Disinformation and Democracy: Media Strategies and Audience Attitudes

Editors

Pere Masip, Bella Palomo and Guillermo López

COGITATIO

COGITATIO

Table of Contents

Article

Disinformation in Facebook Ads in the 2019 Spanish General Election Campaigns

Lorena Cano-Orón [1,*], Dafne Calvo [2], Guillermo López García [1] and Tomás Baviera [3]

[1] Department of Language Theory and Communication Sciences, University of Valencia, 46010 Valencia, Spain;
E-Mails: lorena.cano@uv.es (L.C.-O.), guillermo.lopez@uv.es (G.L.G.)
[2] Department of Early Modern History, Modern History and History of America, Journalism and Audiovisual
Communication and Advertising, University of Valladolid, 47007 Valladolid, Spain; E-Mail: dafne.calvo@uva.es
[3] Department of Economics and Social Sciences, Universitat Politècnica de València, 46022 Valencia, Spain;
E-Mail: tobapui@upv.es

* Corresponding author

Abstract
As fake news elicits an emotional response from users, whose attention is then monetised, political advertising has a significant influence on its production and dissemination. Facebook ads, therefore, have an essential role in contemporary political communication, not only because of their extensive use in international political campaigns, but also because they address intriguing questions about the regulation of disinformation on social networking sites. This research employs a corpus of 14,684 Facebook ads published by the major national political parties during their campaigns leading up to the two Spanish general elections held in 2019. A manual content analysis was performed on all the visually identical ads so as to identify those containing disinformation and those denouncing it. The topics addressed in these ads were then examined. The results show that the political parties' Facebook ad strategies were akin to those of conventional advertising. Disinformation messages were infrequent and mainly posted by Ciudadanos and VOX. Nonetheless, it is striking that the main topic addressed in the ads was the unity of Spain—precisely the issue of Catalonia's independence. In light of this, it can be deduced that 'traditional' parties are taking longer to renounce classical forms of campaigning than their 'new' counterparts, thus demonstrating that the actors implementing disinformation strategies are not only restricted to the extreme right of the ideological spectrum.

Issue
This article is part of the issue "Disinformation and Democracy: Media Strategies and Audience Attitudes," edited by Pere Masip (University Ramon Llull, Spain), Bella Palomo (University of Malaga, Spain) and Guillermo López (University of Valencia, Spain).

1. Introduction

The phenomenon of digital disinformation has become particularly relevant in recent years, not only due to its widespread use in many countries (Bradshaw & Howard, 2019), but also because of the Internet's capacity to amplify it, which, in turn, has repercussions for political life insofar as it undermines reliable information sources (Benkler, Faris, & Roberts, 2018; Bennett & Livingston, 2018). Although some of those repercussions are unforeseen, others are clearly intentional. Campaigns of this sort are mainly launched on social networking sites such as Facebook, Twitter, Instagram and WhatsApp, and the main techniques are based on the use of bots (automated accounts that mimic human behaviour) and trolls (fake accounts managed by humans with specific intentions;

Bradshaw & Howard, 2019). Actually, it is not easy to identify who promotes these campaigns, although there are always political interests behind them.

Political parties are also active on social media. They post content like any other user, but also leverage advertising techniques, inserting sponsored content in news feeds in order to take advantage of the microtargeting capabilities of social media. Due to the Cambridge Analytica scandal, Facebook made sponsored content available to any user through its Facebook Ad Library (Hern, 2018), thus allowing the public to scrutinise the paid messages posted by political parties on this social networking site. As such content was inaccessible only a few years ago, research on digital disinformation has yet to explore political advertising on Facebook.

Some of the first elections to be called after this disclosure were held in Spain, specifically general elections on 28 April and 10 November 2019. So the election campaigns running up to them were preceded by the disinformation scandals in the United States, the United Kingdom and Brazil. Studies of this electoral cycle have revealed that disinformation was mainly spread via Twitter, Facebook and WhatsApp (Paniagua Rojano, Seoane Pérez, & Magallón-Rosa, 2020; Rodríguez-Fernández, 2020).

Accordingly, the aim of this article is to contribute to the research on digital disinformation by analysing the messages conveyed by the major Spanish political parties in the 2019 general election campaigns, using the Facebook Ads tool. The main research question posed here is whether or not those parties implemented any disinformation strategy as regards their sponsored content on Facebook. As far as can be gathered from the literature, the dissemination of political disinformation via Facebook advertising is still an uncharted research avenue. Whereby the originality of this study, whose purpose is to fill that research gap, lies in the fact that it is presumably the first to focus on the Facebook Ads Library.

The presence of disinformation in the advertising messages of the major political parties in Spain, albeit rather thin on the ground, is by no means a minor object of study, inasmuch as those posting them have a lot to gain from this in the political contest. Even though Spanish political ads are by and large conventional, the findings of this study suggest they can include false information with an eye to manipulating public opinion.

This article is structured as follows. First, the theoretical framework is developed, distinguishing between several concepts relating to disinformation, such as fake news and propaganda. This is followed by a description of how the 14,684-ad database was created and of the content analysis methodology employed. Lastly, the findings are presented and discussed.

2. Theoretical Framework

The development of computational techniques for content creation and distribution (Tucker, Theocharis,

Roberts, & Barberá, 2018) has contributed to place disinformation at the centre of political processes at an international level. The growing academic literature has explored disinformation operations in diverse contexts and at different hierarchical levels. This is the case of the use of Twitter by the president of the United States, Donald Trump (Marwick & Lewis, 2017; Ross & Rivers, 2018); the use of bots during the French presidential elections (Ferrara, 2017); the distribution of emotional and polarised information in the United Kingdom during the Brexit referendum campaign (Bastos & Mercea, 2018; Cervi & Carrillo-Andrade, 2019); the referendum on self-determination in Catalonia (Stella, Ferrara, & De Domenico, 2018); and the recent publication of false information on Covid-19 in order to bolster Jair Bolsonaro's leadership (Ricard & Medeiros, 2020). All these processes have highlighted the crucial role played by reliable information in any democratic system. Consequently, terms such as disinformation, fake news or even propaganda have entered into the public debate on politics. Nevertheless, these concepts are not easy to distinguish, as they are used in different ways in diverse scenarios (Magallón-Rosa, 2019).

There still is not a unanimous consensus on the definition of fake news. After the popularisation of this expression by the media, the concept has become even more vague (Bennett & Livingston, 2018; Ireton & Posetti, 2018). In their review of the academic literature on fake news, Tandoc, Lim, and Ling (2018) discovered that the term was indeed used to define up to six different types: news satire, news parody, fabrication, manipulation, advertising and propaganda. In addition to the efforts that have been made to identify and describe the disinformation tactics of parties and public authorities, some authors have contributed to research on this phenomenon by establishing taxonomies that, as analytical tools, shed light on the boundaries between propaganda and disinformation. These notions have been considered to be interchangeable by some authors, as is the case with the terms propaganda and publicity (Tandoc et al., 2018). According to Woolley and Howard (2016, p. 4886):

> Computational propaganda involves software programs that are interactive and ideologically imbued….They are ideological, first, in that they are programmed to promote a particular perspective in politically charged conversations and, second, in that they are artifactual evidence of the idea that technology can be used to influence politics.

For Jack (2017), propaganda is a pejorative term per se, as opposed to other concepts such as advertising, public relations and public diplomacy. The first term refers to a deliberate intention to manipulate or deceive, which "can refer to political communications, advertising, and even junk mail" (Jack, 2017, p. 7).

The epistemological difficulty in differentiating between disinformation and propaganda stems from

two of the main characteristics that both concepts share: the hyperpolarisation of their content and their creation for political purposes (Nielsen & Graves, 2017). Therefore, although it cannot be called fake news in the strict sense of the word, information released by governments, public agencies and parties reflects their agendas in a way that differs from that of objective and evidence-based information (Molina, Sundar, Le, & Lee, 2019). Propaganda and advertising include information that is usually based on facts, but with a bias that helps to present the propagandist or advertiser in a favourable light. The use of native advertising for disseminating this type of content makes it seem more credible, due to its similarity to other media publications or social media posts (Tandoc et al., 2018). In contrast, disinformation is a phenomenon that deliberately seeks to confuse the receiver with false or misleading information, with the aim of achieving persuasive goals or discrediting an opponent. Disinformation, from this perspective, can be regarded as a powerful propaganda tool (Benkler et al., 2018; Martin, 1982). Following Bennett and Livingston's (2018) reasoning, the phenomenon of fabricating and disseminating intentionally false information will be referred to here as disinformation. The intention is to convey its complexity, focusing not only on false information per se, but also on the strategies behind it.

This grey area between disinformation, propaganda and advertising evinces the complexity of this phenomenon in contemporary political processes. It gives a good account of the multiple and sophisticated tactics that are deployed to gain control of the narrative and to manipulate the electorate (Marwick & Lewis, 2017). Such tactics have given rise to a broad debate on their short- and long-term effects on democratic systems, whose legitimacy is currently being questioned, reflected in a decline in the credibility of politicians and public institutions in the eyes of the public at large (Bennett & Livingston, 2018; Tucker et al., 2018).

In this complex digital landscape, social media advertising has been operating as an opaque way to disseminate information. Through these tools, companies can send their ads to specific social media audiences in such a way that only they view the message (Tufekci, 2015; Woolley & Howard, 2018). The capacity of social networking sites to segment audiences is based on the rich digital footprint left by users on them (Kim et al., 2018; Sinclair, 2016). In this vein, several researchers have raised the need to act on sponsored content owing to its potential to disseminate disinformation (Gray, Bounegru, & Venturini, 2020). Particularly in the case of Facebook, some scholars have shown how this platform has been used to divide the population and to misinform, specifically in the case of ads paid for by Russia's Internet Research Agency in the United States (Lukito, 2020; Ribeiro et al., 2019).

Social networking sites are very appealing to advertisers. The business model of these sites, of which Facebook is the paradigm, has been built on their ad services (Dommett & Power, 2019; Kreiss & McGregor, 2018). Not only companies but also political parties' figure among their advertisers. However, the Cambridge Analytica scandal obliged Facebook to modify some of its practices. This social networking site decided, among other things, to put an end to the lack of transparency of the content sponsored by political parties. Accordingly, in April 2018 the Facebook Ad Library was launched with the aim of offering the public access to the paid messages of political parties being disseminated in its news feed (Hern, 2018).

