methaodos.revista de ciencias sociales (2023) 11(1) m231101a01

How I learned to hate you.
Parasocial interactions in echo chambers and their spillover effects

Cómo aprendí a odiarte.
Las interacciones parasociales en las cámaras de eco y sus efectos indirectos

Raluca-Nicoleta Radu | raluca.radu@fjsc.ro
University of Bucharest, Romania
Center for Advanced Internet Studies, Germany

10.17502/mrcs.v11i1.598

Received: 13-09-2022
Accepted: 11-10-2022

Abstract

This article analyzes the most influential posts on Facebook related to COVID-19, for the first two years of the pandemic, to explain how parasocial opinion leaders created echo chambers, in the Romanian public sphere, and to discuss the cumulative spillover effects these echo chambers had on society at large. A database of the 233,242 most influential posts in Romanian about COVID-19, from the first two years of the pandemic, is investigated using a mixed methods approach, to 1) verify statistically if issue-related echo chambers existed and 2) to describe, qualitatively, how they functioned. A special focus is devoted to trolling in the form of reactions to posts, such as haha reactions for messages about COVID-related deaths. Using the literature on parasocial interaction, inoculation theory, online disinhibition effect and echo chambers, the article shows how echo chambers supported trolling behavior, for radicalized Facebook users, how they polluted the public discussion and how they made dialog impossible for social groups that ended up identifying each other as the enemy. Based on these research results, the author proposes two policy recommendations for social platforms.

Keywords: COVID-19, Facebook, inoculation theory, opinion leaders, spillover effects.

Resumen

Este artículo analiza las publicaciones más influyentes en Facebook relacionadas con COVID-19, durante los primeros dos años de la pandemia, con el objeto de explicar cómo los líderes de opinión parasocial crearon cámaras de eco, en la esfera pública rumana, y así discutir los efectos secundarios acumulativos que estas cámaras tenían sobre la sociedad en general. Se investiga una base de datos de las 233.242 publicaciones más influyentes en rumano sobre COVID-19 de los dos primeros años de la pandemia, utilizando un enfoque de métodos mixtos, para 1) verificar estadísticamente si existían cámaras de eco relacionadas con el problema y 2) describir, cualitativamente, cómo funcionaban. Se dedica un enfoque especial al troleo en forma de reacciones a las publicaciones, como reacciones jaja para mensajes sobre muertes relacionadas con COVID. Utilizando la literatura sobre la interacción parasocial, la teoría de la inoculación, el efecto de desinhibición en línea y las cámaras de eco, el artículo muestra cómo las cámaras de eco apoyaron el comportamiento de troleo, para los usuarios de Facebook radicalizados, cómo contaminaron la discusión pública y cómo imposibilitaron el diálogo para los grupos sociales que terminaron identificándose unos a otros como el enemigo. A partir de los resultados de la investigación, el artículo presenta dos recomendaciones normativas para las plataformas sociales.

Palabras clave: COVID-19, Facebook, teoría de la inoculación, líderes de opinión, efectos indirectos.

Summary

How to cite this work

Radu, R.-N. (2023). How I learned to hate you. Parasocial interactions in echo chambers and their spillover effects. *methaodos.revista de ciencias sociales*, *11*(1), m231101a01. http://dx.doi.org/10.17502/mrcs.v11i1.598

1. Introduction

Data from the Digital News Report - DNR 2021 (Newman *et al.*, 2021), on trust in news from social media, correlated with data from the 7th wave of World Values Survey - WVS 2020 (EVS/WVS, 2022), on trust in peers, indicate a paradox: the less one person trusts others, the more he or she trusts social media. There is a negative correlation (R=-0.64) for trust data in social media and in peers, for the 41 markets present in both surveys. The DNR data were collected in January-February 2021 on a sample that represents only digital publics and are negatively influenced by the pandemic evolutions and the worldwide discussions on online and offline disinformation (Newman *et al.*, 2021). The WVS data were collected over a longer time span, 2017-2022, and represent the adult population in different markets. Nevertheless, the negative correlation is quite large, to make us wonder how the paradox can be explained. In the case of Romania, trust in news on social media is above the 41 states average (28.7pp versus 24pp), while interpersonal trust is below average (12.1 percent compared to 29.3 percent). Trust in institutions is also low, over several World Values Survey waves (Tufiș, 2000).

This apparent paradox can be explained with the help of the echo chamber hypothesis: we join discussions with people that share similar perceptions of reality, because we feel more comfortable when we have our views supported by peers (Jamieson & Cappella, 2008). We end up trusting people online, because they think and behave like us. In communities with lower trust in peers and in institutions, echo chambers can have damaging effects on the way public interest issues are interpreted and discussed. What happens when we are confronted with a reality for which we need to negotiate new social representations, such as the onset of COVID-19 pandemic?

This paper follows the evolution of the discussions on COVID-19 on Facebook, the most influential social network in Romania, for the digital public, according to DNR data 2017-2022[1]. The first 10,000 posts related to COVID, ordered by number of interactions, for the period February 2020 - January 2022, were downloaded from Facebook, with the help of CrowdTangle (CrowdTangle Team, 2022). These are posts CrowdTangle, a social media tracking tool owned by Meta, labels as COVID and identifies the language they are written in as Romanian. A first analysis looks at the evolution of the interest for COVID-related posts, expressed through numbers of comments, reactions, and shares, for the whole database and for individual Facebook IDs that supported or contested the generally accepted prevention measurements. A second analysis looks at the first three posts from each month of the study.

Romania followed the general European Union measurements against COVID-19 and the Romanian governments supported generally-accepted prevention and treatment lines, such as mask wearing, social distancing and vaccination. Still, at the end of the summer of 2021, only one in four Romanians was vaccinated (CNCAV, 2021). Romania and Bulgaria had the lowest rates of vaccination against COVID-19 in the European Union. In September 2021, while most European countries were preparing to lift restrictions and to return to normal activities, Romania witnessed the deadliest COVID wave during the COVID-19 pandemic period (WHO, 2022). The interest in COVID-related posts, as expressed by total number of interactions and by the different reactions (see Graph 1), does not follow the COVID waves, but precedes them, and, to some extent, explains the appearance of these waves.

Using the literature on parasocial interaction, inoculation theory, online disinhibition effect and echo chambers, this article shows how echo chambers forced individual positioning for or against prevention rules, as an identity trait, and supported trolling, by radicalized Facebook users. The article also shows how discussions on social media polluted the public space and how the radicalization strategies used by parasocial opinion leaders made dialog impossible for social groups that identify each other as the enemy.

