Fake News in Social Media: Bad Algorithms or Biased Users?

Franziska Zimmer
Department of Information Science, Heinrich Heine University, Düsseldorf, Germany
E-mail: franziska.zimmer@hhu.de

Mechtild Stock
Stock-Kerpen, Kerpen, Germany
E-mail: mechtildstock@gmail.com

Katrin Scheibe
Department of Information Science, Heinrich Heine University, Düsseldorf, Germany
E-mail: katrin.scheibe@hhu.de

Wolfgang G. Stock*
Department of Information Science, Heinrich Heine University, Düsseldorf, Germany
E-mail: stock@phil.hhu.de

ABSTRACT
Although fake news has been present in human history at any time, nowadays, with social media, deceptive information has a stronger effect on society than before. This article answers two research questions, namely (1) Is the dissemination of fake news supported by machines through the automatic construction of filter bubbles, and (2) Are echo chambers of fake news man-made, and if yes, what are the information behavior patterns of those individuals reacting to fake news? We discuss the role of filter bubbles by analyzing social media's ranking and results' presentation algorithms. To understand the roles of individuals in the process of making and cultivating echo chambers, we empirically study the effects of fake news on the information behavior of the audience, while working with a case study, applying quantitative and qualitative content analysis of online comments and replies (on a blog and on Reddit). Indeed, we found hints on filter bubbles; however, they are fed by the users' information behavior and only amplify users' behavioral patterns. Reading fake news and eventually drafting a comment or a reply may be the result of users' selective exposure to information leading to a confirmation bias; i.e. users prefer news (including fake news) fitting their pre-existing opinions. However, it is not possible to explain all information behavior patterns following fake news with the theory of selective exposure, but with a variety of further individual cognitive structures, such as non-argumentative or off-topic behavior, denial, moral outrage, meta-comments, insults, satire, and creation of a new rumor.

Keywords: fake news, truth, information behavior, social media, filter bubble, echo chamber

Received date: April 18, 2019
Accepted date: June 05, 2019

*Corresponding Author: Wolfgang G. Stock
Professor
Department of Information Science, Heinrich Heine University, Universitätsstraße 1 D-40225 Düsseldorf, Germany
E-mail: stock@phil.hhu.de

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestricdly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.
1. INTRODUCTION

It is a truism that false propositions or even deceptions reach their recipients every day and everywhere. Fake news on online press sites and on social media is no exception. However, deceptive information “has had dramatic effect on our society in recent years” (Volkova & Jang, 2018, p. 575). Deceptions and fake news may possibly survive very well in environments of all kinds of social media, be it weblogs, microblogging services, social live streaming platforms, image and video sharing services, or social networking services. “Despite optimistic talk about ‘collective intelligence’, the Web has helped create an echo chamber where misinformation thrives. Indeed, the viral spread of hoaxes, conspiracy theories, and other false or baseless information online is one of the most disturbing social trends of the early 21st century” (Quattrociocchi, 2017, p. 60), leading even to the “emergence of a post-truth world” (Lewandowsky, Ecker, & Cook, 2017, p. 357). Especially, such historically relevant events as the UK’s Brexit vote (Bastos, Mercea, & Baronchelli, 2018), the 2016 presidential election in the United States (Allcott & Gentzkow, 2017), and the excessive use of the term “fake news” by Donald Trump has led to discussions about the role of fake news in society. The related term “post-truth” was named word of the year for 2016 by the Oxford Dictionaries (2016).

In The Guardian, we read “social media filter bubbles and algorithms influence the election” in Great Britain (Hern, 2017). Similarly, for the Observer, “the problem isn’t fake news, it’s bad algorithms” (Holmes, 2016). The University of Amsterdam’s Master of Media blog addresses filter bubbles as algorithms customizing our access to information (Mans, 2016). These three examples clearly demonstrate what the cause of fake news dissemination is: It is bad algorithms. Nevertheless, one may find divergent opinions in the popular press. The New Statesman claims, “Forget fake news of Facebook: the real filter bubble is you” (Self, 2016). Now, the cause of fake news distribution is the misleading information behavior of individual people, i.e. biased users. As filter bubbles and echo chambers are often discussed in the press Bruns (2019) asks, “are filter bubbles real,” and are they overstated?

“Bad algorithms” are related to “filter bubbles,” being applications of personalized information retrieval as well as of recommender systems. They lead the users to receive only an excerpt of (maybe false) propositions instead of the entire spectrum of appropriate information. A source for concrete algorithmic recommendations is the user’s former information behavior, which is recognized by the machine. On the other hand, “bad user behavior” or “biased users” (Vydiswaran, Zhai, Roth, & Pirolli, 2012) refer to “echo chambers,” which are loosely connected clusters of users with similar ideologies or interests, whose members notice and share only information appropriate to their common interests. The information behavior of the user in question in combination with other users’ behaviors (e.g., commenting on posts or replying to comments) exhibits special patterns which may lead to the echo chamber effect (Bruns, 2017).

2. RESEARCH OUTLINE

First of all, the main concepts must be defined. Fake news is information including “phony news stories maliciously spread by outlets that mimic legitimate news sources” (Torres, Gerhart, & Negahban, 2018, p. 3977); it is misinformation (transmitting untrue propositions, nonconsidering the cognitive state of the sender) and disinformation (again, transmitting untrue propositions, but now consciously by the sender) (Shin, Jian, Driscoll, & Bar, 2018). Deception is a kind of disinformation which brings an advantage to the sender. Other authors compare fake news to satire and parody, fabrication, manipulation, and propaganda (Tandoc Jr., Lim, & Ling, 2018). The users’ appraisement of a news story as fake or non-fake depends on the content of the story and—a little bit more—on the source of the transmitted information (Zimmer & Reich, 2018) as well as on the presentation format (Kim & Dennis, 2018).