This issue of social media accountability forms part of a broader debate on seeking solutions to an increasingly polarised, uninformed, and fragmented networking ecosystem (Bakir & McStay, 2018). These sceptical views ultimately beg the question of the extent to which genuine civic engagement with democratic systems can be ensured on platforms that, as in the case of Facebook, rely on user metrics and the imperatives of advertising to turn a profit (Jack, 2017). The aim of this article is to offer an answer to this research question by focusing on the role played by political parties as clients of Facebook and how they address the controversial issue of disinformation.

3. Methodology

The object of study were the ads posted on Facebook by the Spanish Socialist Workers' Party (hereinafter PSOE), the People's Party (hereinafter PP), Ciudadanos, Unidas Podemos—a coalition between Podemos and United Left (hereinafter IU)—and VOX, the five parties with the best election prospects, during the pre- and election campaigns running up to the two general elections held in Spain in 2019 (on 28 April and 10 November). In order to perform an in-depth analysis on the kind of strategies that they were implementing, it was decided to consider Podemos and IU separately because their Facebook ads were mainly posted on their party pages, rather than on the coalition page.

The corpus was obtained through a web crawler written in Python. The Facebook Ad Library displays the ads paid for by a particular organisation in grid-format. This initial presentation provides the basic content of the ads with a link to access the metadata. This metadata, which is displayed on a new webpage, includes run time, advertiser spend, impressions and basic demographic segmentation, namely, sex, age, and autonomous community (geographical region). The first three metadata sets are shown in text format, while the last three are provided in image files. The crawler visited the library page of each political party and, after verifying that the ads corresponded to the two general elections under study, accessed each one of the ad metadata webpages, before downloading the content (text and image/video) and the metadata (text and image files) of the ad in question.

The corpus was made up of 14,684 ads (see Table 1 for their distribution by political party). As the Facebook

Table 1. Information on the ads making up the corpus.

Political Party	No. ads	Estimated spend (€)	Estimated impressions
Ciudadanos	8,560	584,100	70,436,500
IU	18	4,750	1,207,500
PP	4,517	489,350	60,235,500
Podemos	924	995,100	129,992,000
PSOE	621	69,350	22,767,500
VOX	44	2,800	881,500
Total	14,684	2,145,450	285,520,500

Ad Library only provides ranges of values, the data on advertiser spend and impressions were not exact figures. To facilitate comparisons, ad spend, and impressions were estimated, taking into account the mean value of the range. For example, if the spend range of an ad was €100–€499, its mean value €300 was taken. The same procedure was used to calculate ad impressions. As a result, it was estimated that the five political parties spent approximately €2 million on the two election campaigns, with roughly 286 million impressions. Ciudadanos and the PP were the two parties posting the highest number of ads, while Podemos and Ciudadanos made the largest investment.

Although the data are presented here in an aggregated manner, it should be explained that the number of ads posted in both campaigns was not proportional (see Figure 1). This might have been owing to the fact that the campaign for the 10 November elections lasted half as long as that for the 28 April elections, due to both a Spanish law limiting the campaign period for repeated general elections and the more austere cam-

paigns designed by all the parties after a year replete with election calls in Spain.

Following an initial data analysis, it was noted that there were numerous repetitions in the ad content. This was due to the microtargeting capabilities of Facebook: The same content was used in multiple ads, each with a different sociodemographic audience profile, which will be referred to hereinafter as 'visually identical ads' (VIAs). They cannot be regarded as mere duplicates because, albeit with the same content, each VIA had a different target audience (e.g., users in specific geographical locations), spend and number of impressions. Each ad formed part of a complex communication strategy aimed at engaging a predetermined Facebook audience with a particular discourse. To perform the analysis, only these VIAs (1,743) were used.

Two variables were defined for the content analysis. The first is the type of information contained in the ads, which fell into the following two categories: 1) does not contain facts (generic calls to vote or election promises); 2) reports disinformation (the party presents itself as a
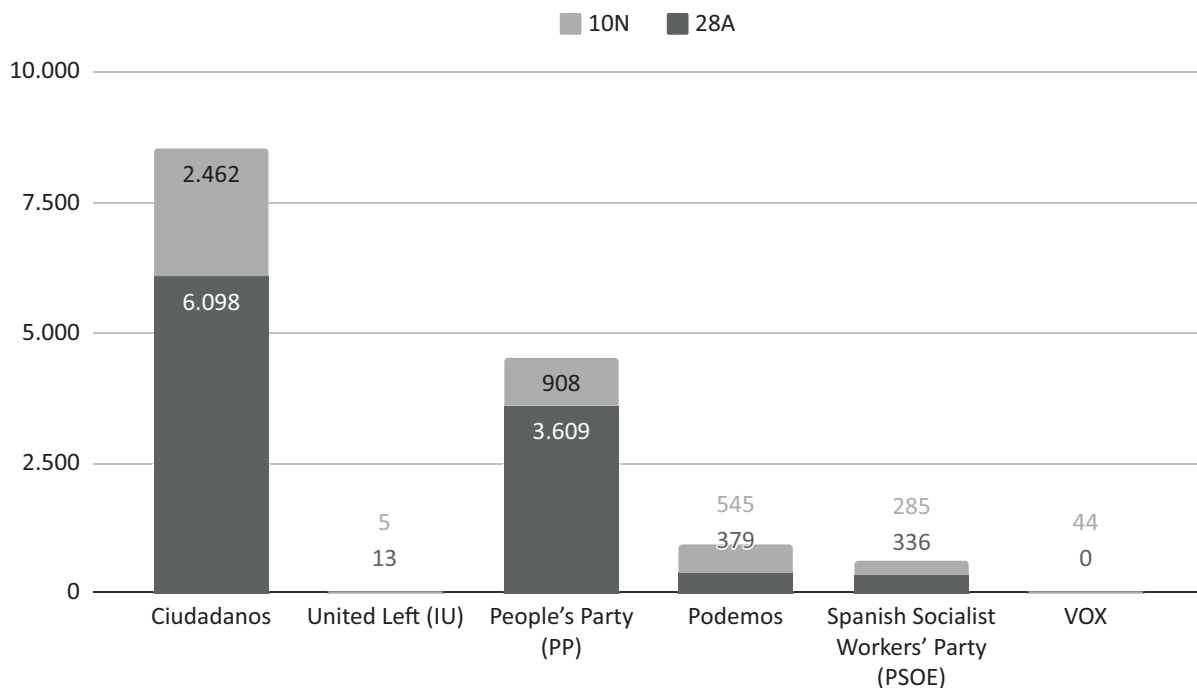

Figure 1. Ads posted by the parties during the 28 April and 10 November election campaigns (2019).

victim of disinformation, manipulated by the media or its political rivals); 3) contains disinformation (when ad content is misleading or directly false); 4) and contains facts (their interpretation may sometimes be more or less biased, but nonetheless cannot be regarded as disinformation). All the ads that contained verifiable facts were fact-checked through news media and other information sources.

The second variable is linked to the main topics of the ads, which were classified in the following categories, designed on the basis of a preliminary study of the corpus: 1) employment; 2) party promotion; 3) Spain's cohesion (Catalan independence issue, glorification of the country); 4) social policy; 5) economic policy; 6) feminism; 7) pacts, coalitions and surveys; 8) education and science; 9) environment; 10) democratic quality (corruption); 11) empty Spain (referring to the depopulation of rural areas in the interior); 12) international policy; 13) immigration; 14) infrastructure; and 15) others.

Different teams, each formed by two of the authors of the research, coded each variable. Inter-rater reliability was evaluated using Krippendorff's alpha, obtaining acceptable values for the (dis)information ($\alpha = 0.808$) and topic ($\alpha = 0.904$) variables.

4. Results

4.1. Use of Disinformation

The presence of disinformation in the electoral propaganda distributed on Facebook by the five main Spanish political parties during the two 2019 general election campaigns was, as can be seen from the data, negligible (see Table 2). There were few occasions on which disinformation appeared in the ads' copy, either as a technique of persuasion or as a complaint.

As can be observed, the vast majority of the ads fell into the first category: they contained phrases that generically appealed to the electorate, attempted to drum up support for the political party in question or set out the party's electoral proposals, the veracity of which cannot be verified, of course, until the party is in a position to implement them either by forming a government or by being able to exert political influence. Of the total number of ads posted, 83.71% (78.03% when including the VIAs) fell into this category.

Regarding the ads that included potentially verifiable claims, most of them were accurate or based on real-ity, thus being closer to propaganda discourse than to disinformation. Of the total number of ads, 14.22% (rising to 20.14%, when including the VIAs) belonged to this category.

All of which means that only 2% of the ads contained disinformation. This is by no means a negligible proportion, but it does indicate that it is not a technique applied by Spanish political parties across the board, neither as a discursive device to report being victims of disinformation operations (12 ads in total) nor as examples of disinformation per se (292 ads, 1.99% of the total).

4.2. Disinformation by Party

Table 3 shows the distribution by party (in both election campaigns) of the different categories used here to detect disinformation or its absence. As before, a distinction can be drawn between the VIAs and the total number of ads posted. It can be seen that the incidence of disinformation parameters was minimal in general and non-existent in the case of IU and the PSOE. The latter developed a campaign focusing on the formulation of electoral proposals or campaign slogans, which did not even include facts that could be checked or verified.

In the main, Ciudadanos was the party posting the highest number of ads containing elements of disinformation, while the PP, Podemos and VOX did so to a lesser extent. Only these last two parties posted ads reporting alleged disinformation practices against them. Table 4 shows the incidence of disinformation in the relative percentage of each party and the corpus as a whole.

Podemos reported disinformation campaigns against it in three different ads. The first recommended a two-minute video explaining how "the state's rotten apples" (Podemos, 2019) were intriguing against the party and its leader, Pablo Iglesias, focusing on the alleged manoeuvres of the Ministry of the Interior controlled at the time by Jorge Fernández Díaz (PP), which in theory were being silenced by the media and the powers that be: political institutions, economic elites and members of the judiciary and law enforcement agencies. This same line of argument was deployed in the other two ads, one of which commented on the party's intention not to "spend hours talking to media outlets owned by banks" (Podemos, 2019), while the other invited readers to receive the party's information directly on their cell phones to avoid "fake news against Podemos" (Podemos, 2019).