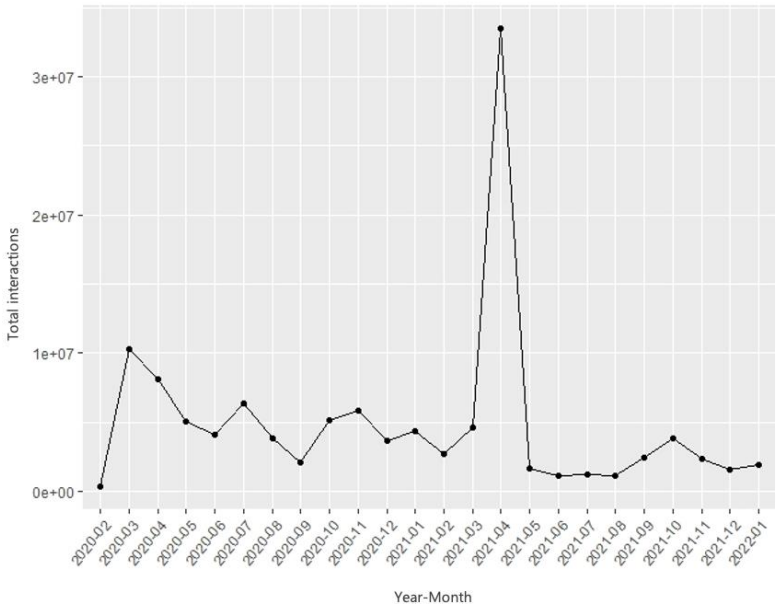
1.1. Echo chambers and social media

Two metaphors were prominently used in the academic and lay discourses about social media: *echo chambers* and *filter bubbles*. In his 2001 book *Republic.com*, Cass R. Sunstein prepared the ground, talking about a near future when people have the 'ability to «filter» what they want to read, see and hear' (Sunstein, 2001, p. 3) both in terms of subjects and in term of world views. Sunstein introduced the notion of 'echo

[1] The Reuters Institute Digital News Report data are available here: www.digitalnewsreport.org.

chambers', to cover the idea of enclaves of like-minded people and the concept of a selection of news on a specific topic, that is consumed exclusively and that obliterates all other topics that might be of interest for a healthy civic life - the 'Daily Me' (Sunstein, 2001, p. 208). Echo chambers, warned Sunstein (2001), affect democracy through fragmentation of media content and media audiences and through group polarization on critical themes in the public sphere. Ten years later, Eli Pariser (2011) proposes the concept of 'filter bubbles', to discuss the effects algorithms have on our perception of reality, by controlling what we see and what we do not see. Algorithms select content, from a list of topics, leaving out all other subjects that might be of interest, at some point in time. Citizens live in a filter bubble, with no information on what is going on, except for the already chosen themes. Pariser was thus splitting the former notion of 'echo chambers' into two distinct concepts: one referring to groups of people gathered around an ideology (the *echo chambers*), and one referring to the selection of information and opinion that narrows our access to relevant knowledge about reality (the *filter bubble*).

Graph 1. Total number of interactions for the first 10,000 most influential COVID-related post in Romanian, on Facebook, over a period of two years. For February 2020, only 3242 posts were identified.



Source. Own elaboration based on Crowdtangle Team, 2022.

Both Sunstein and Pariser wrote compelling accounts of how offline and online media function and how they affect our lives, and various researchers followed suit. In different points in the development of social media, using audience data, researchers tried to validate the existence of echo chambers or filter bubbles, reaching conflicting results. These results were influenced by audience socio-demographics, media systems, types of media consumed – offline or online, and even types of social media, and topic (see, for example, Zollo *et al.*, 2017; Fletcher & Nielsen, 2018; Toff & Kalogeropoulos, 2020; Cinelli *et al.*, 2021). Research using big data analysis of interactions on social platforms supports the echo chambers hypothesis, indicating the existence of 'homophilic clusters' of like-minded peers that share information among themselves (Cinelli *et al.*, 2021, p. 5), but the filter bubble hypothesis is more and more vigorously resisted (Fletcher & Nielsen, 2018; Dahlgren, 2021). Peter Dalhgren (2021) reaches the conclusion that many previous findings on selective exposure to mass media contradict the filter bubble hypothesis, that people have a narrow access to information and do not know what others think and talk about. He says that this hypothesis might be, in fact, based on a technological determinism frame as old as the printing press itself.

C. Thi Nguyen (2020) attempts a clarification of concepts, proposing a discussion on 'echo chambers' and 'epistemic bubbles': `[b]oth are structures of exclusion – but epistemic bubbles exclude through omission, while echo chambers exclude by manipulating trust and credence' (Nguyen, 2020, pp. 141-142). More specifically, epistemic bubbles are based on 'coverage gaps' (Nguyen, 2020, p. 150), that can be addressed by individuals in an active pursuit of information.

The echo chambers are more difficult to escape, explains Nguyen, as they are 'contingent social structures' that embed the individual (Nguyen, 2020, p. 152), governing the manner the world is defined by the inner group through 'epistemic control' (Nguyen, 2020, p. 149). The methods used inside an echo chamber are similar to 'cult indoctrination' (Nguyen, 2020, p. 147), and define group identity by proactively discrediting alternative sources of information, world views and counterarguments.

Nguyen does not distinguish between echo chambers created in actual life, inside groups of peers that meet and discuss face to face, and echo chambers fashioned by media personalities and content creators. In communication studies, the debate around the theories of one step flow, two-step flow and even multistep flow of media influence on the audience is a classical one (Weimann, 2017). Sociological audience research after the Second World War showed that the powerful effects theory of media, acting as a magic bullet in the heads of a passive audience – the one step flow - is not supported by data collected in periods that are not marked by crises. People tend to get information to support their decision-making processes, such as voting for a particular party, from opinion leaders that are, in fact, their peers with which they have a real interpersonal relationship – in a two-step flow of information (Katz, 1957). These opinion leaders are regular members of a given community, more active in searching information on a certain domain and thus, more knowledgeable. They act as intermediaries between mass media and their group of peers, selecting and interpreting information. While performing their intermediary role, they have to preserve the norms and the values of their community, in order to maintain credibility (Katz, 1957; Rogers, 2003). The importance of the interpersonal influence in access to information and decision-making processes is the basis of the two-step flow of information theory (Katz, 1957).

Yet, celebrity studies showed that for certain publics, under certain conditions, connections similar to the interpersonal ones are built with media people, such as anchors, athletes, singers or actors, and even with fictional characters, from movies or cartoons. These parasocial interactions are based on social attractiveness and salient characteristics – the media character is perceived as a friend or acts as a social reference for certain audience members and may influence attitudes and behaviors on a limited period (Giles, 2002). Further research results on connectivity and engagement, with journalists, on mainstream media, and celebrities, on social media, explained how an `illusion of intimacy` (Schickel, 1985, apud Marwick & boyd, 2011) is strategically created through an 'informal «fireside» approach of the listener' (see Giles, 2002, p. 290) and through an elaborate creation of the self, for the public eyes (Marshall, 2010). Thus, 'private experiences' are recreated for `the public consumption` (Marshall, 2010, p. 37), on a three layers deep construction: the 'public self', that is, the official persona, linked to public activities like sport or music, the 'public private self', based on a strategic exposure of the individual's less public life, and the 'transgressive intimate self', represented, for example, by the public manifestation of a temporary, strong emotion, that would normally be shared just with a limited number of family and friends (Marshall, 2010, pp. 44-45).

Using a conversational tone and an elaborate character that combines public and private elements, celebrities manage to communicate with their audiences better than established newsrooms and can influence some people's actions, either immediately, making them part of a campaign, or on a longer time span, by influencing, for example, arguments audience members would use in a conversation, later (Marshall, 2010). For these media based social influencer, Stehr and her colleagues (2015) propose the label of 'parasocial opinion leaders'.

1.2. Strategic discourses in echo chambers

Parasocial relationships are created among politicians, TV and radio hosts or social media personalities and members from their fan base. In echo chambers, in particular, the actual techniques used by public figures to influence audience members go beyond a layered construction of the public self and a direct address and have farther reaching effects. Jamieson and Cappella (2008) documented, in conservative media of the 2000s, outraged discourses based on ideologically selected facts, that reinforced inner group identity, ridiculed the

outer group representatives and framed them as the enemy. Very strong negative emotions were associated with these strategically drafted discourses. In different media, representatives of the same echo chambers legitimized each other in a circular manner and delegitimized the ideologically incompatible public figures, by drafting counter arguments to the outer group positions. Jamieson and Cappella (2008) also documented the effects of this type of discourse strategies on the audiences of conservative media, and the effects included confirmation of existing opinions, for conservative publics, and increased polarization across different media audiences. The public of the conservative radio host Rush Limbaugh was 'engaged, outraged and with strong opinions [...] the byproduct of Limbaugh's success in creating a virtual community of like interests and like minds' (Jamieson and Cappella, 2008, p. 139). Their analysis refers to the inoculation theory (see Jamieson and Cappella, 2008, p. 142), without referencing any of the psychological studies on the matter.

