ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

ECHO CHAMBERS ON FACEBOOK

Walter Quattrociocchi
Antonio Scala
Cass R. Sunstein

Discussion Paper No. 877

09/2016

Harvard Law School
Cambridge, MA 02138

Echo Chambers on Facebook

Walter Quattrociocchi[1,2], Antonio Scala[1,2], Cass R. Sunstein[3]

1 Laboratory of Computational Social Science, IMT Lucca, 55100, Lucca Italy
2 Institute of Complex Systems, CNR, 00100, Rome Italy
3 Harvard Law School, Cambridge, MA, US

Abstract

Do echo chambers actually exist on social media? By focusing on how both Italian and US Facebook users relate to two distinct narratives (involving conspiracy theories and science), we offer quantitative evidence that they do. The explanation involves users' tendency to promote their favored narratives and hence to form polarized groups. Confirmation bias helps to account for users' decisions about whether to spread content, thus creating informational cascades within identifiable communities. At the same time, aggregation of favored information within those communities reinforces selective exposure and group polarization. We provide empirical evidence that because they focus on their preferred narratives, users tend to assimilate only confirming claims and to ignore apparent refutations.

Introduction

Do echo chambers exist on social media? To answer this question, we compiled a massive data set to explore the treatment of two distinct narratives on Facebook, involving the spread of conspiracy theories and scientific information.

It is well-established that many people seek information that supports their current convictions[1,2] - the phenomenon of "confirmation bias," That phenomenon significantly affects decisions about whether to spread content, potentially creating informational

[1] Mocanu, Delia et al. "Collective attention in the age of (mis) information." *Computers in Human Behavior* 51 (2015): 1198-1204.
[2] Bessi, Alessandro et al. "Science vs conspiracy: Collective narratives in the age of misinformation." *PloS one* 10.2 (2015): e0118093.

cascades within identifiable communities[3,4]. In these circumstances, online behavior can promote group polarization [5,6,7].

To explore the role of confirmation bias in the selection of content, we test how users who are interested in information involving conspiracy theories respond to a) intentionally false claims that deliberately mock conspiracy stories, even though they apparently confirm their narratives and to b) debunking information – i.e. attempts to correct unverified rumors [8,9]. We find that intentionally false claims are accepted and shared, while debunking information is mainly ignored. As a result, exposure to debunking information may even increase the commitments of users who favor conspiracy theories. We also compare the reception of scientific information to the reception of conspiracy theories, showing how Facebook users create communities of like-minded types.

The paper is structured as follows. In section 1 we describe the datasets. In section 2 we discuss the evidence of echo chambers on Facebook in both Italy and the United States. In section 3 we show the power of confirmation bias by measuring the susceptibility of conspiracy users to both confirming and debunking information.

Setting Up the (Data) Experiment

Conspiracy theories often simplify causality and reduce the complexity of reality. Such theories may or may not be formulated in a way that allows individuals to tolerate a certain level of uncertainty. Of course some conspiracy theories turn out to be true. Scientific information disseminates advances and exposes the larger public to how scientists think. Of course some scientific information turn out to be false.

The domain of conspiracy theories is exceptionally wide, and sometimes the arguments on their behalf invoke explanations designed to replace scientific evidence. The conspiracy theories traced here involve the allegedly secret plots of "Big Pharma"; the power and plans of the "New World Order"; the absence of a link between HIV and AIDS (and the conspiracy to make people think that there is such a link); and cancer

[3] Anagnostopoulos, Aris et al. "Viral misinformation: The role of homophily and polarization." *arXiv preprint arXiv:1411.2893* (2014).
[4] Del Vicario, Michela et al. "The spreading of misinformation online." *Proceedings of the National Academy of Sciences* 113.3 (2016): 554-559.
[5] Zollo, Fabiana et al. "Emotional dynamics in the age of misinformation." *PloS one* 10.9 (2015): e0138740.
[6] Bessi, Alessandro et al. "Trend of Narratives in the Age of Misinformation." *PloS one* 10.8 (2015): e0134641.
[7] Bessi, Alessandro et al. "The economy of attention in the age of (mis) information." *Journal of Trust Management* 1.1 (2014): 1-13.
[8] Zollo, Fabiana et al. "Debunking in a World of Tribes." *arXiv preprint arXiv:1510.04267* (2015).
[9] Bessi, Alessandro et al. "Social determinants of content selection in the age of (mis) information." *Social Informatics* (2014): 259-268.

cures. By contrast, the scientific news reports on the most recent research findings, such as the discovery of gravitational waves and the Higgs boson.