Table 2. Overall results of the disinformation analysis.

Disinformation	No. VIAs	% VIAs	Total no. ads	% Total ads
Does not contain facts	1,360	78.03%	12,292	83.71%
Reports disinformation	11	0.63%	12	0.08%
Contains disinformation	21	1.20%	292	1.99%
Contains facts (not disinformation)	351	20.14%	2,088	14.22%
Total	1,743	100.00%	14,684	100.00%

Table 3. Distribution of the disinformation variable by party (%).

Party	Does not contain facts		Reports disinformation		Contains disinformation		Contains facts (not disinformation)	
	VIAs	Total	VIAs	Total	VIAs	Total	VIAs	Total
Ciudadanos	10.96	46.84	0	0	0.70	1.89	2.58	9.57
IU	0.69	0.11	0	0	0	0	0.11	0.01
PP	30.75	26.65	0	0	0.23	0.06	14.17	4.05
Podemos	12.97	5.69	0.34	0.04	0.06	0.02	3.10	0.54
PSOE	22.38	4.23	0	0	0	0	0	0
VOX	0.29	0.20	0.29	0.04	0.11	0.02	0.17	0.04
Total	78.03	83.71	0.63	0.08	1.20	1.99	20.14	14.22

Regarding VOX, the high proportion of VIAs (33%, albeit accounting for a scant proportion of the total number of ads and therefore not representative) in which this party presented itself as a victim of disinformation is noteworthy. In all cases, the party used the same text—"At VOX we play fair, we do not use the left's black propaganda techniques or spread hoaxes as others do" (VOX, 2019)—together with an image corresponding to the main message, in an attempt to counter calls for tactical voting, namely that voting for VOX in most provinces meant giving seats to the PSOE or Podemos.

When the data in this category is crossed with the thematic classification (see Table 5), it can be seen that Podemos reported disinformation about issues pertaining to democratic quality and party promotion, whereas VOX, as well as mentioning the same topics, also referred to the cohesion of Spain and the environment. Likewise, considering the difference between the VIAs and the total, it can be observed that this kind of message was not amplified (i.e., repeated in the corpus) and that it was a marginal issue.

4.3. Ads with Disinformation

The incidence of ads incorporating elements of disinformation is shown in Table 6. The parties that resorted most to this kind of discourse were VOX (13.33% of the total number of VIAs) and Ciudadanos (5.6%). Whereas Podemos and the PP were the two parties that posted this type of message a lot less frequently.

As to the PP, there was hardly any disinformation in its ads. Instead, the party chose to include information that best contrasted its management with that of the Sánchez government (PSOE). For example, in those provinces with rising unemployment rates, it was Sánchez's fault; in those where jobs were being created, it was thanks to

Table 4. Reports of cases of disinformation against parties (%).

Party	VIAs	VIAs (n = 1,743)	Total (N = 14,684)
Ciudadanos	0	0	0
IU	0	0	0
PP	0	0	0
Podemos	2.09	0.34	0.04
PSOE	0	0	0
VOX	33.33	0.29	0.04
Total		0.63	0.08

Table 5. Reports of disinformation classified by party and by topic (%).

Theme	Podemos		VOX		Total
	VIAs	Total	VIAs	Total	
Spain's cohesion	0.00	0.00	13.33	0.01	0.01
Democratic quality	1.05	0.02	6.67	0.01	0.03
Environment	0.00	0.00	6.67	0.01	0.01
Party promotion	1.05	0.02	6.67	0.01	0.03
Total	2.09	0.04	33.33	0.04	0.08

![Cogitatio logo]

the efforts of "everyone," the "entrepreneurs." The same can be said about Podemos. By and large, this party did not cross the line separating an interpretation of reality that was favourable to its interests but fact-based, from disinformation or the dissemination of false information. Both parties posted a single ad that can be regarded as having contained disinformation.

The PP's ad, which was repeated on several occasions, read as follows: "Whenever the PSOE governs, it gives rise 🚀 to uncertainty. Something called risk premium increases ☑. We all pay for it. Vote for the People's Party #SafeValue. ☞ www.ppvalorseguro.es" (Partido Socialista, 2019).

Although Spain's risk premium (the rate of return that a country issuing sovereign bonds has to pay over and above the risk-free rate of return) rose during the PSOE governments (especially in the last years of the Rodríguez Zapatero government), this has been by no means a systematic trend. Indeed, during the first year of the Rajoy government, Spain's risk premium hit a euro-era record (Expansión, n.d.).

There was also disinformation in an ad about the wage gap:

> If you're a woman, you're paid 22.3% less than a man for doing the same job. This means that it's like working for FREE for two months of the year. On 28 April, your vote for Unidas Podemos can guarantee equal pay by law. (Podemos, 2019)

On average, women still earn 22.3% less than men. Be that as it may, this should be qualified. Women do not earn 22.3% less than men for doing the same job, but, as a rule, men have better-paid jobs. Other similar ads, plus the video accompanying this same ad, did explain this correctly.

Regarding Ciudadanos, its use of disinformation can be directly associated with its often-aggressive discourse. This belligerence, combined with a more varied, sophisticated and segmented discourse than that of the PP, meant that it was more liable to include inaccuracies and fallacies in its ads than the PP, whose ads tended to be more traditional and ingratiating.

One such example can be found in the large number of ads posted by Ciudadanos on occasion of its recruitment of Edmundo Bal, the public prosecutor who was removed from the case against the pro-independence 'procès,' due to his discrepancies with the Government's line of action. Bal had sought a conviction for rebellion versus the accusation of sedition for which the Attorney General's Office pressed and for which the majority of the prosecuted politicians were ultimately condemned. According to these ads, "🗨Sánchez removed him as a favour to Torra, Rufián and Puigdemont so as to continue in the Moncloa" (Ciudadanos, 2019), something that has yet to be substantiated.

In other ads that, with slight variations, were also repeated, Ciudadanos contended: "While the PP and

the PSOE were making a pact with Pujol and ERC, Inés Arrimadas and Albert Rivera were fighting for the freedom and equality of all Spaniards in Catalonia" (Ciudadanos, 2019). However, the last time the PP struck an agreement with Jordi Pujol was in the year 2000. The Catalan leader abandoned active politics in 2003, three years before Ciudadanos appeared on the scene. In short, in this ad, as in the previous one, the truth (the agreements that the PP and the PSOE had brokered with the nationalists in the past) was distorted by mixing it with false or misleading information.

Ciudadanos also posted seven more VIAs that were reported by various media for their misleading content (Sarabia, 2020). These ads were aimed at the Andalusian provinces, except for Cordoba, encouraging the electorate to vote for the party as it was just a few votes away from winning a seat at the expense of Podemos. As the fact-checker Maldita ("La publicidad de Ciudadanos," 2019) pointed out, however, it was impossible for the party to know this and, therefore, to claim:

> 👀 In the province of Huelva, Ciudadanos is 190 votes away from winning a decisive seat at the expense of Podemos. We all want to oust Sánchez from the Moncloa, but in Huelva, voting for Ciudadanos is the best way to achieve this. (Ciudadanos, 2019)

Lastly, despite the party's scant visibility on Facebook, VOX also included disinformation in its ads. In 15 different ads, most of which were integrated into the same VIA, there were elements of disinformation in two. The first read as follows: "The PP, the PSOE and Ciudadanos: different paths, same objectives. Politically, socially and now also linguistically imposing Catalan on the Balearic Islands, with VOX as the only opposition. 🗨ES #VOXGenuineOpposition" (VOX, 2019).

From this ad it can be deduced that the PP, the PSOE and Ciudadanos had all voted against VOX to "impose Catalan on the Balearic Islands" (Loureiro, Muro, & Alonso, 2019). It was indeed a block of amendments proposed by VOX which, according to the law, had to be voted on as a whole. In this block, VOX included some amendments that were acceptable to ideologically like-minded parties (as is the case of language policy) and others that were totally unacceptable to them. This placed those parties in a dilemma, for if they rejected the block of amendments, VOX could claim that they were in cahoots with the left, whereas if they accepted them, the left could contend that they were indistinguishable from VOX. It is something that this party does very often because it knows that the news coverage will, in one way or another, be to its advantage (Loureiro et al., 2019).

Another ad recommended a video of an interview with Santiago Abascal, the leader of VOX, in which he referred to an unaccompanied foreign minor accused of sexually abusing a 10-year-old girl in Calella (Barcelona). Abascal, among other things, stated the following: "An unaccompanied foreign minor sexually abused a

10-year-old girl. The abuser is now at liberty. With a restraining order not to approach the victim, he will be able to approach others, we imagine" (VOX, 2019). From this statement it can be inferred that the accused had already been sentenced when, in reality, he had been sent to a juvenile facility, under the restraining order not to approach the victim, pending trial ("Detenido un menor," 2019).

Likewise, when viewed from a thematic perspective (see Table 6 and Figure 2) these data reveal the focus of these disinformation messages. In all the categories, the topic of Spain's cohesion, referring to the Catalan independence issue and the glorification of the country, stands out. The subject of democratic quality, relating to corruption, was the second most frequent theme with respect to the dissemination of disinformation.

The most used topics to denounce disinformation were democratic quality and party promotion, normally associated with a recommended information diet, which tended to be the party's own channel.

Concerning the core topics of the ads, it should be stressed that 13 VIAs posted by Ciudadanos about Spain's cohesion or democratic quality corresponded to 228 and 49, respectively, of the total number of ads making up the corpus. Ciudadanos was the party that aired these topics most, while the PP focused on economic policy (two VIAs out of seven), Podemos referred most to feminism (one VIA out of three) and VOX put the accent on 'other subjects' (one VIA out of two). As can be seen in Table 6, Ciudadanos posted most disinformation (1.55% of 1.99%) in the corpus of ads.