The inoculation theory is the base of a persuasion strategy proposed by the social psychologist William J. McGuire in the 1960s and verified by research experiments afterwards (Compton & Pfau, 2005). As the name indicates, the inoculation theory resembles immunization through vaccination: communicators may build attitude resistance to future persuasion in advance, just like vaccines build individual resistance to future viral illnesses. Individuals are presented a threat, the catalyst to resistance, possible arguments that may be used by the opposing parties and may result in attitudinal change, and specific counterarguments, to help them built a defense. This persuasive strategy is based on cumulative effects: it needs time to become effective and has fading results, if not used consistently. Recently, the inoculation theory was identified as a pre-debunking strategy in critical contexts, by Lewandowsky and his colleagues (2020).

Jamieson and Cappella also noted an 'emotionally charged vocabulary' (2008, p. 179) inside the echo chambers. Strong emotions are a trait of the transgressive intimate self. Media show hosts and their guests feel so angry, that they publicly expose their most extreme feelings. Rage, grief or incivility are usually kept to private space (Radu, 2015) and are in general not accepted in public discourse, because they influence rational decision-making and can have spillover effects, such as violence and hatred against individuals or groups (Sorial, 2017).

In summary, there are several dangerous effects of echo chambers, identified by the authors quoted above: fragmentation, cult-like indoctrination, polarization, hatred, and violence (Sunstein 2001; Jamieson & Cappella, 2008; Sorial, 2017; Nguyen, 2020).

In avoidance of fragmentation, the critical issue of access to a similar set of facts about reality is achieved under special conditions, Katz (1996) sustains: either a unique, trusted source of information, such as the single channel of the public service broadcasting (the case of Israel, for two decades), or an occasional media event such as televised debates during elections, ceremonies with a global appeal (the Moon landing) or catastrophic occurrences (the Persian Gulf war). The case of COVID-19 is that of a catastrophic occurrence, where access to critical information is shared by the large majority of individuals.

Several researchers maintain that engagement with media content is temporary, and it is perceived by audience members as having no consequences. On social media, the actual contact with content may be of several seconds, as the feed is composed of tragic news and outraged commentaries, mixed with funny pet posts and advertising. Content is shared with millions of users, but the consumption is private and public reactions to it, no matter how outrageous, feel also private. Perceived anonymity, invisibility, and lack of eye-contact lead to toxic online disinhibition effect (Suler, 2004): facing no immediate consequences, some people are inclined to become rude and to show anger, behaving violently and even threatening others. Niche echo chambers feed communities of trolls, such as groups of individuals who react with *LOL (laughing out loud)* at online grief (Phillips, 2011). While some researchers believe trolling actions are harmless or represent negative feedback for public communication transgression (Phillips, 2011), trolling behavior has actual consequences on targeted individuals, be they athletes (Kavanagh, Jones & Sheppard-Marks, 2016) or journalists (Waisbord, 2020).

The threat is real as result of what Giles (2002, p. 295) calls 'first order parasocial interaction': media personalities and their audiences share a parasocial relationship, with chances of actual contact. The consequences at society level are also real, especially for political echo chambers (Waisbord, 2020). Strategic discourses in echo chambers have spillover effects. Media spillovers or externalities are positive or negative consequences of communication over people, groups or the community at large, that were not part of the initial communication and were not in direct contact with the media content (see Radu, 2019). Trolling is a spillover effect.

Related to COVID discussions online, Velásquez and his colleagues (2021) documented hate clusters, on individual social platforms, and hate communities, across platforms, that spread and amplify violent narratives. On their turn, Xu and his colleagues (2021) showed that hate speech on COVID themes has real consequences, that range from stigma to hate crimes. Stigmatization and hate crimes against a certain race or a certain nationality, during the sanitary crisis, are also spillover effects.

The COVID-19 pandemic was a catastrophic occurrence that, at first, diminished social fragmentation: in the very first months of the generalized lockdown, in 2020, people all over the world sought similar information on the illness, on prevention measures and on fatalities. In newsrooms, all journalists, from sports to foreign affairs, were working on the same COVID-19 beat. Two years later, some people, convinced COVID-19 never existed, laugh at news about the evolution of cases and fatalities. Others, afraid to have one of the many vaccines available, since the beginning of 2021, die because of a now preventable disease (WHO, 2022). In order to understand why so many people do not acknowledge the existence of the disease and decide not to follow generally accepted prevention rules, this article explores the issue on echo chambers related to COVID-19, in the Romanian public space. More specifically, the research aims: 1) To verify the existence of COVID-19 echo chambers in Romanian language, on Facebook. If the existence of echo chambers is validated, the next steps are: 2) To see how these echo chambers function; 3) To discuss their possible effects of COVID-related echo chambers on individuals and on society at large.

2. Method

Facebook was the most popular social platform for general use and for news consumption, for the Romanian digital public, before and during the pandemic (Newman *et al.* 2021; Newman *et al.* 2022). It is a social platform where individuals and companies alike post content and interact with content by commenting, sharing and reacting. Accounts can be aggregated into groups, that are either open, so that anyone can see group discussions, or closed. Also, posts can be shared with a global audience or can be restricted to direct contacts, called friends. On a public post, shares, reactions and comments are also public. The editor of a post can moderate comments, to eliminate or hide the ones considered inappropriate. Inside groups, accounts can be banned. The other types of interactions – shares and reactions – cannot be moderated. Since 2016, the range of reactions was diversified by Facebook, to include five emoji icons for 'love, laughter, surprise, sadness or anger', besides 'the like button' (Mark Zuckerberg, 2016).

Meta, the owner of Facebook, does not allow automatic harvesting of information on posts and users, but grants access to a selection of data for researchers and journalists, through CrowdTangle. The available data include the content of posts and the number of reactions, shares and comments from 'all public Facebook pages with more than 50K likes (automated via API), all public Facebook groups with 95k+ members, all US-based public groups with 2k+ members, and all verified profiles'[2]. For research purposes, databases of the most influential posts, ordered by total interactions, can be downloaded and analyzed.
The first 10,000 posts related to COVID, ordered by interactions, on each month, for the period February 2020 - January 2022, were downloaded from Facebook, with the help of CrowdTangle (CrowdTangle Team, 2022). The total database includes 233,242 posts in Romanian with information on the account, the date of the post, the type of content, the actual texts associated with the post, and the number of different types of interaction: reactions, comments, and shares.

In order to verify the existence of the echo chambers, with opposing positions regarding the prevention for COVID-19, eight Facebook accounts presenting and/ or supporting accepted prevention methods and eight Facebook accounts opposing these methods and/ or presenting an alternative were identified. The sample has 7620 post with 13,910,228 total interactions (3.27 pp from the total number of posts received 11.86 pp of the total number of interactions).