This paper follows the well-known definition of social media by Kaplan and Haenlein (2010, p. 61): “Social Media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content.” Social Media includes, among other systems, weblogs, social networking services (such as Facebook), news aggregators (such as Reddit), knowledge bases (such as Wikipedia), sharing services for videos and images (such as YouTube and Instagram), social live streaming services (such as Periscope), and services for knowledge exchange (such as Twitter) (Linde & Stock, 2011, pp. 259ff.). In contrast to such media as newspapers, radio, or TV, in social media there is no formal information dissemination institution (as, say, The New York Times, CBS Radio, or NBC); thus, disintermediation happens. All social media are not immune from fake news (Zimmer, Scheibe, Stock, & Stock, 2019).

A user of Internet services acts as consumer (only receiving content), producer (producing and distributing content), and participant (liking or sharing content) on all kinds of online media (Zimmer, Scheibe, & Stock, 2018). In classical
communication science one speaks of the audience of media; nowadays, especially on social media, audience members are called “users.” Algorithms are sets of rules defining sequences of operations; they can be implemented as computer programs in computational machinery. In this article, the term “algorithm” is only used in the context of computer programs running on “machines.”

Filter bubbles and echo chambers are metaphorical expressions. For Pariser (2011), a filter bubble is a “unique universe of information for each of us.” Pariser lists three characteristics of the relationship between users and filter bubbles, namely (1) one is alone in the bubble, (2) the bubble is invisible, and (3) the user never chose to enter the bubble. We will critically question Pariser's characteristics. For Dubois and Blank (2018, p. 3) a filter bubble means “algorithmic filtering which personalizes content presented on social media.” Davies (2018, p. 637) defines filter bubbles as “socio-technical recursion,” i.e. as an interplay between technologies (as, for instance, search engines or social media services) and the behavior of the users and their social relations.

An echo chamber describes “a situation where only certain ideas, information and beliefs are shared” (Dubois & Blank, 2018, p. 1). Echo chambers occur “when people with the same interests or views interact primarily with their group. They seek and share information that both conforms to the norms of their group and tends to reinforce existing beliefs” (Dubois & Blank, 2018, p. 3). Users in echo chambers are on a media or content “diet” (Case & Given, 2018, p. 116) or in “ideological isolation” (Flaxman, Goel, & Rao, 2016, p. 313) concerning a certain topic. Such isolation may result from selective exposure of information (Hyman & Sheatsley, 1947; Liao & Fu, 2013; Spohr, 2017) and a confirmation bias (Vydiswaran, Zhai, Roth, & Pirolli, 2015; Murungi, Yates, Purao, Yu, & Zhan, 2019). There are different manifestations of selective information exposure; its strongest form is “that people prefer exposure to communications that agree with their pre-existing opinions” (Sears & Freedman, 1967, p. 197). A special kind of selective exposure of information is “partisan selective exposure,” which is related to political affiliations and not—as general selective exposure—based on ideologies or opinions (Kearney, 2019).

Both basic concepts are closely related; however, an echo chamber is more related to human information behavior and a filter bubble is more associated with algorithmic information filtering and results' presentation in online services.

Social media documents are skipping the intermediation process; indeed, “social media enabled a direct path from producers to consumers of contents, i.e., disintermediation, changing the ways users get informed, debate, and shape their opinions” (Bessi et al., 2015, p. 1). Prima facie, this sounds great. However, if we take a look at the other side of the coin, “confusion about causation may encourage speculations, rumors, and mistrust” (Bessi et al., 2015, p. 1). The disappearance of intermediation has not only "fostered a space for direct meetings in a sort of online Habermasian public sphere" (Törnberg, 2018, p. 17), but has also fostered misuse of social media through the publication of fake news by biased users. Habermas himself was always pessimistic about social media (Linde & Stock, 2011, p. 275), as for him weblogs play “a parasitical role of online communication” (Habermas, 2006, p. 423). The disappearance of intermediation also supports the parasitical roles of fake news in social media.

3. RESEARCH MODEL

The different estimations on the causes of fake news dissemination in social media directly lead to our central research question (RQ): Are echo chambers and filter bubbles of fake news man-made or produced by algorithms? To be more precise:

- RQ1: Is the dissemination of fake news supported by machines through the automatic construction of filter bubbles, and if yes, how do such algorithms work?
- RQ2: Are echo chambers of fake news man-made, and if yes, what are the information behavior patterns of those individuals reacting to fake news?

In our research model (Fig. 1), RQ1 is located on the left-hand side and RQ2 on the right hand side. We start searching for false propositions, i.e. fake news, and their dissemination via social media channels. First, we are going to describe processes leading to filter bubbles. A user will be informed of the existence of the false propositions via the push service of the social media platform. The selection of the documents which are shown to the user is controlled by the service's algorithms, which in turn are fed by the user's information behavior patterns and their behavior on the specific service (e.g., forming friendships, giving likes, etc.). It is possible that the interaction between the algorithms and the former user behavior clips only certain aspects of information content while neglecting all other content, thus forming a filter bubble. On Facebook, it is difficult to handle a bypass of the systems' algorithms. However, on other social media services, for instance, weblogs, there is a direct push of (fake) news to users. Following, we direct our attention to echo chambers. The same user can comment on the false propositions or reply to comments about such fake news.
His or her cognitive information behavior patterns may lead to different reactions such as confirmation, denial, moral outrage, and satire. In combination with other users' information behavior (replying to the user's comments or replies, liking them, sharing them, and so on) echo chambers of like-minded users may appear.