To produce the data set, we built a large atlas of Facebook pages, with the assistance of various groups (Skepti Forum, Skeptical spectacles, Butac, Protesi di Complotto), which helped in labeling and sorting both conspiracy and scientific sources. (We emphasize that other kinds of data sets may not show the particular patterns that we observe here.) To validate the list, all pages have been manually checked looking at their self-description and the type of promoted content. We analyzed users' interaction through Facebook posts with respect to these two kinds of information over a time span of five years (2010-2014) in the Italian and US contexts (see Table 1 for a breakdown of the dataset). Note that the list refers to public Facebook pages dedicated to the diffusion of claims from the two kinds of narratives. Some examples of science pages are https://www.facebook.com/ScienceNOW (2 million likes); https://www.facebook.com/livescience (1.2 million likes); and https://www.facebook.com/sciencenews (2.5 million of likes). Some examples of conspiracy theory pages are https://www.facebook.com/TheConspiracyArchives (200k likes) and https://www.facebook.com/CancerTruth-348939748204/ (250k likes). Numbers reported in Table 1 refer to the posts total number of likes and comments to each page's post on the overall time window.

We measured the reaction of Facebook users who are exposed to different posts, in particular:

a) For Italy, troll posts, sarcastic, and paradoxical messages mocking conspiracy thinking (e.g., chem-trails containing Viagra)

b) For the United States, debunking posts, involving information attempting to correct false conspiracy theories circulating online.

FB ITALY	TOTAL	SCIENCE	CONSPIRACY	TROLL
Pages	73	34	39	2
Posts	271,296	62,705	208,591	4,709
Likes	9,164,781	2,505,399	6,659,382	40,341
Comments	1,017,509	180,918	836,591	58,686
Likers	1,196,404	332,357	864,047	15,209
Commentsers	279,972	53,438	226,534	43,102

FB USA	TOTAL	SCIENCE	CONSPIRACY	DEBUNKING
Pages	478	83	330	66
Posts	679,948	262,815	369,420	47,780
Likes	603,332,826	453,966,494	145,388,117	3,986,922
Comments	30,828,705	22,093,692	8,304,644	429,204
Likers	52,172,855	39,854,663	19,386,131	702,122
Commentsers	9,790,906	7,223,473	3,166,726	118,996

Table 1. Breakdown of the Italian and US Facebook datasets grouped by page category.

Polarized Communities

On Facebook, actions like "share," "comment," or "like" have distinctive meanings. In most cases, a "like" stands for a positive feedback to the post; a "share" expresses the desire to increase the visibility of a given information; and a "comment" reflects a contribution to an online debate, which may contain negative or positive feedback to the post.

Our analysis shows that in these domains, users are highly polarized and tend to focus their attention exclusively on one of the two types of information. We also find that users belonging to different communities tend not to interact and that they tend to be connected only with like-minded people.

More precisely, we analyzed users' engagement with respect to content as the percentage of a user's "likes" on each content category. We considered a user to be polarized in science or conspiracy narratives when 95% of his "likes" is on either conspiracy or science posts. With this stringent test, we find that the most users are highly polarized, with especially high percentages on conspiracy posts: there are 255,225 polarized users on scientific pages (76.79% of all users who interacted on scientific pages), and there are 790,899 polarized users on conspiracy pages (91.53% of all users who interacted with conspiracy posts).

Figure 1 shows the probability density function (PDF) of users' polarization. We found that there are distinct communities that correspond to the two sharp peaks near ρ = -1 (science) and ρ = 1 (conspiracy).

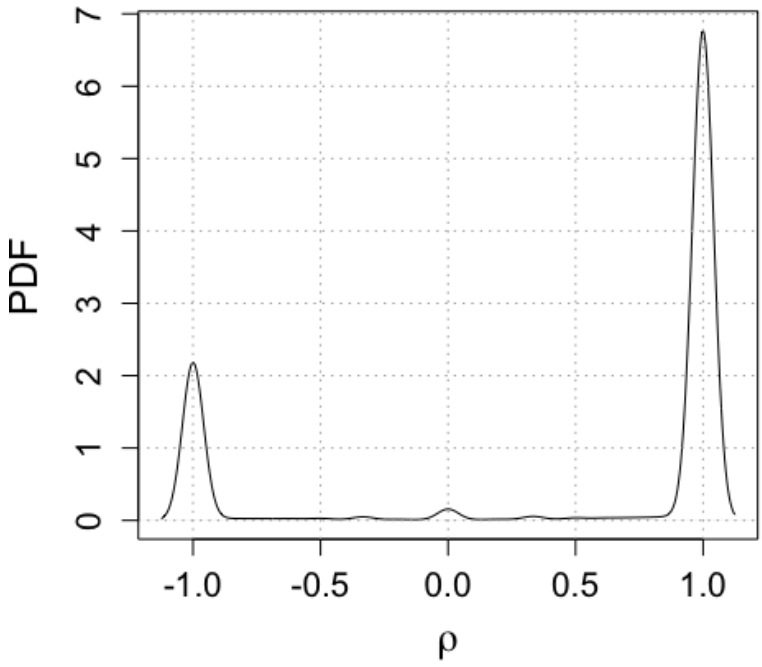


Figure 1: Users are polarized. The probability density function (PDF) of the frequency that a user has polarization ρ is remarkably concentrated in two peaks near the values ρ = -1 (science) and ρ = 1 (conspiracy), indicating that users are split into two distinct communities.