To understand how Facebook users interacted with the selected Facebook IDs, the means of total interactions, likes, laughter and angry reactions for each account, each month, was analyzed, by comparing the data for the two groups. Several types of reactions were analyzed to see if posts receive supportive reactions or trolling reactions. A clear-cut analysis is, nevertheless, difficult if reactions are not understood in the context of a post. For example, a laugher reaction (*haha*) for a post that factually refers to an increase in

[2] https://help.crowdtangle.com/en/articles/4201940-about-us (accessed Aug. 30, 2022).

the number of casualties is a trolling reaction, but the same reaction for a post that ridicules the opposing party is a supportive one. The monthly means for the different reactions, for each of the 16 Facebook IDs, do not have a normal distribution and there are different outliers, from one month to another. In order to keep all the data for the analysis (as I use a selection for the two positions regarding COVID prevention, not all Facebook IDs) a logarithmic transformation was performed, to normalize the distribution and to prepare the data for an ANOVA analysis. The software used for data visualization and analysis is: R version 4.1.3 (2022-03-10) - "One Push-Up" (R Core Team, 2022), tidyverse 1.3.1 (Wickham *et al.*, 2019), psych_2.1.6 (Revelle, 2021) and writexl_1.4.0 (Ooms, 2021).

Besides the statistical analysis, a qualitative exploration was performed on the first three most influential individual posts, each month, for 72 posts. The coding is based on the literature about inoculation theory, echo chambers and celebrity studies. For the inoculation theory, these categories were used: identifying an enemy (1/0), identifying a threat related to COVID prevention (1/0), rebutting an opposing frame, related to COVID prevention (1/0). For example, a post with a fragment from a TV program identifies the authorities as the enemy with: 'Do we still have lives if we are terrorized [to respect COVID prevention rules]?', identifies a threat with 'the panic was induced by media [...] they try to destroy in us everything that is human' and rebuts the framing of COVID as a deadly virus with 'people do not die OF covid, they die WITH covid'.

For the echo chambers, these categories were used: validation of inner group members (1/0), emotionally charged vocabulary (1/0), ridicule (1/0), strong negative emotions (1/0). A post based on an interview, that has both the interviewer and the interviewee agreeing on all issues and supporting each other's position, just like posts that promote an individual with a strong positioning regarding COVID prevention, were labeled as validating inner group members. A post in which a member of the European Parliament compares Europe to a concentration camp was labeled as using an inflammatory choice of words and creating strong negative emotions. In addition, a post that says that one does not have to believe in the vaccine, because there is nothing mystical about medicine, was labeled as using ridicule.

For the three layers of the public image of parasocial opinion leaders, three labels related to content were used: *public self*, for posts about public, official activities, *public private self*, for posts on private life, such as visits to the hospital, to get vaccinated, and *transgressive intimate self*, for posts about strong feelings and intimate thoughts that are usually expressed in a very close circle of persons. For example, the post about the vaccination of the Romanian president, with pictures of the medical procedure, was labeled as 'public private self', and the post in which an artist is bitterly complaining about trolls or when a member of the Parliament shouts at the health minister, in the street, as 'transgressive intimate self'.

Additionally, the posts were color coded to identify support (blue) or lack of support (red) for accepted prevention methods *at the time*, since the knowledge of the disease and of effective prevention methods evolved in the period 2020-2022.

3. Results

The statistical analysis of reaction data for content posted by 16 influent accounts, selected from the most influential 10,000 monthly posts, identified with the help of CrowdTangle, indicates that on Facebook, in the period February 2020 - January 2022, there were at least two echo chambers in Romanian, discussing prevention methods for COVID-19. These two echo chambers developed independently and eventually became mutual enemies, attacking each other.

The ANOVA analysis of the log of monthly means per selected Facebook ID shows that Facebook users interacted differently with the two groups of Facebook selected IDs, based on their positions of support or opposition of accepted prevention methods at the time. Both for data on total interactions (including comments, reactions and shares) and on *like* reactions, the difference between the two groups, one supporting and one contesting prevention methods, is statistically significant (Table 1). A further statistical analysis of laughter and anger reactions, that can be linked to trolling behavior, shows that for *haha* reactions, the groups were not statistically different based on their position regarding COVID prevention, but at the monthly level, the data vary significantly. For the *angry* reactions, data are fluctuating both by position and by month, in a statistically significant way.

Table 1. P values - ANOVA analysis on type of reaction and selected group (by month and by position).

Log of monthly means per selected Facebook IDs	Position (supporting/ protesting prevention)	Month (Febr. 2020-Jan. 2022)
Total interactions	2.814e-09 ***	0.2225
Likes	2.374e-10 ***	0.2952
Haha	0.460691	0.009366 **
Angry	0.0001939 ***	0.0016707 **

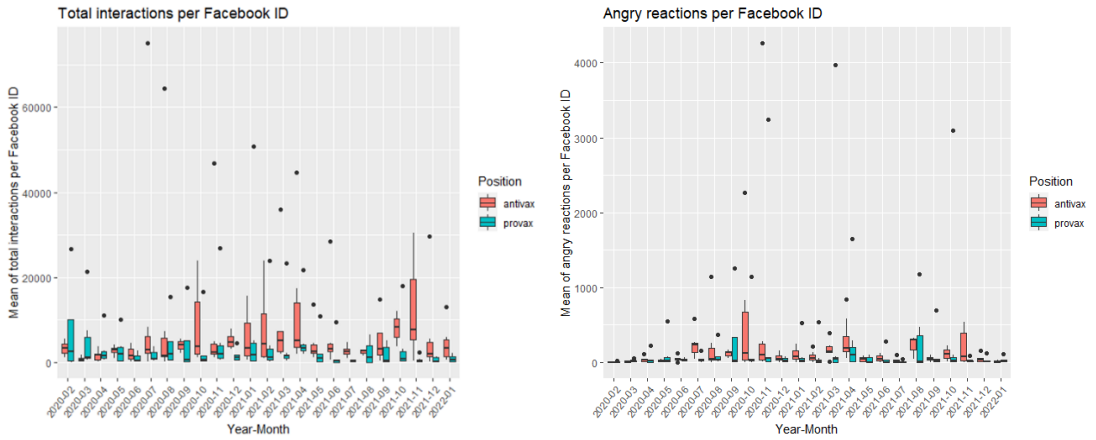
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Source. Own elaboration based on CrowdTangle Team, 2022. For full results, see *Data availability*.

The boxplot representations for monthly total interactions, grouped by position (Graph 2a) show that, in the first half of 2020, the two groups attracted a comparable number of reactions, with a slight support for posts about accepted methods of prevention, in the very first two months of the lockdown. Starting with July 2020, the group contesting accepted prevention methods begins to receive more attention (see *Data availability* for a visualization of selected accounts against all Facebook IDs, in different months). If in the first half of the year, the outliers are pro-prevention accounts (two medical doctors and the Romanian president), in July the outlier is a politician campaigning online against COVID prevention, in preparation for the Parliamentary elections of December 2020 (George Simion, the leader of a new right-wing party).

When it comes to angry reactions, the group against prevention methods begins to differentiate itself also from July 2020 on (Graph 2b). Comments to anti-prevention accounts show that angry reactions were coming both from supporters, that duplicate the arguments of the posting, and from trolls, that contest the message.

Graph 2a & 2b. Comparation of monthly means for interactions to posts, Febr. 2020-Jan. 2022, for the two groups of Facebook IDs



Red/ antivax: accounts constantly posting against accepted prevention methods at the time and blue/ provax: accounts constantly posting to support accepted prevention methods at the time.

Source. Own elaboration based on CrowdTangle Team, 2022.

The group presenting and promoting accepted prevention methods receives an increased number of angry reactions in September 2020, the month for local elections in Romania, in April 2021, a month of street demonstrations against masks and the medical personnel, and in August 2021, a month marked by discussions about a second pandemic school year and by the onset of the deadliest COVID wave in Romania.