5. Discussion

This aim of this study had been to analyse the use of disinformation by the leading Spanish political parties in the
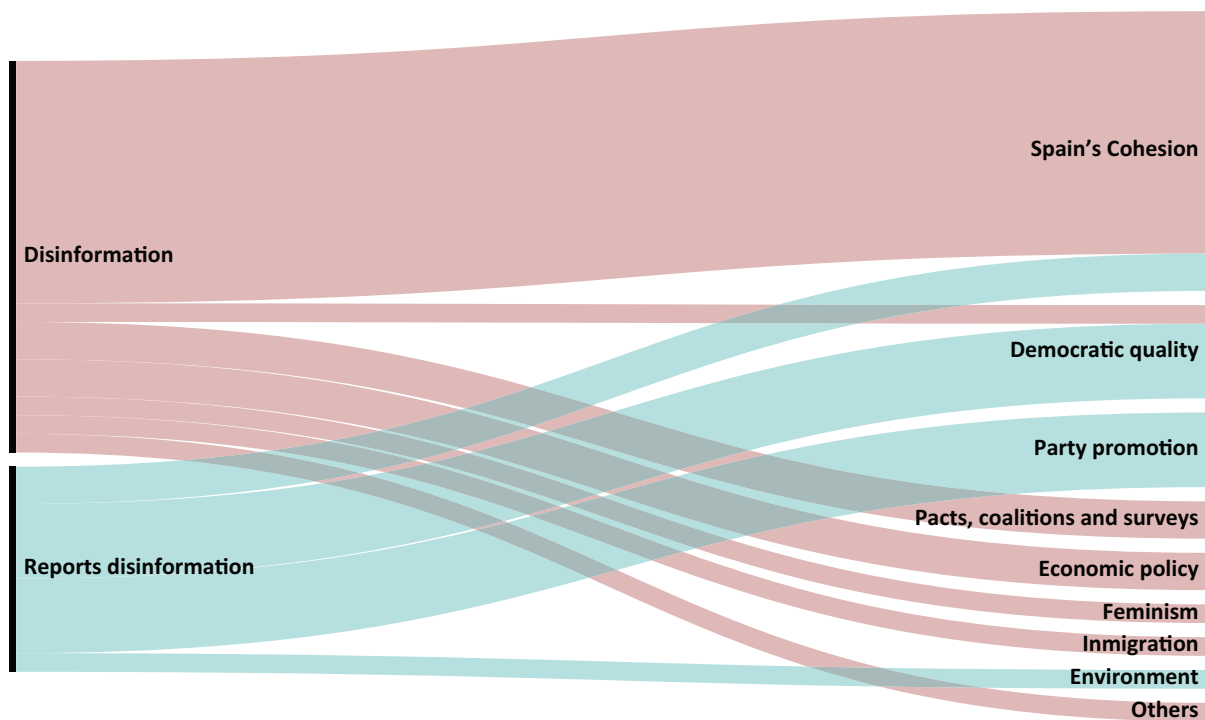
Figure 2. Reporting and use of disinformation by topic.

Table 6. Use of disinformation by political party and by topic (%).

Topic	Ciudadanos		PP		Podemos		VOX		Total
	VIAs	Total	VIAs	Total	VIAs	Total	VIAs	Total	
Spain's cohesion	5.20	1.55	0.00	0.00	0.00	0.00	0.00	0.00	1.55
Pacts and surveys	0.00	0.00	0.25	0.01	0.00	0.00	0.00	0.00	0.01
Others	0.00	0.00	0.00	0.00	0.00	0.00	6.67	0.01	0.01
Economic policy	0.00	0.00	0.25	0.05	0.00	0.00	0.00	0.00	0.05
Feminism	0.00	0.00	0.00	0.00	0.35	0.02	0.00	0.00	0.02
Democratic quality	0.40	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.33
Immigration	0.00	0.00	0.00	0.00	0.00	0.00	6.67	0.01	0.01
Total	5.60	1.89	0.51	0.06	0.35	0.02	13.33	0.02	1.99

two 2019 general election campaigns. The corpus was extracted from the Facebook Ad Library, which was not available before these elections were held. Therefore, this comprehensive and novel corpus contributes to pave the way for future research on political communication and related computational phenomena (Tufekci, 2015; Woolley & Howard, 2018).

When performing our analysis on the election campaign ads on Facebook, we were interested in confirming whether or not the Spanish political parties leveraged disinformation tactics identified in previous political processes at an international level (Marwick & Lewis, 2017) and specifically in targeted advertising on Facebook (Bakir & McStay, 2018; Gray et al., 2020), to persuade particular sectors of the electorate.

However, the results of our analysis show that the parties' approach was much more conventional (López García, 2016), as it generally copied that of traditional campaigns, focusing on electoral proposals and promises, in which in most cases the bias of the facts narrated did not tend to qualify as falsehood (thus respecting the rules of democracy). Our findings reveal that the Facebook Ad Library per se is not enough to audit disinformation campaigns on social networking sites and that greater efforts should be made in this respect at various levels (Tucker et al., 2018). As a matter of fact, it could be claimed that our research shows that the open and transparent nature of the Facebook Ad Library may be a factor that explains the relatively low percentage of ads in our corpus that included elements of disinformation, because such practices could easily be detected.

The presence of disinformation in the parties' political advertising messages was rather scarce, albeit not negligible. Although they certainly did not put Facebook's capabilities to a widespread use to discriminate, segment or intoxicate the public debate with false or misleading information. Even though we did not perform an ad segmentation analysis (which was beyond the scope of our study), we have shown that in the context of the 2019 general election campaigns disinformation revolved around a highly polarised issue in Spain, as previously identified in similar studies (Stella et al., 2018): The country's cohesion or, more generally, the autonomy of each one of its regions.

As can be deduced from the specific analysis of each party, we have also revealed that there were significant differences between them in terms of their use of advertising. We could establish a dividing line between the traditional parties, in which we have not detected disinformation (PSOE and IU) or to a minimum extent (PP), and the new parties that have emerged as a result of the crisis of the two-party system linked to the recession that began in 2008 (Ciudadanos, Podemos, and VOX).

The discursive approaches of the latter were very different from those of the former. However, at different levels and from various approaches they revealed the two elements of disinformation that we have dealt with here: the political party as a victim and agent of disinformation. As can be seen from previous election campaigns, it seems that the 'traditional' parties are taking longer to abandon the classic forms of campaigning than their 'new' counterparts (López García, 2016), which shows that, rather than being reduced to the extreme right of the ideological spectrum, there is a large variety of political actors who resort to disinformation tactics (Bennett & Livingston, 2018).

Finally, our results yet again highlight the difficulties in making a distinction between advertising, propaganda and disinformation. The ads analysed here were similar in that they were publications appearing in the Facebook timeline, which were labelled as advertising (Tandoc et al., 2018) and promoted by a specific political organisation to persuade the electorate (Molina et al., 2019). The partisan rhetoric of these ads sometimes included elements of disinformation and incorrect data to win votes. Although they should be understood as being integrated in a complex system (Marwick & Lewis, 2017) in which misleading information reflects different discourses, ad and distribution strategies and authors. The phenomenon of disinformation is inextricably linked to the analysis of the messages of networked politicians (Marwick & Lewis, 2017), even when they take the shape of ads that are freely accessible to the public on the platform attempting to audit them.

At a theoretical level, we concur with those authors who stress the epistemological difference between disinformation and fake news (Bennett & Livingston, 2018; Ireton & Posetti, 2018). The centrality of fake news in the analysis of political processes may lead to an oversimplification of content deliberately intended to spread disinformation and the actors disseminating it. Our empirical findings show that disinformation is not always produced and distributed by the media and is indeed a propaganda tool (Benkler et al., 2018; Martin, 1982). Internet technologies have provided a new setting for understanding the tactics implemented by political parties to influence public opinion, which not only include ads containing false information, but also bots, polarised groups on social networking sites, media manipulation and even the regular use of social media by political leaders. Future research should approach these phenomena holistically so as to gain a better understanding of disinformation strategies and a more accurate perspective of the links between disinformation and propaganda in contemporary politics. For example, it would be interesting to monitor not only the agents—in this case, political parties—but also specific keywords or topics that are shaping a more complex information landscape.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

References

Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, *6*(2), 154–175. https://doi.org/10.1080/21670811.2017.1345645

Bastos, M., & Mercea, D. (2018). The public accountability of social platforms: Lessons from a study on bots and trolls in the Brexit campaign. *Philosophical Transactions of Royal Society*, *376*(2128), 1–12. https://doi.org/10.1098/rsta.2018.0003

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford: Oxford University Press.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Bradshaw, S., & Howard, P. N. (2019). *The global disinformation order: 2019 global inventory of organised social media manipulation*. Oxford: Project on Computational Propaganda.

Cervi, L., & Carrillo-Andrade, A. (2019). Post-truth and disinformation: Using discourse analysis to understand the creation of emotional and rival narratives in Brexit. *ComHumanitas: Revista Científica de Comunicación*, *10*(2), 125–149. https://doi.org/10.31207/rch.v10i2.207

Ciudadanos. (2019). Facebook ad library. *Facebook*. Retrieved from https://www.facebook.com/ads/library/?active_status=all&ad_type=all&country=ES&q=Ciudadanos&view_all_page_id=74078667754

Detenido un menor tutelado fugado por abusar de una niña de 10 años en Calella [An escaped minor in custody was arrested for abusing a 10 year-old-girl in Calella]. (2019, November 6). *El Periódico*. Retrieved from https://www.elperiodico.com/es/sociedad/20191106/mena-abuso-nina-10-anos-calella-7716463

Dommett, K., & Power, S. (2019). The political economy of Facebook advertising: Election spending, regulation and targeting online., *Political Quarterly*, *90*(2), 257–265. https://doi.org/10.1111/1467-923X.12687

Expansión. (n.d.). Prima de riesgo de España 2020 [Risk premium in Spain, 2020]. *Expansión*. Retrieved from https://datosmacro.expansion.com/prima-riesgo/espana

Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, *22*(8). https://doi.org/10.5210/fm.v22i8.8005

Gray, J., Bounegru, L., & Venturini, T. (2020). 'Fake news' as infrastructural uncanny. *New Media & Society*, *22*(2), 317–341. https://doi.org/10.1177/1461444819856912

Hern, A. (2018, November 29). Facebook to require proof that political ads come from UK. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2018/nov/29/facebook-to-require-proof-that-political-ads-come-from-uk

Ireton, C., & Posetti, J. (2018). *Journalism, 'fake news' & disinformation: Handbook for journalism education and training*. Paris: Unesco Publishing.