In short, the Facebook users we studied mainly focus on a single type of narrative, at least in the contexts studied here. As a further step, we tested whether different narratives present different information consumption patterns. Figure 2 shows the statistics CCDF (Complementary Cumulative Distribution Function) for likes, comments, shares and post lifetime for both types of information.

The shape of the CCDFs is remarkably similar, indicating that conspiracy and scientific information on Facebook is consumed in essentially the same way. The same pattern holds if we look at the liking and commenting activity of polarized users (Figure 3). The bottom right panel of Figure 3 shows the few posts - 7,751 (1,991 from science news and 5,760 from conspiracy news) – that were commented on by polarized users of the two communities.

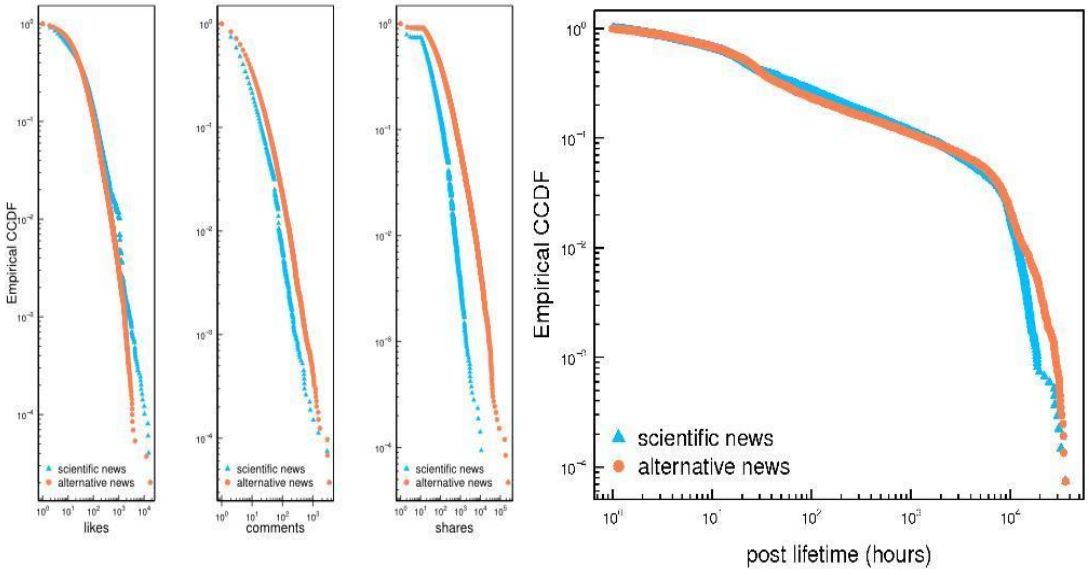
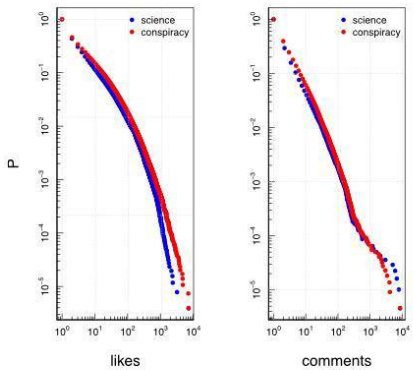
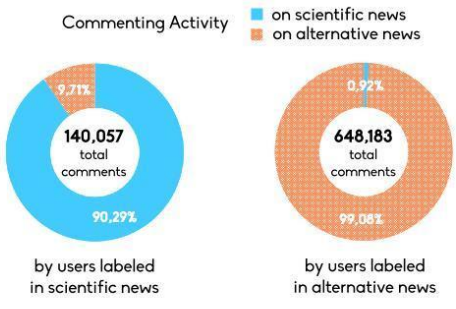


Figure 2: Scientific and conspiracy narratives experience of similar user interactions regarding the statistics of likes, comments, shares and post lifetimes.