As there are two different research questions, this study applies different methods answering them. RQ1 will be evaluated by analyzing the sorting and presentation algorithms of social media by the example of Facebook. For RQ2 the authors performed empirical case study research applying content analysis of comments and replies on fake news distributed via social media channels. The channels disseminating the fake news were a weblog (The Political Insider) and two subreddits of the news aggregator Reddit, namely r/The_Donald and r/worldpolitics. We choose the blog from The Political Insider as it published the fake story on our case (“Hillary Clinton sold weapons to the Islamic State”) for the first time; the subreddit r/The_Donald is clearly addressed to supporters of Donald Trump, while r/worldpolitics is a more liberal subreddit. As a result of this selection we were able to analyze comments from different ideological orientations.

How is our article structured? In the next paragraph, we define our basic terms. As fake news disseminate false propositions, it is necessary to discuss the concept of “truth” in relation to knowledge and information as well as to mediated contexts. In order to analyze and answer RQ1 this paper introduces relevance, pertinence, and ranking algorithms and describes Facebook’s sorting algorithm in detail. To work on RQ2, we empirically studied patterns of cognitive processes of human information behavior in response to fake news. A case study provides us with empirical data of user comments and replies. Then, we describe the applied methods (case study research and content analysis), the empirical findings, and the data analysis. The final paragraph summarizes the main results, confesses limitations, and gives an outlook on further research.

4. KNOWLEDGE, INFORMATION, AND TRUTH

If we want to distinguish between fake (misinformation and disinformation) and non-fake (knowledge) we should know what knowledge, information, and truth are. The corresponding discipline is philosophy, more precisely epistemology. What follows is an excursus on the philosophical foundations of truth. The aim of this paragraph is to show that the definition of truth and the assignment of truth values to empirical statements are anything but easy.

Only a proposition is able to be true or false. In epistemology, one kind of knowledge (“knowing that” in contrast to “knowing how”) is based on true propositions. Chisholm (1977, p. 138) defines knowledge:

$h$ is known by $S$ =df $h$ is accepted by $S$; $h$ is true; and $h$ is nondefectively evident for $S$,

where $h$ is a proposition and $S$ a subject; =df means “equals by definition.” Hence, Chisholm demands that the subject $S$ accepts the proposition $h$ (as true), which is in fact the case.
(objectively speaking) and that this is so not merely through a happy coincidence, but precisely 'nondefectively evident.' Only if all three determinants (acceptance, truth, and evidence) are present, knowledge can be seen as well and truly established. In the absence of one of these aspects, such a statement can still be communicated—as information—but it would be an error (when truth and evidence are absent), a supposition (if acceptance and evidence are given, but the truth value is undecided) or a lie, fake, or deception (when none of the three aspects apply).

Knowledge cannot be transmitted as such; it is in need of a sender, data to be transmitted, a channel, and a receiver. Information dynamically sets knowledge “into motion.” Knowledge always has a truth claim. Is this also the case for information, if information is what sets this knowledge in motion? Is there something like true or false information (Stock & Stock, 2013, p. 39)? Apart from knowledge, there are further, related forms of dealing with objects. If beliefs, conjectures, or fakes are put into motion, are they not information? “Information is not responsible for truth value,” Kuhlen (1995, p. 41) points out. Backlund (1991, p. 50) remarks, “we are unable to say confidently of anything that it could not be information.” Maybe the proposition which is transmitted by information is true or “contingently truthful” (Floridi, 2005); and many information scientists “will generally ignore any distinction between truth or falsity of information” (Case & Given, 2018, p. 67). The task of checking the truth value of the knowledge, rather, must be delegated to the receiving subject S. She or he then decides whether the information retrieved represents knowledge, conjecture, or untruth. Therefore, it is terminologically very problematic to speak of “true/false information,” as only propositions are truth bearers.

Propositions, linguistically presented by declarative sentences, can be true or false. Here, one basic philosophical question arises. Even Pontius Pilate once famously asked “What is truth?” to which Jesus responded—with silence. Truth is a relation between a proposition and a reference object. There are different truth theories working with different reference objects, namely reality, praxis, other propositions in the same system, acceptance inside a community; and, finally, a person’s internal state.

The classical approach to analyze truth is the correspondence theory (David, 1994) theorizing the relation between a proposition and a concrete fact in space and time. Although there are similar definitions of correspondence already in Aristotle’s work, the canonical form of this truth theory originates from the early twentieth century. Bertrand Russell states, “(thus a belief is true when there is a corresponding fact and is false when there is no corresponding fact” (Russell, 1971, p. 129). A person, who will make true propositions on a certain state of affairs in reality, must perceive (watch, hear, etc.) this part of reality personally, in real-time, and on site. In our context of journalism and social media, the person reporting on a state of affairs makes a true proposition ("true" for his self-consciousness) when he luckily is in the right spot at the right time. In times of social media, the term "journalist" includes professional investigative journalism as well as citizen journalists reporting via channels like Facebook, Reddit, Twitter, or Periscope. For the audience of those journalists, there is no chance to verify or to falsify the correspondence between the read or heard proposition in the newspaper, the tweet, or the TV broadcast, and the part of reality, since they simply were not there. This is the reason why the correspondence theory of truth only plays a minor role, if any, in the context of fake or alternative news (Muñoz-Torres, 2012).

Accordance with objective reality and personal awareness is the key factor of the theory of reflection. Whether the human mind contains truth is not a question of theory, but of praxis. In praxis (working, any decision procedure), humans have to prove the truth of their thinking in their practical behavior (Pawlow, 1973). A sentence is true if its proposition works in practice. The problem with the theory of reflection is that it is impossible to consider all facts because they are always a product of selection. A problem of the media is that it sometimes takes a while to gather all facts to accurately use them in practice. By the time the facts were gathered the media momentum has passed.