Jack, C. (2017). *Lexicon of lies: Terms for problematic information*. New York, NY: Data & Society Research Institute.

Kim, Y. M., Hsu, J., Neiman, D., Kou, C., Bankston, L., Kim, S. Y., . . . Raskutti, G. (2018). The stealth media? Groups and targets behind divisive issue campaigns on Facebook. *Political Communication*, *35*(4), 515–541. https://doi.org/10.1080/10584609.2018.1476425

Kreiss, D., & McGregor, S. C. (2018). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 US presidential cycle. *Political Communication*, *35*(2), 155–177. https://doi.org/10.1080/10584609.2017.1364814

La publicidad de Ciudadanos en Facebook: No hay datos que avalen que están "a 190 votos" de quitarle un escaño a podemos en Andalucía [Ciudadanos Facebook ads: No data support that they are "190 votes" away from winning a seat at the expense of Podemos in Andalusia]. (2019, April 26). *Maldita*. Retrieved from https://maldita.es/malditodato/2019/04/26/la-publicidad-de-ciudadanos-en-facebook-no-hay-datos-que-avalen-que-estan-a-190-votos-de-quitarle-un-escano-a-podemos-en-andalucia

López García, G. (2016). New' vs 'old' leaderships: The campaign of Spanish general elections 2015 on Twitter. *Communication & Society*, *29*(3), 149–168. https://doi.org/10.15581/003.29.3.149-168

Loureiro, M., Muro, M., & Alonso, M. (2019, October 8). PP y CS no apoyan la propuesta de Vox para tumbar la inmersión lingüística en Baleares [PP and CS do not support Vox's proposal to veto language immersion in the Balearic Islands]. *Libertad Digital*. Retrieved from https://www.libertaddigital.com/espana/politica/2019-10-08/vox-jorge-campos-lengua-parlamento-balear-espanol-catalan-1276645930

Lukito, J. (2020). Coordinating a multi-platform disinfor-

mation campaign: Internet Research Agency activity on three U.S. social media platforms, 2015 to 2017. *Political Communication*, *37*(2), 238–255. https://doi.org/10.1080/10584609.2019.1661889

Magallón-Rosa, R. (2019). *Unfaking news: Cómo combatir la desinformación* [Unfaking news: How to combat disinformation]. Madrid: Ediciones Pirámide.

Martin, J. (1982). Disinformation: An instrumentality in the propaganda arsenal. *Political Communication*, *2*(1), 47–64. https://doi.org/10.1080/10584609.1982.9962747

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. New York, NY: Data & Society Research Institute.

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2019). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*. https://doi.org/10.1177/0002764219878224

Nielsen, R. K., & Graves, L. (2017). *"News you don't believe": Audience perspectives on fake news*. Oxford: University of Oxford.

Paniagua Rojano, F., Seoane Pérez, F., & Magallón-Rosa, R. (2020). Anatomía del bulo electoral: La desinformación política durante la campaña del 28-A en España [An anatomy of the electoral hoax: Political disinformation in Spain's 2019 general election campaign]. *Revista CIDOB d'Afers Internacionals*, *124*, 123–145. https://doi.org/10.24241/rcai.2020.124.1.123

Partido Socialista. (2019). Facebook ad library. *Facebook*. Retrieved from https://www.facebook.com/ads/library/?active_status=all&ad_type=all&country=ES&q=PSOE&view_all_page_id=189318235003

Podemos. (2019). Facebook ad library. *Facebook*. Retrieved from https://www.facebook.com/ads/library/?active_status=all&ad_type=all&country=ES&q=Podemos&view_all_page_id=269212336568846

Ribeiro, F. N., Saha, K., Babaei, M., Henrique, L., Messias, J., Benevenuto, F., . . . Redmiles, E. M. (2019). On microtargeting socially divisive ads: A case study of russia-linked ad campaigns on Facebook. In *Proceedings of the conference on fairness, accountability, and transparency* (pp. 140–149). New York, NY: Association for Computing Machinery.

Ricard, J., & Medeiros, J. (2020). Using misinformation as a political weapon: Covid-19 and Bolsonaro in Brazil. *The Harvard Kennedy School Misinformation Review*, *1*(2), 1–6. https://doi.org/10.37016/mr-2020-013

Rodríguez-Fernández, L. (2020). Junk news y "medios de desinformación" en la campaña electoral del 10-N [Junk news and "disinformation media" in the 10-N election campaign]. *Revista Mediterránea de Comunicación*, *11*(2). https://doi.org/10.14198/MEDCOM2020.11.2.19

Ross, A. S., & Rivers, D. J. (2018). Discursive deflection: Accusation of "fake news" and the spread of mis- and disinformation in the tweets of President Trump. *Social Media+Society*, *4*(2), 1–12.

Sarabia, D. (2020). Ciudadanos engañó a los votantes con campañas pagadas en Facebook para llamar al voto con datos falsos [Ciudadanos tricked voters with paid campaigns on Facebook to encourage them to cast their ballot with false information]. *El Diario*. Retrieved from https://www.eldiario.es/tecnologia/Ciudadanos-votantes-campanas-pagadas-Facebook_0_893811190.html

Sinclair, J. (2016). Advertising and media in the age of the algorithm. *International Journal of Communication*, *10*, 3522–3535.

Stella, M., Ferrara, E., & De Domenico, M. (2018). Bots sustain and inflate striking opposition in online social systems. *PNAS*, *115*(49), 12435–12440. https://doi.org/10.1073/pnas.1803470115

Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2018). Defining "fake news": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tucker, J. A., Theocharis, Y., Roberts, M. E., & Barberá, P. (2018). From liberation to turmoil: Social media and democracy. *Journal of Democracy*, *28*(4), 46–59. https://doi.org/10.1353/jod.2017.0064

Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Journal on Telecommunications & High Technology*, *13*, 203–218.

VOX. (2019). Facebook ad library. *Facebook*. Retrieved from https://www.facebook.com/ads/library/?active_status=all&ad_type=all&country=ES&q=VOX%20Espa%C3%B1a&view_all_page_id=467127060059387

Woolley, S. C., & Howard, P. N. (2016). Political communication, computational propaganda, and autonomous agents: Introduction. *International Journal of Communication*, *10*, 4882–4890. https://doi.org/1932-8036/20160005

Woolley, S. C., & Howard, P. N. (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford: Oxford University Press.

About the Authors

Lorena Cano-Orón is a Postdoctoral Fellow at the University of Valencia. She holds a PhD in Communication and Interculturality with an International Mention (University of Valencia, 2019). She is a Member of the research groups Mediaflows and ScienceFlows. Her lines of research are oriented towards the study of content flows in social networks—in the field of health and political communication—as well as the study of the effects of new technologies from a technopolitical perspective.

Dafne Calvo is a PhD Candidate in Communication at the University of Valladolid. She has visited research centres in the United States (Northeastern University), Uruguay (University of the Republic) and Mexico (Benemérita Autónoma University of Puebla). Since 2014, she has been a Member of the working team in different competitive research projects at a national level. Her doctoral thesis addresses digital political participation and the free culture movement in Spain.

Guillermo López García holds a PhD in Audiovisual Communication (2002) from the University of Valencia, where he has worked as an Associate Professor of Journalism since 2008. Most of his research focuses on political and online communication. He coordinates the R&D Group Mediaflows, a project focusing on the analysis of communication flows among political parties, the media and citizens in processes of political mobilization. He has authored or edited 12 books and numerous papers and book chapters in scientific publications.

Tomás Baviera is a Senior Lecturer in Marketing at the Universitat Politècnica de València. He holds an MSc in Telecommunication Engineering from the same university, a MA degree in Applied Social Research Methods from the Universitat Autònoma de Barcelona, as well as a PhD in Communication from the Universitat de València. He is a Researcher in the Mediaflows research group, which focuses on political communication during election campaigns. His research interests include digital communication, online marketing and organisational behaviour.

Article

Gender Differences in Tackling Fake News: Different Degrees of Concern, but Same Problems

Ester Almenar *, Sue Aran-Ramspott, Jaume Suau and Pere Masip

Blanquerna School of Communication and International Relations, Ramon Llull University, 08001 Barcelona, Spain;
E-Mails: esteraat1@blanquerna.url.edu (E.A.), suear@blanquerna.url.edu (S.A.-R), jaumesm@blanquerna.url.edu (J.S.),
peremm@blanquerna.url.edu (P.M.)

* Corresponding author

Abstract

In the current media ecosystem, in which the traditional media coexists with new players who are able to produce information and spread it widely, there is growing concern about the increasing prominence of fake news. Despite some significant efforts to determine the effects of misinformation, the results are so far inconclusive. Previous research has sought to analyze how the public perceive the effects of disinformation. This article is set in this context, and its main objective is to investigate users' perception of fake news, as well as identify the criteria on which their recognition strategies are based. The research pays particular attention to determining whether there are gender differences in the concern about the effects of fake news, the degree of difficulty in detecting fake news and the most common topics it covers. The results are based on the analysis of a representative survey of the Spanish population (N = 1,001) where participants were asked about their relationship with fake news and their competence in determining the veracity of the information, and their ability to identify false content were assessed. The findings show that men and women's perception of difficulty in identifying fake news is similar, while women are more concerned than men about the pernicious effects of misinformation on society. Gender differences are also found in the topics of the false information received. A greater proportion of men receive false news on political issues, while women tend to more frequently receive fake news about celebrities.