In more than one in every three cases, the official account of the Romanian President is an outlier for *angry* reactions monthly means. Also, half of the monthly outliers for the *haha* reactions are linked to the same account. In January 2021, the post that showed Klaus Iohannis, the President of Romania, receiving his first vaccine against COVID-19 (Klaus Iohannis, 2021), had 1pp anger reactions and 10pp laughter reactions, out of all reactions gathered that month. For a post that says that the vaccine is safe and effective, and that restrictions and vaccination will help end the pandemic, most of the 7643 *haha* and *angry* reactions are, in fact, trolling.

The total monthly interest for COVID related posts variates also, with two peaks – March 2020, when the President declared the state of emergency and the Government established a generalized lockdown, and in April 2021, when dissatisfaction fueled by one right-wing member if the Parliament, Diana Ivanovici-Șoșoacă, resulted in street protests against medical personnel and against accepted prevention methods. In March 2020, there were 10,267,494 total interactions to COVID-related posts, while in April 2021 the number is three times bigger: 33,496,136 (see Graph. 1).

The number of total interaction per months is not linked to actual evolution of the disease (see *Data availability* for data and visualization), but with summer vacation, when the interest drops, with the electoral period in the last three months of 2020, when the interest increases, and with a campaign run by Diana Ivanovici-Șoșoacă, that was excluded from a right wing party in February 2021 and tried to redress her image with a violent anti-mask and anti-vaccination discourse. Șoșoacă's campaigned fueled street protests, in April 2021, and the discussions created around these street protests forced most people to take a position, for or against prevention methods, at the beginning of the general vaccination campaign.

In March 2020 the average number of total reactions, for the first 10,000 most influential posts, is more than 1000, and in April 2021, it grows to more than 3300. The lowest interest, with an average of less than 125 interactions per post per month, is shown in the summer of 2021, which prepared the disastrous forth wave of COVID-19 in Romania. In September 2021, posts about media coverage of COVID deaths get *haha* reactions from different Facebook users – a clear indication that trolls were crossing echo chambers borders on a regular basis.

A second step of the research was done to understand how these echo chambers functioned. A qualitative analysis looks at the first three most influential posts from each month of the study, labeling them as supportive or not supportive for generally accepted prevention instruments against COVID-19 and coding them according to previous research on inoculation theory, echo chambers and parasocial opinion leaders.

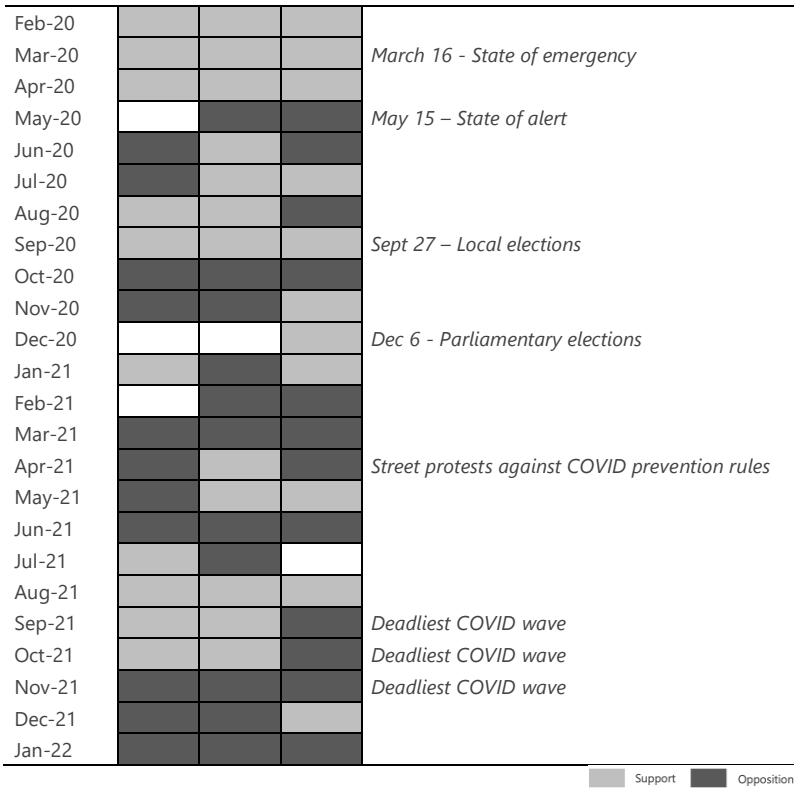
Out of the 72 most influential posts, more than a half were in the form of videos. CrowdTangle identified the search term COVID in the text of the post. It is highly probable that the selection did not include posts related to COVID, that were uploaded by individuals or by organizations, on Facebook, as video content, and had no text description using the search term COVID. Out of the 72 posts, one was incorrectly identified as COVID-related and one used the search word COVID as a very general context, but the subject was different. Nine of the 16 Facebook IDs selected for the quantitative analysis were present also in the corpus of most influential three posts, monthly.

Posts opposing accepted prevention rules existed since March 2020 in the general 10,000 monthly sample, but they make the top three most influential only in May 2020, after the President lifted the state of emergency and the Government replaced it with the state of alert (see Fig. 1.). George Simion, the leader of the right-wing Alliance for the Union of Romanians (AUR), entered the top three in July 2020, with a discourse against accepted rules. Five months later, AUR becomes the fourth party in the Romanian Parliament and the mainstream media acknowledged Simion's Alliance went under the general interest radar, gathering support online, on social platforms such as Facebook. The deadliest COVID wave, September-November 2021, was preceded by an increasing number of months with very influential posts against accepted prevention rules.

The only month in which older posts got in the top three is April 2021: one of the posts was supporting the prevention rules, two were against, and all three were from 2020. April 2021 is also the month with most total interactions for the first 10,000 most influential COVID-related post in Romanian, on Facebook, over a

period of two years, showing that in the month with street protests, people had a more intensive activity of searching, sharing, commenting and reacting to content on Facebook.

Figure 1. Support or opposition to prevention measurements against COVID



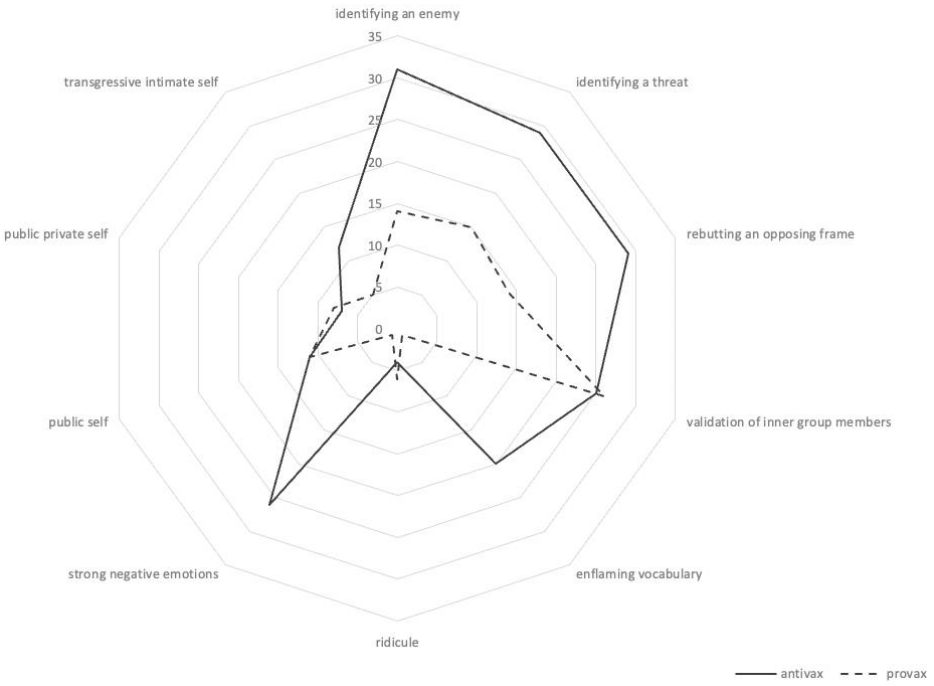
The first three posts each month, color-coded to show support or opposition to generally accepted prevention measurements against COVID, at that time. Each cell represents a post. The last column indicates noticeable events during the first two years of the pandemic.