The coherence theory of truth declares that one statement corresponds with another statement, or with the maximal coherent sum of opinions and accepted clauses of statements (Neurath, 1931). There cannot be an opposite statement within an already accepted system of statements. If the statement can be integrated, it is true, otherwise it is false. However, instead of rejecting the new statement, it is possible to change the whole system of statements to integrate the latest one into the system. The statements need to be logically derivable from each other.

The definition of the consensus theory of truth states that truth is what is agreed upon by all people in a group. First, the speakers need to be clear about what they are saying to ensure everyone understands what they mean, they insinuate each other’s truthfulness, and their words are accurate. A discourse needs to determine if the claim of the speaker is indeed to be accepted. Everyone needs to have the same level of influence to rule or to oppose (Habermas, 1972). Relying only on the consensus theory of truth is difficult and does not necessarily lead to the truth in the sense of the correspondence theory.

Brentano (1930) describes the evidence theory of truth, “When I have evidence, I cannot err.” A judgement is true if
it expresses a simple quality of experience. Brentano adheres to the traditional view that there are two different ways for a judgement to be evident; either it is immediately, or it is evident insofar as it is inferable from evident judgements by applications of evident rules. But, evidence is a primitive notion; it cannot be defined, it is only experienceable, and thus, found in oneself.

The philosophical truth theories illustrate that truth or lies are in the eye of the beholder (evidence theory), the praxis (theory of reflection), the community (consensus theory), or in the system of accepted propositions (coherence theory). As the correspondence theory of truth is not applicable in the environments of journalism and social media, we have big problems in stating what exactly is true and what is not. If we do not know what the truth is, we also cannot know exactly what “fake news” is. It is the individual person who decides, based on a (probably unknown) truth theory, what is considered as truth, as lies, as “true news,” and as “fake news.” By the way, attempts of automatic semantic deception detection (e.g., Conroy, Rubin, & Chen, 2015) are faced with the same problems, especially when they rely on the coherence or the consensus theory of truth.

5. FAKE NEWS DISSEMINATION THROUGH ALGORITHMIC FILTER BUBBLES (RQ1)

The concept of relevance is one of the basic concepts of information science (Saracevic, 1975). Users expect an information system to contain relevant knowledge, and many information retrieval systems, including Internet search engines and social media services, arrange their results via relevance ranking algorithms. In information science, researchers distinguish between objective and subjective information needs. Correspondingly to these concepts, we speak of relevance (for the former) and pertinence (for the latter), respectively.

Since relevance always aims at user-independent, objective observations, we can establish a definition: A document, for instance, a website, a blog post, a post on Facebook or Reddit, or a microblog on Twitter (or, to speak more precisely, the knowledge contained therein) is relevant for the satisfaction of an objective (i.e. subject-independent) information need.

A research result can only be pertinent if the user has the ability to register and comprehend the knowledge in question according to his or her cognitive model. Soergel (1994, p. 590) provides the following definition: “Pertinence is a relationship between an entity and a topic, question, function, or task with respect to a person (or system) with a given purpose. An entity is pertinent if it is topically relevant and if it is appropriate for the person, that is, if the person can understand the document and apply the information gained.” Pertinence ranking presupposes that the information system in question is able to identify the concrete user who works with the system; it is always subject-dependent personalized ranking (Stock & Stock, 2013, pp. 361ff).

We describe only one paradigmatic example of ranking in social media, namely the algorithms of Facebook as the most common social media platform. Facebook's sorting of posts is a pertinence ranking algorithm; it works with the three factors affinity, weighting, and timeliness. According to these three aspects, a user will see posts on her or his Facebook page with the posts sorted in descending order of their retrieval status values (Zuckerberg et al., 2006). Affinity is concerned with the user’s previous interactions on the posting pages, whereas different interactions are weighted variously. If a user X frequently views another user’s (say, user A) posts, likes them, comments on them, or shares them. A’s future posts—depending on their weights (resulting from the numbers of likes, shares, and comments)—get a higher weight for user X. Facebook also considers the position of the creator of the post (is this user often viewed, annotated, etc.? and the nature of the post (text, image, or video). The timeliness states that a contribution becomes more important the newer it is. However, other factors play a role, and the algorithm is constantly being adapted. For example, an already viewed ranked list is not displayed a second time in exactly the same order (i.e., the criteria for the sorting are each slightly modified) in order to make the lists more interesting. Also, posts from people (as opposed to those from companies) are weighted higher, and the spatial proximity between the receiver and the sender of the post plays an important role. In particular, the affinity causes a user to see the one source at the top of his or her list, which he or she has often viewed in previous sessions.

Ranking on Facebook is always personalized and based on the user’s common interests, her or his information behavior on the service, and her or his Facebook friends (Tseng, 2015; Bakshy, Messing, & Adamic, 2015). The more a user repeatedly clicks on the posts of the same people, the more the selection of posts stabilizes, which always appear at the ranking’s top positions. Thus, in a short time—with high activity on Facebook—an information diet may occur that presents users only those posts on top of their pages, whose creators they prefer. So it can be assumed that such personalized content representation leads to “partial information blindness (i.e., filter bubbles)” (Haim, Graefe, & Brosius, 2018, p. 330).