Keywords

disinformation; fake news; gender; misinformation; perception; Spain

1. Introduction

The relationship between screens and their influence on individuals has been further complicated by the emergence of digital platforms. In the current media ecosystem, in which the traditional media coexist alongside new media and also new players who are capable of producing information and spreading it widely, there is a growing concern about the increasing prominence of fake news. Despite some significant efforts to determine the effects of disinformation, the results are inconclusive, and there are many research gaps which have still to be addressed (Tucker et al., 2018). More specifically, recent research has addressed the consequences of exposure to disinformation in different societal groups or individuals with diverse characteristics. For example, fake news has been studied in relation to political beliefs or ideology (Allcott & Gentzkow, 2017; Guess, Nagler, & Tucker, 2019), news consumption or social media use (Wagner & Boczkowski, 2019) or national feelings (Khaldarova

& Pantti, 2016). Although studies of the reception of fake news are still at an early stage (Jankowski, 2018), we have identified a clear absence of a gender-based approach to the topic. Issues of gender have been taken into account in relation to fake news but normally from a perspective based on the content or message analysis (Stabile, Grant, Purohit, & Harris, 2019) or the strategies of far-right groups in what has been defined as cultural wars (Mudde, 2019). As will be further explained in the literature review section, this article draws on reception studies on the dimension of gender differences in news consumption (Fortunati, Deuze, & de Luca, 2014; Toff & Palmer, 2019) as well as the previously identified gender attitudinal differences with regard to the use of new communication technologies (Bond, 2009; Cai, Fan, & Du, 2017; Renau, Carbonell, & Oberst, 2012). From this theoretical basis, this article aims to investigate whether there are gender differences in the users' perception of fake news. More specifically, we will focus on several issues which have already been researched in connection with fake news, but where the gender dimension has been normally disregarded. These include concerns about the spread of fake news, the degree of difficulty in detecting it and the topics received.

Although authors like Wardle and Derakhshan (2017) distinguish between misinformation, disinformation and mal-information, and it is well known that the use of the term 'fake news' is problematic, for the purpose of this article we will use the terms disinformation and fake news interchangeably, particularly in the results section. Scholars tend to avoid using the term 'fake news' because they consider that it is inadequate to describe the complex phenomena of false information (Tandoc, Lim, & Ling, 2018). However, for ordinary people, fake news is the most popular and frequently used expression to refer to the nebula of false information, viral lies, conspiracy theories, and other forms of misleading information spread on social networks and some news media (in Spanish, *notícias falsas*). Since the research is based on a survey of a representative sample of the Spanish population, we decided to use the most familiar word in order to avoid misunderstanding.

2. Reception Studies, Disinformation and the Gender Dimension

From a theoretical perspective, Audience Reception Studies could help us to investigate individual's perceptions and understanding of disinformation from a gender perspective. Reception studies began in the 1980s as a reaction to the widespread assumption that the audience was 'passive' in its media consumption, portraying audiences as easy to manipulate and homogeneous in their behavior and characteristics (Livingstone, 1998). Reception studies implied a change in viewing audiences as heterogeneous and resistant in their media consumption. Hence, audiences were then defined both by their personal characteristics as well as by their social contexts

(Mattelart & Neveu, 2003). Since the 1980s reception studies have evolved and adapted to an ever-changing media system. Nowadays the new view of the audience as 'active' is highly relevant, thanks to the possibilities of the new communication technologies, which allow greater user interactivity with media content (Banaji & Buckingham, 2013; Jenkins, 2006; Jenkins & Carpentier, 2013; Scolari, 2012).

In principle, misleading content is not considered as a gender-specific media product. However, rather than conceptualizing audiences as passive consumers, or victims of this content, we suggest that we should conceptualize them as active in that they could challenge or re-appropriate the content. In this sense, it would be relevant to analyze whether gender plays a role in how individuals receive and react to fake news. Re-appropriation must always be considered within the context of the social situation in which it takes place as well as within the routines of everyday life and personal characteristics, which are so important at the beginning of the feminist audience studies according to Cavalcante, Press, and Sender (2017). Even before the digital hybrid media system (Chadwick, 2013), audience research stresses the importance of media consumption patterns in everyday life (Bird, 2003; Morley, 1992) as well as the social and relational roles in which this consumption takes place (Boczkowski, 2010). The gender approach is very relevant in these areas of research. Previous research showed significant differences in how women and men find and consume news, as well as the topics that most interest them (Fortunati et al., 2014). It seems that women may be more interested in news and reporting which is directly related to daily life, such as the weather, health and similar topics, and less interested in 'hard' political news (Poindexter & Harp, 2008). A 'news consumption gender gap' has also been identified (Toff & Palmer, 2019) that plays an important role in shaping how women consume less news through patterns of news avoidance, news-is-for-men perceptions and structural inequalities that shape individuals' everyday media consumption habits.

It seems then, that with regard to news' consumption, there are substantial differences between genders, affecting which types of news are more or less consumed by both women and men. Previous research has addressed the issue of topics of disinformation, but disregarded the gender differences. There is broad agreement that disinformation has a strong political component, linked to partisanship and identity politics (Mourão & Robertson, 2019). Hence, a high proportion of fake news relates to such topics and narratives. However, researchers have also found differences between countries, with regard to specific national political issues, divisions or media system characteristics. For example, Humprecht (2019) found great differences in comparing the US and UK with Germany and Austria. While the former have a higher level of partisan-based disinformation content, in German-speaking countries

sensationalist stories are more prominent. Hence, disinformation mirrors national news agendas and political debates. Furthermore, contextual events could also modify which types of fake news are most frequently consumed. For example, the outbreak of Covid-19 implied that disinformation about health issues became more relevant and widespread, as recent research in Spain shows (Masip et al., 2020; Salaverría et al., 2020). Despite these interesting patterns, the gender aspect has not been fully addressed. Hence, our first research question will be to see if the previously identified 'news consumption gender gap' can be extended to fake news reception:

> RQ1: Are there gender differences in the perceptions of the most common topics of fake news that are received?

However, the 'news consumption gender gap' is not the only gender bias found by previous research. As Cai et al. (2017) point out, as regards the new communication technologies, there still exists a gender attitudinal gap as a consequence of many different factors, including the general conception that technology is a male-dominated area, that men are more competent users of technology, and other social and cultural norms and factors. This leads to many different relationships with digital news content between men and women. As Bond (2009) highlights, the two genders exhibit different motivations for engaging in social media use, a feature that can be explained through behavioural patterns of socialization (Renau et al., 2012). As Toff and Palmer (2019, p. 1565) observe: "According to the theories of socialization, gender roles and news consumption habits tend to originate in the home and are reinforced and modeled in school, among peers, and in the media." As previous research has emphasized, the approach to news media has gender differences. This goes further than the gender preferences on the volume, content and topics or the ways to find news, to the socialization processes, modeled from childhood and reinforced by, among other agents, the media system, and includes a structural inequality. This structural difference seems to be more of a historical constant than an anomaly and implies what has been defined as 'cognitive costs': If we consider that women are, in general, less educated, especially about political matters, the 'cognitive costs' of deciphering and focusing on politics might be higher (Benesch, 2012). Hence, if media consumption and use of technology are so strongly influenced by gender-related issues, we believe that it is pertinent that audience researchers pay attention to the possible gender-based media perception with regard to the criteria of trust in the digital field.

One interesting issue, therefore, is how these 'cognitive costs' might relate firstly to the capacity of individuals to be concerned about disinformation, firstly, and secondly to detecting this kind of content. Can we talk about the different cognitive costs between women and men regarding the capacity to detect fake news?

Does the lack of interest in political news make women less likely to be worried about the spread of misleading content? Existing research on disinformation has not fully addressed this a gender approach, with just a few exceptions (Ştefăniţă, Corbu, & Buturoiu, 2018). When researching fake news, it has been found that individuals do not clearly distinguish such content from news, although the results are better if they need to identify propaganda or advertising (Nielsen & Graves, 2017). It seems that for many individuals, fake news might be synonymous with 'bad journalism.' However, as comparative results from the Digital News Report show (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018), individuals in many different countries (US, Spain, Greece, Brazil, etc.) do recognize that they are frequently exposed to disinformation although their interpretation of what exactly this means might differ from academics definition of it. With regard to individuals' capacity to detect disinformation, previous research has been based mostly on cognitive perspectives in order to attempt to distinguish which individuals are more likely to believe fake news or to share this sort of content (Pennycook & Rand, 2018; Schulz, Werner, & Müller, 2018). Although results differ greatly depending on the country concerned (Benkler, Faris, & Roberts, 2018; Humprecht, Esser, & Van Aelst, 2020), there is a common ground in most of them, defined in previous studies as the 'third person effect' (Jang & Kim, 2018). This means the widespread belief that 'other people' are the ones who are deceived by fake news rather than oneself or those in one's immediate circle. Age or digital literacy skills are also often cited as predictors of how well or badly individuals might be able to spot fake news, with research often giving different or contradictory results (Jones-Jang, Mortensen, & Liu, 2019; Pennycook & Rand, 2018; Wagner & Boczkowski, 2019).

Taking into consideration the findings of previous research with regard to disinformation as well as the current gap in addressing the gender approach, we will frame our second and third research questions as follows:

> RQ2: Are there gender differences in concern about the spread of fake news?

> RQ3: Are there gender differences in the perceived degree of difficulty in detecting fake news?

3. Methodology

The design of this research is based on a national survey on a sample of N = 1,001 Spaniards over 18 years of age. The data were collected by a market research firm (Gesop) through a questionnaire administered online. The sample consisted of 1,001 completed questionnaires, with the sampling stratified by age, sex, autonomous community (Spain is divided into 17 autonomous communities) based on the actual

distribution of the Spanish population. The margin of sampling error is ± 3.2 with a 95% level of confidence and p = q = 0.5. The fieldwork was carried out from 24–30 April 2019. Table 1 shows the demographics of the participants.

To carry out the research, a broad questionnaire was prepared with five main blocks: sociodemographic data, media consumption, trust in the media, use of social networks, disinformation and social networks. The questions were of two types: with multiple choice single answer or multiple choice. Questions based on a Likert scale (1–7, with 1 being none and 7 a lot) were used in the questions asking participants for their assessment of the proliferation of false news.

For the present investigation, the answers obtained in the following questions have been taken into account: i) Are you concerned about the spread of fake news? ii) Is it difficult for you to identify fake news? iii) How often do you check content you think could be false? iv) What factors do you take into account to assess the reliability of content received on social networks? and v) Which are the most frequent topics of fake news you receive? In addition, we take into account the gender of the respondents and their self-reported use of social networking sites.

The results obtained in these questions were analyzed with the student's test statistic (bilateral), T-test, for independent samples, Chi Square tests were used to verify the existence of dependency relationships between variables. With regard to dependency, the standardized residuals were analyzed. Finally, in those questions with multiple answers, the results were analyzed by comparing proportions for independent samples. The statistic for the significance tests used was the Z (bilateral), in pairwise comparisons. In all cases, the 95% confidence interval level, which is used in the social sciences, has been used.