	Users Classified	% Users Classified	Comm on their category	Comments on the opposite category
Science	225,225	76,79	126,454	13,603
Conspiracy	790,899	91.53	642,229	5,954



Consumption patterns for polarized users. Cumulative Distribution Function for likes (panel a) and comments (panel b) of polarized users. Also the analysis on polarized users show some similarities in consumption patterns.



Commenting activity of polarized users. **7551** posts (**1991** in Science and **5790** in Conspiracy) polarized users on the opposite category in terms of comments.

Figure 3: The user polarization of scientific and conspiracy narratives shows statistically similar interactions respect to the number of "likes" and "comments" to a post. The number of users debating with the other community is a very small fraction of the polarized users.

Figure 4 shows that the more active a polarized user is on a specific content, the higher the number of friends who display the same behavior For each polarized user, we consider the fraction of y friends who share the same polarization and compare it with that user's engagement θ (number of likes) on the specific narrative. Social interaction is "homophily driven" – i.e., users with similar polarization tend to aggregate together. It follows that the two groups of polarized users (science and conspiracy) share not only similar information consumption patterns but also a similar social network structure.

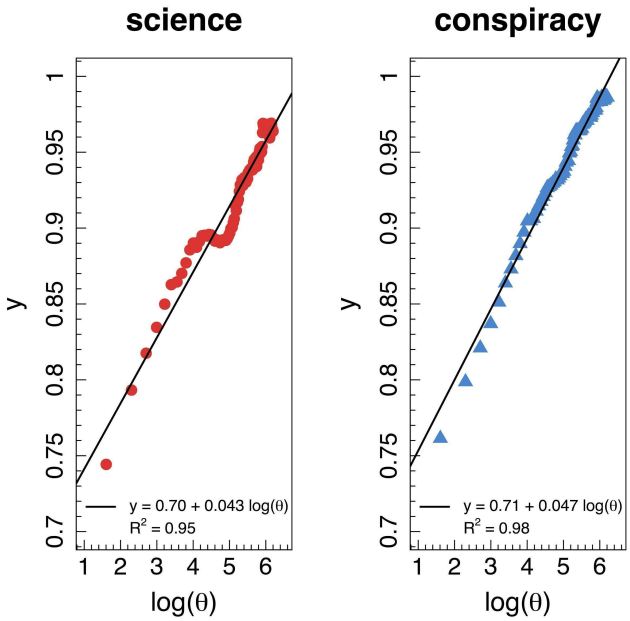


Figure 4: Homophily and activism: the more a polarized user is active (larger θ), the more the user has friends with similar profiles (larger y).

To check whether the observed effects might be limited to the Italian Facebook, we perform a similar analysis on the conspiracy and science page of Facebook US. As expected, we find the same patterns. Contents of the two narratives aggregate users into different communities. The consumption patterns of users' communities are very similar. Figure 5 provides a summary of the polarization and consumption patterns in terms of likes, comments, shares. and lifetime for US Facebook users.

In the top panel, we can observe the users' polarization histograms sorted on the basis of both likes (on the left) and comments (on the right). As in the Italian case (Figure 1), polarization is sharply bimodal, with most of the users concentrated around the extreme values $\rho(u) = -1, 1$ (respectively science and conspiracy). The bottom left panels of Figure 5 show that the CCDFs of the number of likes, comments, and shares are heavy

tailed and similar for both groups, thus indicating similar activity and consumption patterns for both types of users. Finally, in the bottom right panel of we plot the lifetime of posts belonging to conspiracy and scientific news, and even here they are hardly distinguishable.