It depends on the user to form a “friendship” on Facebook, and it is on the user to often select certain friends’ posts
for reading, liking, sharing, and commenting. Facebook’s pertinence ranking algorithm indeed may amplify existing behavioral patterns of the users into filter bubbles and then into echo chambers, whereby the information behavior of the users plays the important primary role. In contrast to the assumptions of Pariser (2011) on filter bubbles, (1) no one is alone in the bubble when the bubble leads to echo chambers (where other users are by definition); (2) the bubble is visible to certain users insofar as they figured out Facebook’s ranking methods; for other, rather uncritical users, the bubble is indeed invisible; (3) the users’ behavior feeds the pertinence ranking algorithms; therefore, the users (consciously or unintentionally) cooperate with the service entering the bubble through their own information behavior.

Here we arrive at a first partial result and are able to answer RQ1: Algorithms by themselves do not produce filter bubbles or subsequently echo chambers, they only consolidate the users’ information behavior patterns. Concerning the reception of fake news, it is not possible to argue that they are solely distributed by “bad algorithms,” but by the active collaboration of the individual users. Also, Del Vicario et al. (2016, p. 554f.), for instance, found out that “content-selective exposure is the primary driver of content diffusion and generates the formation of homogeneous clusters, i.e., ‘echo chambers.’” DiFranzo and Gloria-Garcia (2017, p. 33F.) arrive at a similar result: “The related filter-bubble effect is due to the user’s network and past engagement behavior (such as clicking only on certain news stories), that is, it is not the fault of the news-feed algorithm but the choices of users themselves.” There are results concerning fake news and the algorithms of Facebook: “While this criticism has focused on the ‘filter bubbles’ created by the site’s personalisation algorithms, our research indicates that users’ own actions also play a key role in how the site operates as a forum for debate” (Seargeant & Tagg, 2019, p. 41). Although algorithms are able to amplify human information behavior patterns, obviously, the users play the leading role concerning construction and maintenance of those bubbles of (fake) news. Indeed, there are filter bubbles; however, they are fed by users’ information behavior and—more important—they are escapable (Davies, 2018).

6. FAKE NEWS DISSEMINATION THROUGH MAN-MADE ECHO CHAMBERS (RQ2)

6.1. Our Approach

When we want to analyze echo chambers of fake news and also believing as well as mistrusting such false propositions by individual persons, we have to study their cognitive processes in detail. In our research study, we apply case study research and content analysis. As we want to investigate which concrete cognitive information behavior patterns concerning fake news exist, we start our endeavors with the help of concrete cases. Case study researchers “examine each case expecting to uncover new and unusual interactions, events, explanations, interpretations, and cause-and-effect connections” (Hays, 2004, p. 218f.). Our case includes a (probably fake) post and comments as well as replies to it. It is a story on Hillary Clinton selling weapons to the Islamic State. With the help of this singular case study (Flyvbjerg, 2006) we try to find cognitive patterns and to understand users’ information behavior at the time shortly after the publication of fake news.

To analyze the cognitive patterns of the commenting users, we look upon the results of the cognitive processes, i.e. the texts (as we are not able to measure the human cognitive patterns directly) and apply quantitative and qualitative content analysis (Krippendorff, 2018) of posts in social media. In quantitative content analysis, the occurrence of the categories in the coding units is counted and, if necessary, further processed statistically; the qualitative content analysis turns to the statements within the categories, namely the “manifest content” (Berelson, 1952) and the “deeper meaning” (such as subjective senses), as well as formal textual characteristics such as style analysis (Mayring & Fenzl, 2019). In order to create the appropriate categories for the content analysis, we applied both (1) inductive (or conventional) as well as (2) deductive (or directed) measures (Elo & Kyngas, 2008; Hsieh & Shannon, 2005). By (1) applying the conventional approach with a first and preliminary analysis of comments concerning our case, we defined the first codes; and we (2) arrived at codes while studying relevant published literature. The coding unit was the single comment or the single reply. Every coding unit was coded with only one (the best fitting) category. The coding process was led by a short code book and conducted by two of the article’s authors in August 2018, whereas all steps were performed intellectually. In a first round, the coders worked independently (resulting in Krippendorff’s alpha > 0.8, signaling the appropriateness of the code book and the coders’ work); in a second round, the (few) disagreements were discussed and solved (Mayring & Fenzl, 2019, p. 637). In the end, there was an intercoder consistency of 100%, i.e., Krippendorff’s alpha was 1.

Our approach is similar to research in microhistory describing posts and comments on social networking services in order to find information on historically relevant—especially local—events and developments (Stock, 2016, 2017). Similar
to our approach, Walter, Brüggemann, and Engesser (2018) studied user comments in echo chambers concerning the topic of climate change. Gilbert, Bergstrom, and Karahalios (2009) defined agreement as manifestation of an echo chamber. They found that about 39% of all comments agree with the blog author; 11% disagree, and half of all commentators react in other ways. Murungi et al. (2019, pp. 5192f.) found that significant amounts of comments on a concrete political situation (Roy Moore’s candidacy for the U.S. Senate in Alabama in 2017) were non-argumentative.

For our case study, we consulted a weblog (The Political Insider, a right-wing oriented web site; August 2016) (N=43) and Reddit as the current most popular news aggregator (Zimmer, Akyürek et al., 2018). To be more precise, we analyzed Reddit’s subreddits r/The_Donald (a forum “for Trump supporters only”; September 2016) (N=177) and r/worldpolitics (a “free speech political subreddit”; September 2016) (N=246). We checked all comments and all replies to the comments manually. All in all, we analyzed 466 documents. Studying literature and empirical material, we found different patterns of information behavior in response to fake news and applied them as codes for our content analysis:

- Confirmation: broad agreement with post, attempt of verification
- Denial: broad disagreement with post, attempt of falsification
- Moral outrage: questioning the posts, comments and replies from a moral point of view
- New rumor: creation of a new probably false proposition
- Satire: satirical, ironic, or sarcastic text
- Off-topic: non-argumentative, ignoring the discussion, arguing on other topics, broad generalization
- Insult: defamation of other people or groups
- “Meta” comment/reply: discussing the style of another post, offense against a commentator

Additionally, we evaluated the topic-specific orientation (positive, negative, and neutral) for all texts. Positive means an articulated or implicated agreement with the original post. If a comment, for instance, argues, “Clinton should be arrested” in response to the post “Hillary Clinton sold weapons to ISIS,” it is counted as positive. Neutral means that there is no relation to the concrete topic of the triggering post, e.g., “Obama is born in Kenya” as a comment on “Clinton sold weapons.” All other texts were coded as negative, e.g., “What’s there to say? It’s just a vague, unfounded accusation.” We aggregated all generations of replies (replies to a comment, replies to a reply) into the code “reply.”