4. Results

The results obtained show that Spanish people are very concerned about the proliferation of fake news. On a scale of 1 to 7, with 1 being not at all concerned and 7 being very concerned, the mean obtained is 6.12 ($\bar{x} = 6.12$, $\sigma = 1.3$). Women ($\bar{x} = 6.21$, $\sigma = 1.23$) are significantly more concerned than men ($\bar{x} = 6.03$, $\sigma = 1.36$) regarding fake news dissemination (see Table 2).

As Table 3 shows, differences in gender with regard to concern about fake news dissemination does not occur among the youngest, ($\bar{x} = 6.06$, $\sigma = 1.36$ in men Vs $\bar{x} = 6.10$, $\sigma = 1.14$ among women), nor among the older age group ($\bar{x} = 6.18$, $\sigma = 1.25$ in men Vs $\bar{x} = 6.21$, $\sigma = 1.47$ among women). However, it is observed in the 30–59 age range. Despite not being statistically significant it is interesting to see that young men tend to show greater concern than men between 30 and 59. Conversely, among women, concern shows the opposite trend, increasing when participants reach the 30–59 age range.

4.1. Social Media and Concern about Fake News

Regarding the analysis of concern about fake news, respondents were divided into two groups, depending on their responses on the Likert scale from 1 to 7. We considered as 'worried people' those who responded 6 or 7 on the Likert scale, and the unconcerned as those who rated it between 1 and 5.

The role of social networks as disseminators of disinformation could suggest a certain relationship between their use and individuals' level of concern about fake news. However, it has been observed that the intensity of use of social networks is not correlated with concern about fake news. Only the frequent use of Facebook χ^2 (1, N = 791) = 5.25, p < .05 is associated with a high degree of concern. This correlation is not observed

Table 1. Demographics.

	Response	N (%)
Gender	Male	498 (49.7%)
	Female	503 (50.3%)
	Total	1,001 (100%)
Age	16–29	169 (16.9%)
	30–44	310 (30.9%)
	45–59	311 (31.0%)
	≥60	211 (21.1%)
	Total	1,001 (100%)

Table 2. Concern about fake news dissemination.

	1: Not at all	2	3	4	5	6	7: Extremely	Mean	Standard deviation
Total	2.50%	0.80%	1.90%	3.70%	11.39%	26.87%	52.55%	6.12	1.3
Male	3.01%	0.80%	2.01%	4.42%	12.65%	27.51%	49.60%	6.03	1.36
Female	1.99%	0.80%	1.79%	2.98%	10.14%	26.24%	55.47%	6.21	1.23

Table 3. Concern about fake news dissemination by gender and age.

	19–29		30–59		60–74	
	M	F	M	F	M	F
Mean	6.06	6.10	5.96	6.26	6.18	6.21
Standard deviation	1.36	1.14	1.42	1.15	1.25	1.47
T test	0.816		0.008		0.856	

Notes: p = .05.

on Twitter and Instagram. The residue analysis confirms the positive correlation between concern and daily use of Facebook.

On analyzing the results obtained from individuals who show higher levels of concern, we can observe the dependency between gender and use of social networks. Women who worry about fake news use Facebook more frequently (χ^2 (1, N = 630) = 13.75, p < .05) and Instagram (χ^2 (1, N = 424) = 4.37, p < .05) than worried men; the men who express greater concern about disinformation use Twitter more intensively (χ^2 (1, N = 334) = 8.77, p < .05). The analysis of the standardized residuals confirms the results: Facebook: (z score = 3.708) and Instagram (z score = 2.091) are used in a greater proportion by women than men, who make more use of Twitter (z score = 2.962). Among the respondents who stated that they were not concerned, a correlation was detected between daily use and gender in the case of Facebook χ^2 (1, N = 161) = 4.59, p < .05; but not on Instagram or Twitter. The specific analysis of the residuals shows that women who are not concerned about fake news make more daily use of Facebook than men who are not concerned (z score = 2.14).

Despite these results, it is risky to establish a causal relationship between the two variables, since Facebook χ^2 (1, N = 791) = 19.53, p < .05 and Instagram χ^2 (1, N = 590) = 5.66, p < .05) are the social media sites preferred by women, and Twitter (χ^2 (1, N = 419) = 7.10, p < .05) is the most popular social network among men.

In conclusion, the intensive use of one or other social network by those who are concerned seems to relate to factors other than fake news, since it reflects the normal behavior in consumption of social media at the gender level.

4.2. Detecting Fake News

A majority of respondents (67.13%) admit to having difficulties in detecting fake news. However, no significant differences can be established between genders (x^2 (2, N = 1,001) = 1.713, p .05).

Nevertheless, the analysis of the difficulty in identifying fake news in relation to the concern for the spreading of fake news allows us to verify the existence of a correlation between the two variables. 76.8% of concerned citizens report having difficulty in detecting fake news, which drops to 63.58% among the least worried about its dissemination. Therefore, there is a relationship

between concern about the proliferation of fake news and the degree of difficulty in detecting it (χ^2 (2, N = 998) = 14.25, p < .05). We can also view it in reverse: Those less concerned about fake news are the same people who claim to have fewer difficulties in detecting it.

At the gender level, in line with previous results, no behavioral differences between men and women were detected. Differences by gender are not statistically significant when it comes to identifying fake news based on individuals' concern about it. Women are more worried than men, but both exhibit the same difficulty in detecting fake news.

The study of trust and credibility is complex and therefore, outside the scope of this article. However, it is known that the tendency to consider a certain content true or false is conditioned by what is accepted as true (Williams, 2002). This is susceptible to change over time and is conditioned by various antecedents and factors such as the characteristics of the message, the sender or the topics (Fulmer & Gelfand, 2012).

Within this context, research participants were asked what factors they take into account to assess the trustworthiness of the news received through social networks. As it can be seen in Table 4, the source is the most important factor for both men and women, although it is more important for men (z score = 2,073).

If the degree of concern is taken into account, the results follow a similar pattern. The most concerned participants attach greater importance to the source of the content—the author—(67.3%) than the less-concerned people (53.7%), (z score = 3.6154). This behavior is repeated in the analysis at the gender level (men concerned vs. less concerned and women concerned vs. less concerned). Among those who are most concerned, there are no differences between men and women.

Greater concern about the spread of fake news could be related to an increased need for verification. Analysis of the data confirms this point. There is a correlation between the degree of concern about fake news and the frequency of checking it: χ^2 (3, N = 771) = 11.94, p < .05. The most concerned check more frequently, and the least concerned 'almost never.'

Around 42% of those who are more concerned frequently check (z score = 2.13), compared to 32.7% of those less concerned, which results in statistically different values (there is no overlap in the confidence intervals). Those who are least concerned reported almost never checking (20.1%), a statistically higher value than

Table 4. Factors used to assess trustworthiness of news.

	M	F	Total	M (%)	F (%)	Total (%)**
Sharer (person who shared the content	251	270	521	50.40	54.00	52.20
Source (author of the content)	337	307	644	67.67*	61.40	64.53
Topic	175	191	366	35.14	38.20	36.67
Ideological affinity with the content	96	90	186	19.28	18.00	18.64
Format	154	149	303	30.92	29.80	30.36
Total	498	500	998			

Notes: * = Statistical difference by gender at p = .05; ** = multi answer question.

10.6% of those concerned (z score = −3.22). This behavior does not vary by gender.

4.3. Topics and Fake News

Lastly, an analysis was carried out on whether there were gender differences with regard to the fake news topics that the interviewees considered that they have received. Politics is by far the most frequent topic, both among men and women. Statistically significant differences are observed as men receive fake news about politics more frequently (z score = 2.02). There are no differences regarding the other subjects, except in the case of sport, which is much more frequent among men than among women (z score = 1.96). Also, there are no differences in the topics received in relation to the degree of concern (see Table 5).

Respondents were asked to cite up to three topics on which they believe they receive fake news most frequently (Table 6). When analyzing the three responses, politics remains the main topic, although does not show any differences between genders. In contrast, nationalism emerges as the second most common topic among men (42.0%), a statistically higher percentage than women (z score = 3.33). Politics, nationalism and immigration are the most common topics for perceived fake news. At a lower level of importance, news about celebrities is also mentioned (29.4% of women vs. 22.1% of men, z score = −2.64) and sport, significantly more frequently among men than women (9.4% vs. 3.4%, z score = 3.89). It was also observed that the degree of

concern does not have any relationship with the topic of fake news received.

5. Conclusion and Discussion

The main aim of this article is to contribute to the limited existing literature on gender perceptions of fake news. After reviewing the available research on reception, our starting point was to assume that individuals are also 'active' in resisting fake news. Our intention was to analyze whether identified gender differences in news reception, as well as the cognitive costs associated with gender issues in relation to news (Toff & Palmer, 2019) might have some translation into individuals' perceptions of fake news, and more specifically, with regard to the spread of fake news, the degree of difficulty in detecting it and topics of fake news received.

Our research shows that contrary to what one may expect, there are few differences between genders with regard to disinformation. While in other related matters such as social media use or news consumption, gender differences have been clearly identified, the differences are very subtle on disinformation. The main point is that women are more concerned than men regarding the spread of disinformation which is similar to what happens in other aspects of daily life (i.e., Xiao & McCright, 2012). However, women and men have similar problems in detecting false content, they use similar factors in assessing trustworthiness, and they receive misleading material about the same topics, mainly politics.

Table 5. Topics and fake news (first response).

	M	F	Total	M (%)	F (%)	Total (%)
Politics	197	167	364	39.56*	33.53	36.47
Immigration	46	64	110	9.24	12.85	11.02
Nationalism	44	39	83	8.84	7.83	8.32
Science and Technology	7	10	17	1.41	2.01	1.70
Feminism/Gender equality	28	25	53	5.62	5.02	5.31
Society**/Crime report	46	44	90	9.24	8.84	9.02
Sports	10	3	13	2.01*	0.60	1.30
Celebrities	31	46	77	6.22	9.24	7.72
DK/NO	89	102	191	17.87	20.48	19.14
Total	498	500	998	100.00	100.40	100.00

Notes: * = Statistical difference by gender at p = .05; ** = includes health, education, and the local news.