Source. Own elaboration based on CrowdTangle Team, 2022.

There are 33 posts supporting generally accepted COVID prevention rules and 34 not supporting these rules – so the distribution is quite similar. The large majority of posts not supporting COVID prevention rules use the inoculation theory techniques, by identifying an enemy, identifying a threat related to COVID prevention and presenting a counterargument, to rebut possible opposing framing for COVID-19 (Graph 3). There is an active validation of inner group members. Most influential posts use an inflammatory choice of words and elicit strong emotions. Four posts out of 34 ridicule representatives of the pro-prevention group. People presented in the posts are during their public activities (public self - 11), are presenting part of their private lives, publicly (public private self - seven posts) or are so outraged, that can no longer control their feelings in public (transgressive intimate self – 12 posts).

In comparison, posts supporting rule following show individuals during their public activities (11 posts), or during private activities that are made public (eight posts). Five posts can be labeled as presenting transgressive intimate self moments. Out of 33 posts supporting COVID prevention rules, less than half identify an enemy or a threat related to COVID prevention. Fourteen posts are rebutting an opposing frame, related to COVID-19 prevention. The majority of these posts (26) are validating the members of the inner group, yet only one uses inflammatory vocabulary and elicits strong negative emotions. Six of the rule-following posts are using ridicule against the opposing group.

Graph 3. Content comparison for two echo chambers: a summary of the coding of most influential 72 COVID-related posts, Febr. 2020 – Jan. 2022.



Source. Own elaboration based on CrowdTangle Team, 2022. For full results, see *Data availability*.

Rule-opposing posts have almost four times more angry reactions (a mean of 2753 vs. 742, for rule-following posts), ten times more sad reactions (2835 vs. 263), and more that 50 percent more shares (32,366 vs. 20,271). *Angry* reactions mean disapproval for the rule-following posts, that are trying to keep a positive note, and approval and disapproval, for the rule non-following posts, that are talking about conspiracies and a fake illness. Rule following posts have more than two times more love reactions, as compared to the other group (2770 vs. 1237), but also almost six times more laughter reactions, even after controlling for comedy posts (a mean of 877 vs. 150). The content of rule-opposing posts has an angrier and more engaged audience, that are trolling more intensely rule following posts, with *haha* and *angry* reactions.

4. Discussion

A society needs to agree on basic facts in order to discuss solutions, but echo chambers filter facts ideologically, function as interpretive communities, increase cognitive distance among ideologically different groups and make dialog difficult - a phenomenon called polarization in the academic literature (Jamieson & Cappella, 2008; Kavanagh & Rich, 2018). In a public health crisis centered on an easily spreading virus, such as the COVID-19 pandemic, individual solutions do not function. People must act as a community that understands and follows prevention rules based on sound medical knowledge, in order to save as many lives as possible. In a context of limited trust in peers and in institutions, echo chambers can have damaging effects that are not limited to a toxic environment for public debates, but go as far as making innocent victims. This article verifies the existence of echo chambers in Romanian, on Facebook, on the issue of COVID prevention measures. Additionally, the article shows how echo chambers functioned and what were the cumulative spillover effects of echo chambers, on individuals and on the Romanian society, in general.

4.1. COVID-related echo chambers

This research aimed to verify, using statistical instruments, if echo chambers related to COVID-19 prevention existed on Facebook, in Romanian. The issue of echo chambers created on Facebook, to communicate with and to engage the ordinary social media user, is relevant on several dimensions, in the discussion about the COVID-19 pandemic. These dimensions include access to information, interpretation of relevant information and decision-making processes.

Echo chambers are understood as groups of individuals that share similar opinions, reinforced though constant interactions, leading to fragmentation, cult-like indoctrination, polarization, hatred, and violence (Sunstein 2001; Jamieson & Cappella, 2008; Sorial, 2017; Nguyen, 2020). The pandemic was a catastrophic event (Katz, 1996) that reduced fragmentation during the first months of the generalized lockdown: all individuals had access to similar information on illness and prevention measures, no matter what type of media they consumed on a regular basis. Yet, when vaccines became available in 2021, a small part of Romanians accepted to be immunized trough vaccination, a clear indication that decision making processes were negatively influenced by a series of factors. Romania and Bulgaria had the lowest vaccination rates in the European Union, that resulted in one of the deadliest COVID wave in the autumn of 2021, when other European countries were preparing to return to normal pre-pandemic activities.

There are several possible explanations for the lower rate of vaccination. This article explores only one of these possible explanations – the interactions on social media about COVID. Romanians have a lower-than-average level of trust in peers and in institutions, but a higher-than-average trust in social media (Newman et al, 2021; EVS/WVS, 2022). This apparent paradox can be explained by the existence of online echo chambers, that use discourse strategically, to define the inner group identity and to manipulate a lack of trust in member of the outer-groups (Sunstein 2001; Jamieson & Cappella, 2008; Nguyen, 2020). Vaccination campaigns are based on trust, thus a manipulation of trust resources, inside echo chambers, is a relevant research topic. Facebook was chosen because in Romania this social platform was the most used one, for news and for general purposes, during the sanitary crisis (Newman et al, 2021; Newman et al, 2022).

The article presents a mixed methods approach that showed that two Romanian language echo chambers, supporting and opposing generally accepted prevention rules, existed on Facebook in the first two years of the pandemic. The analysis also shows that parasocial opinions leaders used communication strategically to increase the cognitive distance between people accepting prevention rules and people contesting prevention rules, and some of these leaders had political aims.

The quantitative and the qualitative analyses, on a dataset of 233,242 posts labeled as COVID-related by CrowdTangle, in the period February 2020 – January 2022, support the echo chambers hypothesis, that people join online discussions with individuals that share similar perceptions of reality and, in time, develop similar behaviors. In the COVID-19 case, after an initial harmonious period of clarification about the illness and for effective prevention methods (February - June 2020), clear borders were demarcated among people who demonstrated institutional trust and accepted rule following for prevention and treatment of COVID and people who did not. The congealing of separate, opposing echo chambers was forced by politicians that instrumentalized a medical issue, during electoral periods (in the second half of 2020) and during a rearrangement of political leaders on the right (February-April 2021).

4.2. The functioning of echo chambers

In the two echo chambers similar communication methods are used, at different intensities, a qualitative analysis of the most influential three posts each month showed. The analysis used codes based on echo chambers: validation of members of in-group, inflammatory vocabulary, ridicule and strong negative emotions (Jamieson & Cappella, 2008; Nguyen, 2020), on inoculation theory: identifying an enemy, identifying a threat, preemptive counter-argumentation (Compton & Pfau, 2005; Lewandowsky et al, 2020) and on parasocial opinion leaders, with the three layers of public and private personae (Giles, 2002, Marwick & boyd, 2011; Marshall, 2010; Stehr et al, 2015).