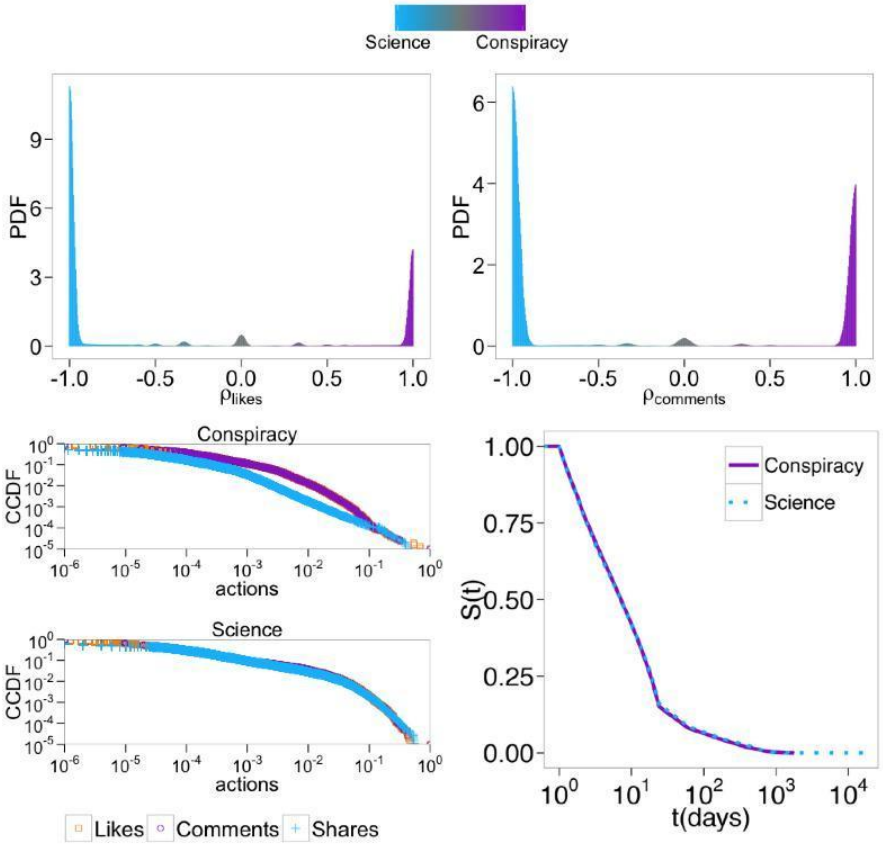


Figure 5: Analysis of US Facebook: as in the Italian case, contents related to distinct narratives aggregate users into different communities and users' attention patterns are similar in both communities in terms of interaction and attention to posts. Top panels: histograms of users' polarization calculated compared to the number of likes (left) and the number of comments (right). Left bottom panels: the statistics of likes and comments are similar both for science and for conspiracy users. Right bottom panel: the Kaplan-Meier estimates of survival functions of posts in science and conspiracy (measuring the fraction of posts which are still active after a given time from their publication) are hardly distinguishable.

Information Spreading and Emotions

Cascades

Thus far, we have considered users' interaction with information. We now focus on the spread of information among users. We show how homophily produces informational cascades and how these are mostly confined inside the echo chambers. We start the analysis by looking at the statistical signatures of cascades related to science and conspiracy news.

Measuring the distance in time between the first and last user sharing a post can approximate the lifetime of cascade effects. In Figure 6 we show the PDF of the cascade lifetime (using hours as time units) for science and conspiracy. In both categories we find a first peak at 1–2 hours and a second one at 20 hours. Temporal patterns are similar. We also find that a significant portion of the information diffuses rapidly (24% for science news and 21% for conspiracy rumors diffused in less than 2 hours, and 39% of science news and 41% of conspiracy theories in less than 5 hours). Only 27% of the diffusion of science news and 18% of conspiracy lasts more than one day.

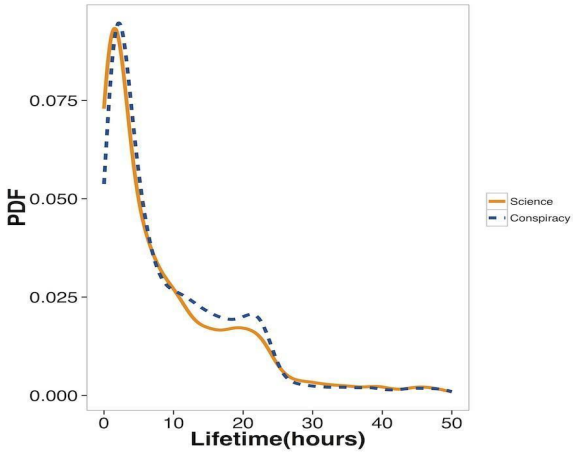


Figure 6: cascade lifetimes for science and conspiracy are very similar.

In Figure 7, we show that the majority of shares pass from users with similar polarization, i.e. users belonging to the same echo chamber. In particular, the average edge homogeneity (measuring the users' similarity) of all cascades shows that it is highly unlikely that a path might include users from different groups. Contents tend to be confined only within echo chambers, and the cascade size is well approximated by the dimension of the echo chamber.