6.2. Results

Tables 1-3 exhibits our descriptive results for the three selected sources, namely The Political Insider, r/The_Donald, and r/worldpolitics. Concerning our case study, most comments on The Political Insider are confirmations of the (false) proposition; likewise, the comments’ orientation is predominantly positive (Table 1). In both analyzed subreddits most comments (about 40% to 50%) and even more replies (about 70% to 80%) are non-argumentative or off-topic (Tables 2 and 3). In the subreddit r/The_Donald we found about 40% agreement with the fake proposition for the comments; however, only 8% existed for the replies.

About half of the comments in r/The_Donald express a neutral orientation, and the other half a positive one; while most of the replies were neutral. Most comments and more than 80% of the replies in r/worldpolitics are off-topic and express no orientation concerning the given topic (i.e., the triggering post). The authors of r/worldpolitics are more critical than those of r/The_Donald as about 30% of all comments were classified as denial (in contrast to 0% in r/The_Donald).

The dominating cognitive patterns are non-argumentative or arguments being off-topic. The very first comment on r/worldpolitics was “time to put up or shut up,” which diverse authors regarded as an invitation to speculate on different political topics with loose or no relationship to the content of the post. We can find rather senseless texts as, e.g., “LOL who knew,” “Holy shit!!”, or “Trump was right all along” (all from r/The_Donald). However, most of the off-topic comments and replies pursue a similar tendency, most notably attacking Obama and praising Trump in r/The_Donald or discussing the DNC (Democratic National Committee) in r/worldpolitics. Confirmations of the fake news are frequent in The Political Insider and r/The_Donald, but not in r/worldpolitics. Here are some examples: “Done, done, DONE! Round up his people”—“Traitors are hanged from the highest tree!”—“His eyes were always cold to me … soulless. It is no surprise that Obama would be the founder of ISIS, really.” Confirmations culminate in death threats: “Put him [i.e., Obama] to death. Period. Let the left cry. They will never agree that they are wrong, that he was a criminal. It doesn’t matter. He is a traitor to this country, and if these allegations are true, he needs to be appropriately punished” (all from r/The_Donald).

Sometimes, commentators are dissatisfied with the discussion and argue from a meta position as “I’m really not interested in engaging in a totally off-topic argument with you”; “What? Seriously you believe this?”; or “Why have you sent me an article about how George Bush, the Republican president, may have rigged the 2004 election as evidence that Hillary Clinton, the
Table 1. Users’ cognitive patterns in reactions to fake news: The Political Insider

<table>
<thead>
<tr>
<th>Cognitive pattern</th>
<th>Comments</th>
<th>Replies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmation</td>
<td>33.3%</td>
<td>23.1%</td>
</tr>
<tr>
<td>Denial</td>
<td>3.3%</td>
<td>-</td>
</tr>
<tr>
<td>Moral outrage</td>
<td>3.3%</td>
<td>-</td>
</tr>
<tr>
<td>New rumor</td>
<td>13.3%</td>
<td>15.4%</td>
</tr>
<tr>
<td>Satire</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Off-topic</td>
<td>26.6%</td>
<td>61.5%</td>
</tr>
<tr>
<td>Insult</td>
<td>20.6%</td>
<td>-</td>
</tr>
<tr>
<td>“Meta”</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Positive orientation</td>
<td>73.3%</td>
<td>46.2%</td>
</tr>
<tr>
<td>Negative orientation</td>
<td>3.3%</td>
<td>-</td>
</tr>
<tr>
<td>Neutral orientation</td>
<td>23.3%</td>
<td>53.8%</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>13</td>
</tr>
</tbody>
</table>

Post: “Wikiileaks CONFIRMS Hillary Sold Weapons to ISIS… Then Drops Another BOMBSHELL! Breaking News”

Table 2. Users’ cognitive patterns in reactions to fake news: r/The_Donald

<table>
<thead>
<tr>
<th>Cognitive pattern</th>
<th>Comments</th>
<th>Replies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmation</td>
<td>40.8%</td>
<td>7.9%</td>
</tr>
<tr>
<td>Denial</td>
<td>-</td>
<td>4.0%</td>
</tr>
<tr>
<td>Moral outrage</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>New rumor</td>
<td>5.3%</td>
<td>5.0%</td>
</tr>
<tr>
<td>Satire</td>
<td>1.3%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Off-topic</td>
<td>47.4%</td>
<td>78.2%</td>
</tr>
<tr>
<td>Insult</td>
<td>5.3%</td>
<td>3.0%</td>
</tr>
<tr>
<td>“Meta”</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Positive orientation</td>
<td>48.7%</td>
<td>11.9%</td>
</tr>
<tr>
<td>Negative orientation</td>
<td>-</td>
<td>5.0%</td>
</tr>
<tr>
<td>Neutral orientation</td>
<td>51.3%</td>
<td>83.2%</td>
</tr>
<tr>
<td>N</td>
<td>76</td>
<td>101</td>
</tr>
</tbody>
</table>

Post: “Breaking Assange: Obama & Clinton not only supplied ISIS with a billion dollars worth of weapons annually, they paid these mercenaries salaries! Obama employed ISIS… let it sink in. Obama was the real leader of ISIS!”