Content creators from both echo chambers use inoculation theories techniques, identifying enemies and possible threats and presenting preemptive rebuttals of possible counterarguments to their main discursive line. Both echo chambers are actively validating inner group members, but the anti-prevention parasocial

leaders use, in addition, inflammatory vocabulary to provoke strong negative emotions and coagulate the group further around a common identity and against a common enemy.

In both echo chambers, public and private pieces of information are mixed, to support a connection between online personae and social media users. Thus, posts act as a call to action, making audience members, on both sides, react to, comment and share content they support, and content they do not support.

After a period of social representations clarifications, that lasted less than half a year, members of the echo chambers on prevention methods against COVID reached a common interpretation of available facts and saw others as enemies that presented an active threat to their wellbeing. Echo chamber communication and engagement supported trolling behaviors from both sides.

4.3. Cumulative spillover effects of echo chambers

The influence of parasocial opinion leaders is limited in time. For example, after the spike of interest, in April 2021, the number of total interactions with COVID-related posts drops to an all-time low. Social media users engage and react to content, one month, to act as if nothing happened, the next, on several occasions during the first two years of the pandemic. The data supports previous research on limited engagement during parasocial interactions (Giles, 2002) and the online disinhibition effect (Suler, 2004).

Yet, the effects of the echo chambers are felt on the longer term.

Members of the two groups transgressed the echo chambers borders and interacted with posts that were sharing a different world view. This interaction seems to have resulted in an increased radicalization - a possible explanation for the sarcastic *haha* reactions to COVID deaths mentions and *angry* reactions for all types of posting. In January 2021, when President Iohannis announced he is vaccinated, 10 percent of all reactions were of sarcastic *hahas* and 1 percent were *angry* reactions.

Anger and laughter reactions, with which the accounts of politicians and of media newsrooms are trolled on a daily basis, normalized violent reactions to public interest issues. The political instrumentalization of public health issues polluted the public debate and questioned experts' authority. As a result, the majority of the Romanian population, with little interpersonal and institutional trust to extend, was caught between the radicalized groups and watched in disbelief the transgressions against civility, delaying the vaccination process that led to the public health disaster in the fall of 2021.

We may consider, at first sight, the monthly variations in interactions with Facebook content as an increase or a decrease in polarization. A deeper understanding of processes that support echo chambers show how parasocial opinion leaders use discursive strategies to coagulate groups around an issue and to manipulate trust. In echo chambers, group identity is based on seeing the other as the enemy. Echo chambers normalize border transgression and violent behavior, for radicalized Facebook users, on both sides. Trolls act as a negative feedback loop for other echo chamber members, but their violent behavior is not helping the community reach a state of equilibrium, because dialog is impossible among declared public enemies. In a public health crisis, a state of equilibrium is reached when the crisis resolves itself with the loss of many innocent lives – a dark cumulative effect of low levels of trust in peers and in institutions, but of high level of trust in echo chamber communities, created on social media.

5. Conclusions

Many public debates are carried on offline and online platforms, combined. Social platforms facilitate echo chamber coagulation on public interest issues, polluting the public sphere and radicalizing supporters, that act as trolls, online, and, in extreme cases, recourse at violence, offline. This article documents the coagulation of two echo chambers on Facebook, in Romanian, on the issue of COVID prevention, and trolling activities that used the Facebook reaction buttons and were targeted at the representatives of the opposing group.

The article shows how parasocial opinion leaders produce toxic content for the public debate and how their followers act as trolls, with the help of Facebook reactions. Trolling activities have an array of different instruments on social platforms: reactions and comments on Facebook, retweets and comments on Twitter, comments on Instagram, comments and reactions on YouTube. Politicians and public institutions, alongside

newsrooms and journalists, are expected to maintain an active presence on social platforms, having a direct connection with the regular social media users and with trolls. While some platforms allow for moderation of some type of interactions, such as hiding comments on Facebook or blocking comments that used certain words, on Instagram, other types of interaction can not be controlled.

One *haha* reaction may seem meaningless, for the social media user and for the editor of a social media account. Yet, when a large percent of all reactions to a post promoting vaccines or to a post announcing COVID casualties are laughter reactions, this is the result of a consistent trolling activity in a highly polarized environment. Violence in public debates on key issues has damaging effects on the community and undermines trust in key democratic institutions, from central administration to individual newsrooms. Social platforms should take into account the cumulative spillover effects of all types of trolling activities and should consider limiting their effects.

This article verifies statistically if issue-related echo chambers existed on Facebook, in Romanian, and describes qualitatively, how they functioned. The article presents several features of toxic content online and documents the effects of the spread of this toxic content, looking at trolling activities that use Facebook reaction buttons. Based on the results of this research, two policy recommendations are developed, in order to increase civility in public debates and to nurture trust in fundamental democratic institutions.

1. **Algorithms supporting viralisation of toxic content**. Social platforms should refrain from providing algorithmic support for the viralisation of possibly toxic content, as described by research results. Actual moderation or a limitation of distribution, using algorithms, is not advised, because it may easily escalate to censorship.

2. **Social media features used for trolling.** Social platforms should acknowledge trolling activities and the instruments used for violent activities online and should reconsider the access to these instruments. More specifically, in the case of the reaction buttons, introduced recently by Facebook, a reconsideration of positive and negative effects of the reaction buttons, including a discussion on trolling, based on research results, should be helpful in creating a strategy to limit the damaging effects of trolling in public debates.

References

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, *118*(9). https://doi.org/10.1073/pnas.2023301118.

CNCAV - Comitetul Național de Coordonare a Activităților privind Vaccinarea împotriva COVID-19. (2021). *Actualizare zilnică (31/08) – evidența persoanelor vaccinate împotriva COVID-19*. https://vaccinare-covid.gov.ro/wp-content/uploads/2021/08/Tabel-situatie-vaccinari_31.08.2021.pdf, (accessed 26/08/2022).

Compton, J. A., & Pfau, M. (2005). Inoculation theory of resistance to influence at maturity: Recent progress in theory development and application and suggestions for future research. *Annals of the International Communication Association, 29*(1), 97-146. https://doi.org/10.1080/23808985.2005.11679045.

CrowdTangle Team. (2022). *CrowdTangle*. Facebook, Menlo Park, California, United States. Saved Search ID numbers: 1706708, 1706711, 1706714, 1706721, 1706722, 1706723, 1706740, 1706746, 1706750, 1706752, 1706759, 1706761, 1706776, 1706779, 1706785, 1706789, 1706800, 1706801, 1706805, 1706810, 1706812, 1706813, 1706816, 1706817.

Dahlgren, P. M. (2021). A critical review of filter bubbles and a comparison with selective exposure. *Nordicom Review*, *42*(1), 15-33. https://doi.org/10.2478/nor-2021-0002.

EVS/WVS (2022). European Values Study and World Values Survey: Joint EVS/WVS 2017-2022 Dataset (Joint EVS/WVS). JD Systems Institute & WVSA. Dataset Version 3.0.0. https://doi.org/10.14281/18241.19.

Fletcher, R., & Nielsen, R. K. (2018). Automated serendipity: The effect of using search engines on news repertoire balance and diversity. *Digital Journalism*, *6*(8), 976-989. https://doi.org/10.1080/21670811.2018.1502045.

Giles, D.C. (2002). Parasocial Interaction: A Review of the Literature and a Model for Future Research. *Media Psychology, 4*(3), 279-305. https://doi.org/10.1207/S1532785XMEP0403_04.

Jamieson, K. H., & Cappella, J. N. (2008). *Echo chamber: Rush Limbaugh and the conservative media establishment*.

Oxford University Press.

Katz, E. (1957). The two-step flow of communication: An up-to-date report on an hypothesis. *Public opinion quarterly*, *21*(1), 61-78. https://doi.org/10.1086/266687.