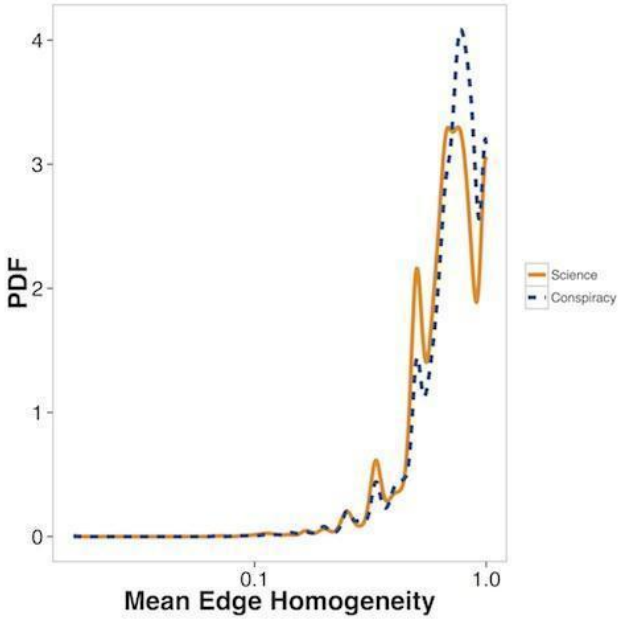


Figure 7: Confinement of cascades within echo chambers: a positive edge homogeneity indicates that information propagates among users with similar beliefs. We do not observe cascades with a negative mean edge homogeneity and that the values are most likely to be concentrated around the maximum edge homogeneity value of ~ 1, indicating a confinement of the cascades within echo chambers.

In Figure 8, we show the lifetime of a cascade as a function of the cascade size, i.e. the number of users sharing a post. Thus far we have seen similar signatures for both the science and the conspiracy echo chambers, but we now observe, for the first time, some differences between the two. In short: For conspiracy-related content, the lifetime of a post shows a monotonic growth respect to the cascade size, but for science news, we observe instead a peak in the lifetime corresponding to a cascade size of $\approx 100 \div 200$ users. News is assimilated very differently. Science news reaches a higher level of diffusion more quickly, and a longer lifetime does not correspond to a higher level of interest but most likely to a prolonged discussion within a specialized group of experts.

By contrast, conspiracy rumors diffuse more slowly and show a positive relation between lifetime and size. Long-lived posts tend to be discussed by a higher number of users.

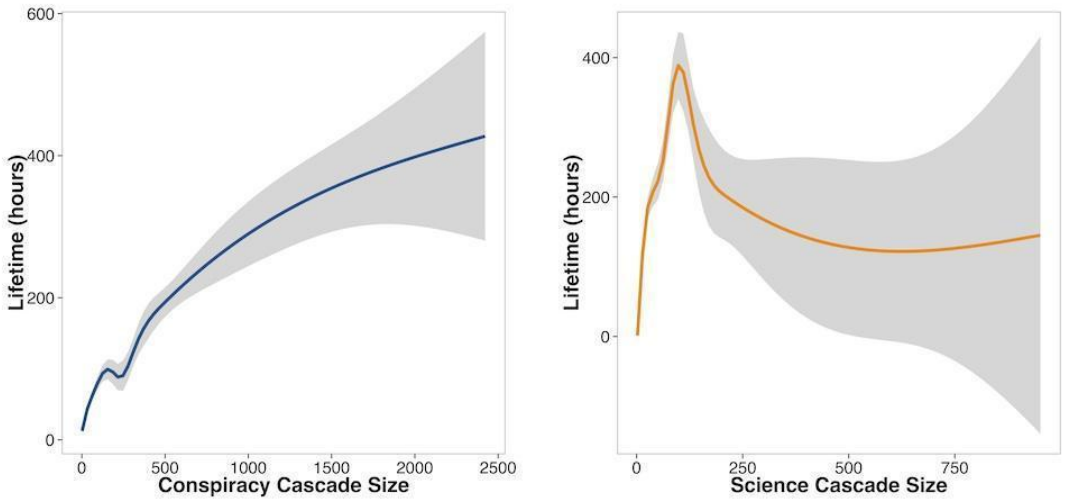


Figure 8: Lifetime of a post in comparison to cascading size (number of users discussing the post). The grey area represents the variability of the lifetime for a given size.

Extremity and Emotions

Consistent with a large body of research, it might be hypothesized that users, discussing issues online, will become increasingly extreme in their beliefs after those discussions. As a result, their views will be reinforced and polarized. We provide some support for this hypothesis by showing the results of the sentiment analysis[10] applied to user discussion (comments) within the Italian Facebook echo chambers.

On Facebook, comments are, of course, the medium of online debate by which users express their views about the post or about the discussion itself. Sentiment analysis is a computational tool to approximate the emotional attitude of users' comments. It is based on a supervised machine learning approach, where we first manually annotate a substantial sample of comments, and then build a classification model. The model associates each comment with one sentiment value: negative, neutral, or positive. The value expresses the emotional attitude of Facebook users when posting comments.

We find that for both kinds of content, the longer the discussion, the more negative is the sentiment. Comments on conspiracy posts tend to be more negative than those on science posts. Moreover, the higher the engagement of a user in the echo chamber, the stronger the probability of expressing negative emotion when commenting -- both on science and conspiracy. In Figure 9 we show the average sentiment of polarized

[10] Sentiment analysis has been performed with supervised training of Support Vector Machines

users as a function of their number of comments. The more active a polarized user is, the more the user tends towards negative values both on science and conspiracy posts.

