Information Is Changing the Way We Think and Know*. New York: Routledge.

Appelman A and Sundar SS (2016) Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly* 93(1): 59–79.

Araujo T (2018) Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior* 85: 183–189.

Araujo T (2020) Conversational agent research toolkit: An alternative for creating and managing chatbots for experimental research. *Computational Communication Research* 2. Available at: https://computationalcommunication.org/ccr/article/view/9.

Babaei M, Kulshrestha J, Chakraborty A, et al. (2018) Purple feed: Identifying high consensus news posts on social media. In: *Proceedings of the 2018 AAAI/ACM conference on AI, ethics, and society*, New Orleans, LA, 27 December, pp. 10–16. New York: ACM.

Bail CA, Argyle LP, Brown TW, et al. (2018) Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences of the United States of America* 115(37): 9216–9221.

BBC News Labs (2019) Bots. Available at: https://bbcnewslabs.co.uk/projects/bots/.

Chaves AP, and Gerosa MA (2019) How should my chatbot interact? *A survey on human-chatbot interaction design*. ArXiv:1904.02743 [Cs]. http://arxiv.org/abs/1904.02743

Corti K and Gillespie A (2016) Co-constructing intersubjectivity with artificial conversational agents: People are more likely to initiate repairs of misunderstandings with agents represented as human. *Computers in Human Behavior* 58: 431–442.

Dale R (2016) The return of the chatbots. *Natural Language Engineering* 22(5): 811–817.

Diakopoulos N (2019) *Automating the News: How Algorithms Are Rewriting the Media*. Cambridge, MA: Harvard University Press.

Dingler T, Choudhury A and Kostakos V (2018) Biased bots: Conversational agents to overcome polarization. In: *Proceedings of the 2018 ACM international joint conference and 2018 international symposium on pervasive and ubiquitous computing and wearable computers*, Singapore, 8 October, pp. 1664–1668. New York: ACM.

Druckman JN (2001) The implications of framing effects for citizen competence. *Political Behavior* 23(3): 225–256.

Eberl J.-M, Meltzer CE, Heidenreich T, Herrero B, Theorin N, Lind F, Berganza R, Boomgaarden HG, Schemer C, and Strömbäck J (2018) The European media discourse on immigration and its effects: A literature review. *Annals of the International Communication Association* 42(3): 207–223. https://doi.org/10.1080/23808985.2018.1497452

Edwards C, Edwards A, Spence PR, et al. (2014) Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior* 33: 372–376.

Epley N, Waytz A and Cacioppo JT (2007) On seeing human: A three-factor theory of anthropomorphism. *Psychological Review* 114(4): 864–886.

Flanagin A and Metzger MJ (2014) Digital media and perceptions of source credibility in political communication. In: Kenski K and Hall-Jamieson K (eds) *The Oxford Handbook of Political Communication*. Oxford: Oxford University Press, pp. 417–436.

Flaxman S, Goel S and Rao JM (2016) Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly* 80(S1): 298–320.

Ford H and Hutchinson J (2019) Newsbots that mediate journalist and audience relationships. *Digital Journalism* 7: 1013–1031.

Freling TH, Vincent LH and Henard DH (2014) When not to accentuate the positive: Re-examining valence effects in attribute framing. *Organizational Behavior and Human Decision Processes* 124(2): 95–109.

Gentzkow M and Shapiro JM (2011) Ideological segregation online and offline. *The Quarterly Journal of Economics* 126(4): 1799–1839.

Go E and Sundar SS (2019) Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior* 97: 304–316.

Graefe A, Haim M, Haarmann B, et al. (2018) Readers' perception of computer-generated news: Credibility, expertise, and readability. *Journalism* 19(5): 595–610.

Guzman AL and Lewis SC (2019) Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society* 22: 70–86.

Hameleers M and Van Der Meer TGLA (2019) Misinformation and polarization in a high-choice media environment: How effective are political fact-checkers? *Communication Research* 47: 227–250.

Hart PS, Feldman L, Leiserowitz A, et al. (2015) Extending the impacts of hostile media perceptions: Influences on discussion and opinion polarization in the context of climate change. *Science Communication* 37(4): 506–532.

Hayes AF (2013) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford Press.

Helberger N (2018) Challenging diversity – Social media platforms and a new conception of media diversity. In: Moore M and Tambini D (eds) *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*. New York: Oxford University Press, pp. 153–175.

Hertz N and Wiese E (2016) Influence of agent type and task ambiguity on conformity in social decision making. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60(1): 313–317.