Katz, E. (1996). And deliver us from segmentation. *The Annals of the American Academy of Political and Social Science*, *546*(1), 22-33. https://doi.org/10.1177/0002716296546001003.

Kavanagh, E., Jones, I., & Sheppard-Marks, L. (2016). Towards typologies of virtual maltreatment: Sport, digital cultures & dark leisure. *Leisure Studies*, 35(6), 783-796. https://doi.org/10.1080/02614367.2016.1216581.

Kavanagh, J. & Rich, M. D. (2018). *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life*. Rand Corporation.

Klaus I. (2021). *M-am vaccinat astăzi împotriva COVID-19. Procedura este simplă, nu doare, iar vaccinul este sigur și eficient. Recomand tuturor vaccinarea, care*. [Status update]. [Pictures]. Facebook. https://www.facebook.com/632184113535510/posts/3654493634637861.

Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E. Rand, D. G., Rapp, D. N., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, S., Vraga, E. K., Wood, T. J., Zaragoza, M. S. (2020). *The Debunking Handbook 2020*. https://sks.to/db2020.

Mark Zuckerberg. (2016, Febr. 24). *Today is our worldwide launch of Reactions -- the new Like button with more ways to express yourself. Not every*. [Status update]. [Video]. Facebook. https://bit.ly/3E7fLhE

Marshall, P.D. (2010). The promotion and presentation of the self: celebrity as marker of presentational media. *Celebrity Studies, 1*(1), 35-48. https://doi.org/10.1080/19392390903519057.

Marwick, A., & Boyd, D. (2011). To see and be seen: Celebrity practice on Twitter. *Convergence*, *17*(2), 139-158. https://doi.org/10.1177/1354856510394539.

Nguyen, C. (2020). Echo chambers and epistemic bubbles. *Episteme*, *17*(2), 141-161. https://doi.org/10.1017/epi.2018.32.

Newman, N., Fletcher, R., Schulz, A., Andı, S., Robertson, C.T. and Nielsen, R.K. (2021). *Digital News Report 2021*. Reuters Institute for the Study of Journalism. https://bit.ly/3fZ5WKl

Newman, N., Fletcher, R., Robertson, C. T, Eddy, K. and Nielsen, R. K. (2022). *Digital News Report 2022*. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf.

Ooms, J. (2021). *writexl: Export Data Frames to Excel 'xlsx' Format. R package version 1.4.0*. https://CRAN.R-project.org/package=writexl.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. The Penguin Press.

Phillips, W. (2011). LOLing at tragedy: Facebook trolls, memorial pages and resistance to grief online. *First Monday*. https://doi.org/10.5210/fm.v16i12.3168.

Radu, R.N. (2015). Protejarea vieții private. In Radu, R.N. (ed.) *Deontologia comunicării publice,* 135-160. Polirom.

Radu, R. N. (2019). Externalities and journalism. In Vos, T.P. & Hanusch, F. (ed.) *The International Encyclopedia of Journalism Studies*, 1-7. John Wiley & Sons. https://doi.org/10.1002/9781118841570.iejs0073.

R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Revelle, W. (2021). *psych: Procedures for Personality and Psychological Research*. Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 2.1.6.

Rogers, E.M. (2003). *Diffusion of Innovations* (5th ed.). Free Press.

Sorial, S. (2017). The Expression of Anger in the Public Sphere. *Journal of Social Philosophy, 48*(2), 121–143. https://doi.org/10.1111/josp.12188.

Stehr, P., Rössler, P., Schönhardt, F., & Leissner, L. (2015). Parasocial Opinion Leadership. Media Personalities' Influence within Parasocial Relations: Theoretical Conceptualization and Preliminary Results. *International Journal of Communication, 9*, 982–1001.

Suler, J. (2004). The Online Disinhibition Effect. *Cyberpsychology & Behavior, 7*(3), 321-326. https://doi.org/10.1089/1094931041291295.

Sunstein, C.R. (2001). *Republic.com*. Princeton University Press.

Toff, B., & Kalogeropoulos, A. (2020). All the news that's fit to ignore: How the information environment does and does not shape news avoidance. *Public Opinion Quarterly, 84*(S1), 366-390. https://doi.org/10.1093/poq/nfaa016.

Tufiș, C. (2020). Despre necesitatea încrederii în instituțiile sistemului politic. Voicu, B., Rusu, H., Tufiș, C., [ed], *Atlasul valorilor sociale*, 46-53. Presa universitară clujeană.

Velásquez, N., Leahy, R., Johnson Restrepo, J., Lupu, Y., Sear, R., Gabriel, N., Jha, O. K., Goldberg, B. & Johnson, N. F. (2021). Online hate network spreads malicious COVID-19 content outside the control of individual social media platforms. *Scientific Reports, 11*, 11549. https://doi.org/10.1038/s41598-021-89467-y.

Weimann, G. (2017). Multistep flow of communication: Evolution of the paradigm. Rössler, P., Hoffner, C.A. & van Zoonen, L. (ed) *The international encyclopedia of media effects*, 1-10. John Wiley & Sons, Inc. https://doi.org/10.1002/9781118783764.wbieme0055.

Waisbord, S. (2020). Mob censorship: Online harassment of US journalists in times of digital hate and populism. *Digital Journalism*, *8*(8), 1030-1046. https://doi.org/10.1080/21670811.2020.1818111.

WHO - World Health Organization. (2022). *WHO Coronavirus (COVID-19) Dashboard*. https://covid19.who.int/data (accessed 20/08/2022).

Wickham *et al*. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686, https://doi.org/10.21105/joss.01686.

Xu, J., Sun, G., Cao, W., Fan, W., Pan, Z., Yao, Z. & Li, H. (2021). Stigma, Discrimination, and Hate Crimes in Chinese-Speaking World amid Covid-19 Pandemic. *Asian Journal of Criminology, 16*, 51–74. https://doi.org/10.1007/s11417-020-09339-8.

Zollo, F., Bessi, A., Del Vicario, M., Scala, A., Caldarelli, G., Shekhtman, L., Havlin, S. & Quattrociocchi, W. (2017). Debunking in a world of tribes. *PloS one*, *12*(7). https://doi.org/10.1371/journal.pone.0181821.

Brief CV of the author

Professor Raluca-Nicoleta Radu (PhD in Communication Studies, University of Bucharest, 2009) is Director of the Journalism Department, Faculty of Journalism and Communication Studies, University of Bucharest, Romania and visiting fellow at the Center for Advanced Internet Studies, Bochum, Germany. She has more than 20 years of research and teaching experience in media ethics, media economics, cultural industries and comparative studies

Funding

The research was supported by a visiting fellowship at the Center for Advanced Internet Studies (CAIS) GmbH, July-September 2022.

Acknowledgements

This paper is an improved version of Radu, R.N. (2022). `How I learned to hate you. Crossing borders on Facebook, before the 2021 COVID sanitary disaster`. *Borders and transfers. The Society for Romanian Studies 2022 conference* in Timișoara, Romania, 15-17 June, 2022. I want to thank Boaz Miller, Aaron Neiman, Leyla Safta-Zecheria, Bogdan Voicu and Galit Wellner for our discussions on this subject.

Data availability

Supplementary materials for this paper include the R code used for data manipulation, visualization and analysis, the qualitative coding, supplementary visuals, four sets of outliers, COVID data from WHO, data on trust in peers and trust in social media, descriptive statistics for the CrowdTangle selected data and ANOVA results. The materials may be found here: https://doi.org/10.7910/DVN/9NEJMZ.