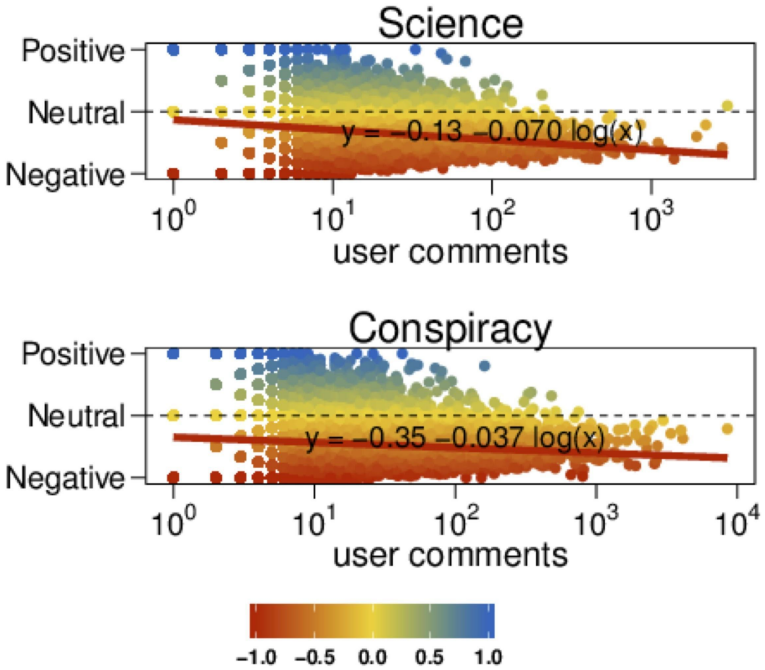


Figure 9: Sentiment (positive, negative, neutral) of polarized users comments as a function of their engagement.

Response to confirmatory and debunking opinions

We have seen that users tend to aggregate around preferred contents and form polarized groups. We now attempt to sharpen this claim by testing how users respond to information that either confirms or debunks their beliefs.

We first test the attitudes of conspiracy and science users in interacting with false information. To avoid biases in the determination of the truth of a post, we have used information that is false *intentionally*. For the Italian data set, we collected a set of Troll posts that were satirical imitations of conspiracy information sources. All these posts contain clearly unrealistic and satirical claims. Examples include posts declaring that a new lamp made of actinides (e.g. plutonium and uranium) might solve problems of energy gathering with a lower impact on the environment, or that chemical analysis reveals that chem-trails contains sildenafil citratum (the active ingredient of ViagraTM). We find that polarized users on conspiracy pages are consistently more active in liking and commenting on intentionally false claims (80% of the pool) when compared to science users (Fig 10).

Users usually exposed to conspiracy claims are more likely to jump the credulity barrier. Even when information is deliberately false and framed with a satirical purpose, its conformity with the conspiracy narrative transforms it into suitable (and welcome) content for the conspiracy echo chamber. Confirmation bias evidently plays a pivotal role in the selection of content.

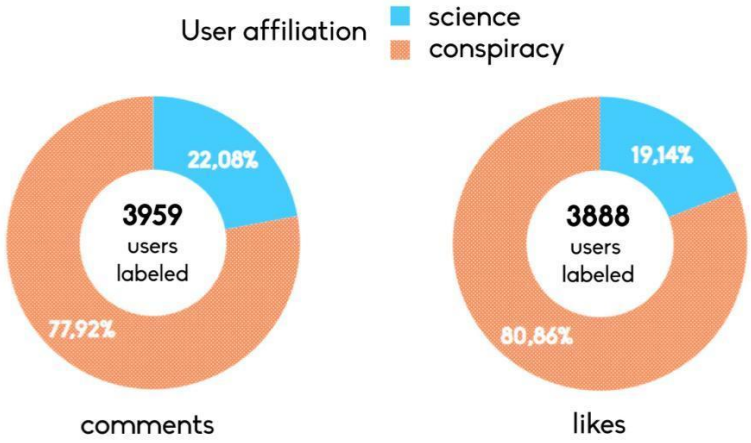


Figure 10: interaction of polarized users with deliberately false information (Trolls' posts). We observe that both in terms of comments and likes conspiracy users represent approximately 80% of the pool.