Hertz N and Wiese E (2018) Under pressure: Examining social conformity with computer and robot groups. *Human Factors* 60(8): 1207–1218.

Ho A, Hancock J and Miner AS (2018) Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication* 68(4): 712–733.

Jones B and Jones R (2019) Public service chatbots: Automating conversation with BBC News. *Digital Journalism* 7: 1032–1053.

Kahan DM, Braman D, Cohen GL, et al. (2010) Who fears the HPV vaccine, who doesn't, and why? An experimental study of the mechanisms of cultural cognition. *Law and Human Behavior* 34(6): 501–516.

Kim Y and Sundar SS (2012) Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior* 28(1): 241–250.

Messing S and Westwood SJ (2014) Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research* 41(8): 1042–1063.

Metzger MJ, Flanagin AJ, Eyal K, et al. (2003) Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. In: Kalbfleisch PJ (ed.) *Communication Yearbook 27*. Mahwah, NJ: Lawrence Erlbaum, pp. 293–335.

Metzger MJ, Hartsell EH and Flanagin AJ (2015) Cognitive dissonance or credibility? A comparison of two theoretical explanations for selective exposure to partisan news. *Communication Research* 47: 3–28.

Moon Y and Nass C (1998) Are computers scapegoats? Attributions of responsibility in human-computer interaction. *International Journal of Human-Computer Studies* 49(1): 79–94.

Mutz DC (2002) Cross-cutting social networks: Testing democratic theory in practice. *American Political Science Review* 96(1): 111–126.

Nass C and Moon Y (2000) Machines and mindlessness: Social responses to computers. *Journal of Social Issues* 56(1): 81–103.

Nass C and Steuer J (1993) Voices, boxes, and sources of messages: Computers and social actors. *Human Communication Research* 19(4): 504–527.

Nass C, Moon Y, Fogg BJ, et al. (1995) Can computer personalities be human personalities? In: *Proceedings of the conference companion on human factors in computing systems – CHI '95*, pp. 228–229. Available at: https://dl.acm.org/doi/10.1145/223355.223538.

Pariser E (2011) *The Filter Bubble: What the Internet Is Hiding from You*. London: Penguin Books.

Reeves B and Nass C (1996) *How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge: CSLI Publications; Cambridge University Press.

Sah YJ and Peng W (2015) Effects of visual and linguistic anthropomorphic cues on social perception, self-awareness, and information disclosure in a health website. *Computers in Human Behavior* 45: 392–401.

Stroud NJ (2010) Polarization and partisan selective exposure. *Journal of Communication* 60(3): 556–576.

Sunstein C (2017) *#Republic: Divided Democracy in the Age of Social Media*. Princeton, NJ: Princeton University Press.

Tamerius K (2019) Opinion: How to have a conversation with your angry uncle over the holidays. *The New York Times*, 18 November. Available at: https://www.nytimes.com/interactive/2018/11/18/opinion/thanksgiving-family-argue-chat-bot.html.

Tversky A and Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211(4481): 453–458.

Valcke P, Sükösd M and Picard RG (eds) (2015) *Media Pluralism and Diversity: Concepts, Risks and Global Trends*. New York: Palgrave Macmillan.

Van Aelst P, Strömbäck J, Aalberg T, et al. (2017) Political communication in a high-choice media environment: A challenge for democracy? *Annals of the International Communication Association* 41(1): 3–27.

Veglis A and Maniou TA (2019) Chatbots on the rise: A new narrative in journalism. *Studies in Media and Communication* 7(1): 1–6.

Waytz A, Heafner J and Epley N (2014) The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology* 52: 113–117.

Wölker A and Powell TE (2018) Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. *Journalism*. Epub ahead of print 18 February. DOI: 10.1177/1464884918757072.

Yeo SK, Cacciatore MA and Scheufele DA (2015) News selectivity and beyond: Motivated reasoning in a changing media environment. In: Jandura O, Petersen T, Mothes C, et al. (eds) *Publizistik und gesellschaftliche Verantwortung*. Wiesbaden: Springer, pp. 83–104.

Zanbaka C, Goolkasian P and Hodges L (2006) Can a virtual cat persuade you? The role of gender and realism in speaker persuasiveness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1153–1162. https://doi.org/10.1145/1124772.1124945

Zarouali B, Van Den Broeck E, Walrave M, et al. (2018) Predicting consumer responses to a chatbot on Facebook. *Cyberpsychology, Behavior, and Social Networking* 21(8): 491–497.