Table 3. Users’ cognitive patterns in reactions to fake news: r/worldpolitics

<table>
<thead>
<tr>
<th>Cognitive pattern</th>
<th>Comments</th>
<th>Replies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmation</td>
<td>12.5%</td>
<td>9.1%</td>
</tr>
<tr>
<td>Denial</td>
<td>29.2%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Moral outrage</td>
<td>-</td>
<td>1.0%</td>
</tr>
<tr>
<td>New rumor</td>
<td>2.1%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Satire</td>
<td>4.2%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Off-topic</td>
<td>43.8%</td>
<td>72.2%</td>
</tr>
<tr>
<td>Insult</td>
<td>2.1%</td>
<td>0.5%</td>
</tr>
<tr>
<td>“Meta”</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Positive orientation</td>
<td>6.3%</td>
<td>10.1%</td>
</tr>
<tr>
<td>Negative orientation</td>
<td>14.6%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Neutral orientation</td>
<td>31.3%</td>
<td>6.6%</td>
</tr>
<tr>
<td>N</td>
<td>48</td>
<td>198</td>
</tr>
</tbody>
</table>

Post: “Julian Assange: ‘1,700 emails’ proves Hillary Clinton sold weapons to ISIS in Syria.”

Democratic candidate, has rigged the upcoming election?” (all from r/worldpolitics).

Some (however few) comments are insults, as, for instance, “Yet more proof that the people at the very top are, for all practical purposes, gangsters” (r/worldpolitics); “Obama is a piece of shit Globalist muslim”; or “Aw, come on. Whadaya expect from a f**kin’ Kenyan ‘born’ in Hawaii, raised in Indonesia, programmed and sponsored by the Saudi Manchurian School for Gifted Leftists?” (both from r/The_Donald).

Here, a further cognitive pattern comes into play: the construction of a new rumor, for example: “The Hawaiian birth certificate (of Obama, a/n) was proven to be a forgery”; “Obama’s entire life is pure fiction, a 100% CIA creation”; “Hillary is the Mother of ISIS”; “They (Obama and Clinton, a/n) wanted this war in Syria, they wanted the refugee influx”; or, “It will take a while before people admit that Obama and Michelle and the supported ‘daughters’ were all fake”; “Malia’s and Sasha’s biological parents have always been nearby while the girls provided a fictional family for Barack and Michelle” (all from r/The_Donald).

Some comments and replies consist of satirical, ironic, or sarcastic text, as for instance: “Of course, president Hussein was the head of Isis. He’s a muzlim [sic]” (r/The_Donald); “Is that really how your brain works? Or you just playin’?”; “Someone feel like pointing to some of those emails? Julian? Anybody? Like most Americans, I am too stupid and lazy to spend four years reading emails”; “This news article is great, and absolutely 100% real. I can’t wait to see this actually real story break worldwide, because Hillary absolutely sold weapons to ISIS in Syria, and this is not at all a conspiracy theory!” (all from r/worldpolitics). Sometimes it is problematic to identify irony; however, considering the context the pattern becomes visible.

In the subreddit r/worldpolitics (but with next to nothing in The Political Insider and r/The_Donald) we found critical denials of the fake news as, for instance, “get suspicious when it’s only niche websites reporting stuff like this. If there were real evidence, every conservative site would make a front page”; or, “1700 mails about Libya proof that Hillary sold weapons to Isis in Syria? I don’t mean to comment on the allegations but I hate it when headlines are clearly bullshit.”

A rather uncommon pattern in this case study is moral outrage, a kind of meta-comment from a moral point of view, for instance: “All of you are blaming Hillary and President Obama. They have to get approval from Congress to do this stuff” (The Political Insider); or, “What’s there to say? It’s just a vague, unfounded accusation” (r/worldpolitics).

There are different distributions of cognitive patterns regarding the level of discussion, i.e. between the first generation
of texts (comments on the triggering fake news) and the next generations (replies to the comments and replies to other replies). There are much more non-argumentative and off-topic replies than off-topic comments (The Political Insider: 62% versus 27%; r/The_Donald: 78% versus 47%; r/worldpolitics: 72% versus 44%). And there are less confirmative replies than confirmative comments (The Political Insider: 23% versus 33%; r/The_Donald: 8% versus 41%; r/worldpolitics: 9% versus 13%). Additionally, the users’ information behavior is drifting from positive or negative orientation at the comments’ level to an enhanced neutral orientation at the replies’ level.

6.3. Are There Indeed Echo Chambers?

What can we learn from our case study? Do users indeed live inside an echo chamber? The answer depends on the concrete operationalization of the “echo chamber.” If we narrowly define this concept as a community with high confirmation rates (in our case: for fake news) in combination with high degrees of positive topic-specific orientation (and further with the creation of new rumors with the same direction as the original fake), there are indeed hints for the existence of such communities. A third of the commentators of The Political Insider and about two-fifths of the commenting audience of r/The_Donald seem to argue inside their echo chambers. However, we can define “echo chamber” more broadly. As we know from the texts, off-topic comments and most of the neutral-orientation texts argue in the same direction as the entire community; therefore the filter bubble may include most of these comments and replies. The content of the specific (false) proposition is entirely clear and taken for granted, so users lose the specific thread (from the triggering post); however, they do not lose the (ideological or political) direction. In the sense of this broad definition, depending on the source, up to about 90% of comments (sum of confirmations and off-topic comments) in r/The_Donald, about 60% in The Political Insider, and about 55% in r/worldpolitics exhibit hints towards the existence of echo chambers in those social media channels. In contrast to Bruns (2019) we found that the problems concerning filter bubbles and echo chambers are not overstated, but basic facts in our contemporary online world.