Debunking information aims to correct falsehoods. For the U.S. data set, we use debunking posts to test their efficacy and, more generally, to characterize the effect of such information on conspiracy users. The first interesting result is that out of 9,790,906 polarized conspiracy users, just 117,736 interacted with debunking posts – that is, commented on a debunking post at least once. Among these conspiracy users, those with persistence in the conspiracy echo chamber greater than one day were only 5,831 -- 5,403 who registered likes, and only 2,851 offered comments. (The latter numbers exceed 5,831 because some people did both.) Hence, the impact of debunking appears marginal; fewer than 1.3% of conspiracy users interact with it.

A good approximation of a user's commitment to the echo chamber is given by his or her daily number of likes and comments (i.e., liking and commenting rates). We use this measure to compare the activity of the user before and after the interaction with a debunking post (dissenting information if the user belongs to the conspiracy echo chamber). Fig. 11 shows that the exposure to debunking actually induces a shift towards higher liking and commenting rates.

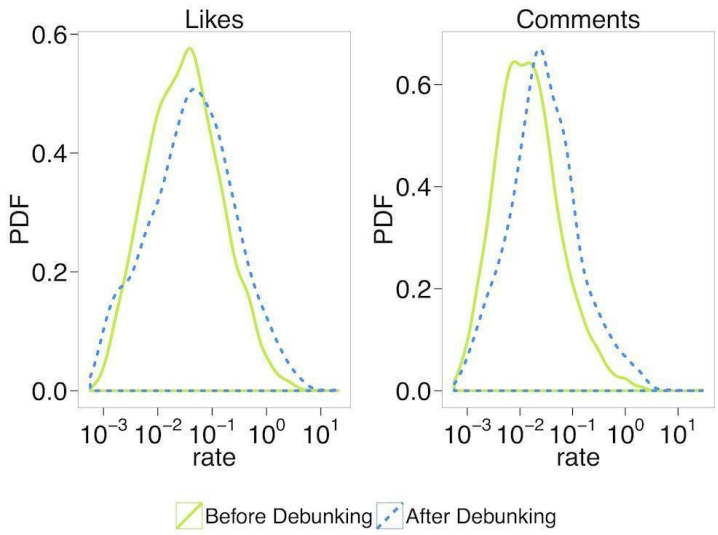


Figure 11: Reinforcement of conspiracy beliefs after exposure to debunking. We have observed that the small fraction of conspiracy users that interact with debunking posts tend to increase their activity within the conspiracy echo chamber

Debunking is ignored (by ~99.98% of conspiracy users) or produces the unwanted effect of reinforcing the very beliefs that it was supposed to correct. It follows that in our data set, attempts to convince conspiracists that their beliefs are false generally seem to fail.

Conclusions

At least in the areas studied here, Facebook users are highly polarized. Their polarization creates largely closed, mostly non-interacting communities centered on different narratives - i.e. echo chambers. The echo chambers are statistically similar in terms of how communities interact with posts. For both scientific information and conspiracy theories, the more active a user is within an echo chamber, the more that user will interact with others with similar beliefs. The spreading of information tends to be confined to communities of like-minded people. We have also found at least indirect evidence of group polarization within those communities.

In the discussions here, users show a tendency to seek out and receive information that strengthens their preferred narrative (see the reaction to trolling posts in conspiracy echo chambers) and to reject information that undermines it (see the failure of debunking) The absorption of trolls' intentionally false conspiracy theories into echo chambers shows how confirmation bias operates to create a kind of cognitive inoculation.

We emphasize that our data sets are limited, of course, to particular data sets, and so we cannot venture any general claims about echo chambers on social media. Contexts differ, and far more research would be necessary to support any such general claims. But

at least in the domains studied here, people are using Facebook to create enclaves of like-minded people, spreading information in strikingly similar ways.

Acknowledgments

For the development of this work we thanks Alessandro Bessi, Michela Del Vicario, Fabiana Zollo, Delia Mocanu, Qian Zhang, Laura Tommaso, Guido Caldarelli and Gene Stanley. A special thanks to Giulia Borrione for her valuable thoughts and insights that greatly contributed to the development of this work.

Furthermore, we want to thank for valuable suggestions and discussion Geoff Hall, "Protesi di Protesi di Complotto," "Che vuol dire reale," "La menzogna diventa verita e passa alla storia," "Simply Humans," "Semplicemente me," Salvatore Previti, Elio Gabalo, Sandro Forgione, Francesco Pertini, and Franco Lechner. Thanks to the Program on Behavioral Economics and Public Policy at Harvard Law School.

MULTIPLEX (Foundational Research on MULTIlevel comPLEX networks and system) grant number 317532, SIMPOL (Financial Systems SIMulation and POLicy Modelling) grant number 610704, DOLFINS (Distributed Global Financial Systems for Society) grant number 640772, and SoBigData (Social Mining & Big Data Ecosystem) grant number 654024.