7. CONCLUSION

As the correspondence theory of truth is not applicable in mediated contexts, there remain truth theories which heavily depend on the community (consensus theory) and on the coherence of propositions (coherence theory), but do not point to the truth. This annoying fact does not make research on fake news easy.

Algorithms (and their mechanisms to form filter bubbles) applied in social media themselves do not form communities purely on their own as they amplify users’ information behavior. The crucial element of fake news and their pathways into social media is mainly the individual users, their cognitive patterns, and their surrounding echo chamber (Zimmer, 2019).

Reading (fake) news and eventually drafting a comment or a reply may be the result of users’ selective exposure to information (Frey, 1986; Sears & Freedman, 1967) leading to preferring news (including fake news) fitting their pre-existing opinions. If users take the (false) proposition as given, discuss it uncritically, ignore other opinions, or argue further off-topic (however, always in the same direction), an echo chamber can be formed and stabilized. In contrast to some empirical findings on echo chambers (Fischer et al., 2011; Garrett, 2009; Nelson & Webster, 2017) we found clear hints for the existence of such communities. Depending on the concrete operationalization of the “echo chamber,” about one third to two-fifths (a narrow definition) and more than half of all analyzed comments and replies (a broad definition) can be located inside an echo chamber of fake news. Explicitly expressed confirmation depends on the stage of discussion. In the first stage (comments), confirmative texts are more frequent than in further stages (replies).

Confirmative information behavior on fake news goes hand in hand with the consensus and the coherence theory of truth. The (in the sense of the correspondence theory of truth basically false) proposition will be accepted “by normative social influence or by the coherence with the system of beliefs of the individual” (Bessi et al., 2015, p. 2). This behavior leads directly to a confirmation bias. Our results are partly in line with the theory of selective exposure of information.

However, it is not possible to explain all information behavior following fake news with the theory of selective exposure, but with a variety of further individual cognitive patterns. We were able to identify cognitive patterns clearly outside of echo chambers as denial, moral outrage, and satire—all in all patterns of critical information behavior.

This study has (as every scientific endeavor) limitations. In the empirical part of the study, we analyzed comments and replies to comments on social media. The publication of a comment or a reply on an online medium follows a decision-making process (should I indeed write a comment or a reply?). With our method, we are only able to gather data on individuals who have written such texts; all others remain unconsidered. We did not talk to the commenting and replying individuals. Therefore, we
were not able to ask for intellectual backgrounds, motivations, and demographic details of the commentators.

In this article, we report about one case study only, so the extent of the empirical data is rather limited. Although we collected and intellectually coded some hundreds of texts, this is like a drop in the bucket when faced with millions of posts, comments, and replies on social media. A serious methodological problem (not only ours, but of all research relying on data from the Internet) is the availability of complete data sets on, for instance, a fake news story and all the comments and replies on the fake news, as users and website administrators often delete discriminating posts, comments, or replies. We indeed found hints for deleted posts, comments, or replies on *The Political Insider* as well as on Reddit. In lucky cases (as in our study: the post and the comments of *The Political Insider*), one will find some deleted data on web archives.

Here are some recommendations for future research. As we only analyzed texts on fake news in order to find cognitive reaction patterns, research should also study in analogous ways reactions to true propositions. Are there the same cognitive patterns? People do not only live in the online world. Of course, their lives in the physical world are influenced by family members, friends, colleagues, and other people. As there are empirical hints on the geographic embedding of online echo chambers (Bastos et al., 2018), it would be very helpful to analyze offline echo chambers and the interplay between online and offline echo chambers as well. We distinguished between comments and replies and found different cognitive patterns of the respective authors. Are there indeed different cognitive patterns while writing posts, formulating comments, and phrasing replies to the comments? How can we explain those differences?

*What is new in this paper?* As algorithms (as, for instance, Facebook’s ranking algorithm) only amplify users’ information behavior, it is on the individuals themselves to accept or to deny fake news uncritically, to try to verify or to falsify them, to ignore them, to argue off-topic, to write satire, or to insult other users. If filter bubbles are made by algorithms and echo chambers by users, the echo chambers influence the filter bubbles; however, filter bubbles strengthen existing echo chambers as well. There are different cognitive patterns of the individual users leading to different reactions to fake news. Living in echo chambers (namely the uncritical accepting of the news due to the users’ pre-existing opinions shared within a group or compared with a set of propositions) indeed is a typical, but not the only cognitive pattern.

Therefore, a “critical user” seems to be the decisive factor in identifying and preventing fake news. Our analysis at the beginning of this paper has shown that there is no satisfying answer to what can be considered *the* truth in media. In the end—and this is in line with Chisholm’s (1977) definition of knowledge—it is just a critical user who compares sources and validates the timeliness and evidence of a contribution before believing, denying, or ignoring it and then deciding whether it is true or false. So, finally, it is on the individual user’s critical literacy, information literacy, digital literacy, and media literacy in order “to help cultivate more critical consumers of media” (Mihailidis & Viotty, 2017, p. 441) and, additionally, on libraries and information professionals to instruct their users “in the fight against fake news” (Batchelor, 2017, p. 143) and to “become more critical consumers of information products and services” (Connaway, Julien, Seadle, & Kasprak, 2017, p. 554). Libraries, next to schools (Gust von Loh & Stock, 2013), are faced with the task to educate and instruct people to become critical users.

**REFERENCES**


Fake News in Social Media: Bad Algorithms or Biased Users?