Table 6. Topics and fake news (3 responses).

	M	F	Total	M (%)	F (%)	Total (%)**
Politics	334	316	650	67.1	63.2	65.1
Immigration	182	201	383	36.5	40.2	38.4
Nationalism	209	159	368	42.0*	31.8	36.9
Science and Tech.	38	33	71	7.6	6.6	7.1
Feminism/Gender equality	139	144	283	27.9	28.8	28.4
Society/Crime report	162	162	324	32.5	32.4	32.5
Sports	47	17	64	9.4*	3.4	6.4
Celebrities	110	147	257	22.1	29.4*	25.8
DK/NO	94	114	208	18.9	22.8	20.8
Total responses	1,221	1,179	2,400			
Total sample	498	500	998			

Notes: * = Statistical difference by gender at $p = .05$; ** = multi answer question.

RQ2 of this research aimed to determine the degree of concern regarding disinformation in Spaniards and whether there are gender differences. In this regard and in line with other research (Eurobarometer, 2018), we can affirm that fake news has become a concern for the Spanish people, women more so than men. It is interesting to observe that the degree of concern does not vary according to gender for young people, which, as a hypothesis and to be explored in future research, could be linked to a higher educational level than previous generations.

The spread of disinformation is closely related to social media use (Tandoc, Lim, & Ling, 2020), and its use is also related to the degree of concern about the spread of fake news. Our research reveals that Facebook is largely perceived as the leading distributor of fake news both by men and women. Only among the group of the most concerned are any gender-related nuances observed. While women claim that they receive fake news through Facebook and Instagram, men believe it reaches them predominantly through Twitter. However, a causal relationship cannot be established, since various studies clearly show that social networks in Spain show gender differences in user distribution. While Facebook and Instagram are more female, Twitter is largely used by men (We are social, 2020).

In the survey, participants were also asked if they had difficulty distinguishing between false and true information (RQ3). This is a complex issue to address insomuch as it depends on individuals' self-perception as being able to identify false news. It is interesting to note how those who are most concerned about fake news also perceive the greatest difficulty in detecting it, which could suggest that greater awareness of the problem leads to a greater perception of difficulty of detection. Accordingly, those with less concern see themselves as having less difficulty in detecting it. Previous research has shown that individuals have a tendency to consider that the media, and also disinformation, have a greater effect on others than on themselves, in what previous research has identified as the 'third person effect' (Jang & Kim, 2018). Interestingly, our research shows that most respondents acknowledge

that they have problems distinguishing between false and true content, and there are no significant gender differences in this regard either. We do not disregard the 'third person effect,' but we would like to emphasize that, although individuals may believe others are easier to trick, to some extent all of them perceive themselves to be vulnerable to fake news. This issue should be addressed in further research.

As it has been seen, the credibility attached to a particular content is conditioned by multiple factors, relating to the message, to the psychological traits of individuals, or even to the source of the content, which will be subject to change over time and according to the context. Our findings show that these factors apply to both men and women in a very similar way. The only significant gender difference relates to the source (the author of the content), which is more relevant for men. In a classic communication context, this result could be linked to the role that journalism/the media plays in the construction of reality, as well as the fact that men in Spain are more extensive users of news outlets than women (AIMC, 2020). However, in the current hybrid media system, the role of the source goes far beyond the traditional media, which share the stage with new actors, including possible creators and disseminators of false content such as political parties, governments, defenders of conspiracy theorists, etc.

This research allows us to identify two patterns of behavior that go beyond gender: i) Concerned and active users, who are genuinely concerned about fake news, are more aware of the difficulty in detecting it and therefore, make a greater effort to check the veracity of the news they receive; and ii) confident and passive users, who feel less concerned about false news, view is as less difficult to detect and, therefore, they verify the content less and take less account of the source. It would be interesting to verify these typologies in other countries. This correlation between concern and perception of difficulty in detecting disinformation suggest a need to focus more closely on fact-checking processes and results, and foster a questioning attitude towards news.

Lastly, our RQ1 was to determine whether there are gender differences in the most common topics of fake news that are received. This reflection originates from what previous research has named the 'news consumption gender gap' (Toff & Palmer, 2019), as it demonstrated significant differences in how women and men find and consume news (Fortunati et al., 2014; Lee, 2013): While women might be more interested in news connected with issues in daily life, men would be more likely to access political news (Poindexter & Harp, 2008; Rosentiel, 2008). Furthermore, it seems that these differences in news consumption originate from patterns of news avoidance (Toff & Palmer, 2019). Hence, we believe it was important to see whether these patterns are reproduced in fake news reception. However, the results show that politics-related fake contents are the most common, and that in line with this, perhaps unsurprisingly, politics is, for both men and women, the main topic of the false news they identify. In general, men and women receive false news on similar topics, although politics is clearly the most frequent and the only topic in which significant differences between genders are observed. As we have already seen in earlier literature (AIMC, 2020) men tend to consume more information than women and are more interested in politics. Nevertheless, this is not reflected in a greater ability of men to identify false news, where men and women have the same difficulties.

This research highlights that men, with a greater interest in politics and a greater consumption of news media, have the same difficulties as women in detecting fake news and that this news is on the same topics as women. The greater interest in politics could explain the higher percentage of politics as the first topic mentioned by men. Nationalism also comes into play when the accumulated three topics are considered, this relevance can be explained by the current Catalan-Spanish conflict.

To sum up, previous research has confirmed that men and women exhibit different behavior in relation to news media use, have different interests with regard to news topics, and different social media use. However, the results of this research show that there are no significant gender differences with regard to the ability to detect fake news. Women are more concerned than men, but both have the same problems when facing disinformation content, which has become a widespread global phenomenon today.

Acknowledgments

Conflict of Interests

References

AIMC. (2020). *Estudio general de medios: Marco general 2020* [General media study: General framework 2020]. Madrid: Asociación para la Investigación de Medios de Comunicación. Retrieved from www.aimc.es/a1mc-c0nt3nt/uploads/2020/01/marco2020.pdf

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236.

Banaji, S., & Buckingham, D. (2013). *The civic web*. Cambridge, MA: MIT Press.

Benesch, C. (2012). An empirical analysis of the gender gap in news consumption. *Journal of Media Economics*, *25*(3), 147–167.

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford: Oxford University Press.

Bird, S. E. (2003). *The audience in everyday life: Living in a media world*. New York, NY: Routledge.

Boczkowski, P. J. (2010). *News at work*. Chicago, IL: The University of Chicago Press.

Bond, B. J. (2009). He posted, she posted: Gender differences in self-disclosure on social network sites. *Rocky Mountain Communication Review*, *6*(2), 29–37.

Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, *105*, 1–13.

Cavalcante, A., Press, A., & Sender, K. (2017). Feminist reception studies in a post-audience age: Returning to audiences and everyday life. *Feminist Media Studies*, *17*(1), 1–13.

Chadwick, A. (2013). *The hybrid media system*. Oxford: Oxford University Press.

Eurobarometer. (2018). *Fake news and disinformation online* (Flash Eurobarometer No. 464). Brussels: European Commission. Retrieved from https://ec.europa.eu/commfrontoffice/publicopinion

Fortunati, L., Deuze, M., & de Luca, F. (2014). The new about news: How print, online, free, and mobile coconstruct new audiences in Italy, France, Spain, the UK, and Germany, *Journal of Computer-Mediated Communication*, *19*(2), 1–20.

Fulmer, C. A., & Gelfand, M. J. (2012). At what level (and in whom) we trust: Trust across multiple organizational levels. *Journal of Management*, *38*(4), 1167–1230.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1), 2375–2548.

Humprecht, E. (2019). Where 'fake news' flourishes: A comparison across four Western democracies. *Information, Communication & Society*, *22*(13), 1973–1988.

Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, *25*(3), 493–516.

Jang, S., & Kim, J. (2018). Third person effects of fake news: Fake news regulation and media literacy interventions. *Computer and Human Behavior*, *80*, 295–302.

Jankowski, N. W. (2018). Researching fake news: A selective examination of empirical studies. *Javnost: The Public*, *25*(1/2), 248–255.

Jenkins, H. (2006). *Convergence culture*. New York, NY: New York University Press.

Jenkins, H., & Carpentier, N. (2013). Theorizing participatory intensities: A conversation about participation and politics. *Convergence*, *19*(3), 265–286.

Jones-Jang, S. M., Mortensen, T., & Liu, J. (2019). Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American Behavioral Scientist*, *65*(2). https://doi.org/10.1177/0002764219869406

Khaldarova, I., & Pantti, M. (2016). Fake news. *Journalism Practice*, *10*(7), 891–901.

Lee, A. M. (2013). News audiences revisited: Theorizing the Link between audience motivations and news consumption. *Journal of Broadcasting & Electronic Media*, *57*(3), 300–317.

Livingstone, S. (1998). Audience research at the crossroads: The 'implied audience' in media and cultural theory. *European Journal of Cultural Studies*, *1*(2), 193–217.

Masip, P., Aran-Ramspott, S., Ruiz-Caballero, C., Suau, J., Almenar, E., & Puertas-Graell, D. (2020). Consumo informativo y cobertura mediática durante el confinamiento por el Covid-19: Sobreinformación, sesgo ideológico y sensacionalismo [News consumption and media coverage during the confinement by Covid-19: Information overload, ideological bias and sensationalism]. *El Profesional de la Información*, *29*(3).

Mattelart, A., & Neveu, E. (2003). Introducción a los estudios culturales [Introduction to cultural studies]. Barcelona: Paidós.

Morley, D. (1992). *Television, audiences and cultural studies*. London: Routledge.

Mourão, R. R., & Robertson, C. T. (2019). Fake news as discursive integration: An analysis of sites that publish false, misleading, hyperpartisan and sensational information. *Journalism Studies*, *20*(14), 2077–2095.

Mudde, C. (2019). *The far right today*. Cambridge: Polity Press.

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., & Nielsen, R. K. (2018). *Reuters Institute digital news report 2018*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf

Nielsen, R. K., & Graves, L. (2017). 'News you don't believe': Audience perspectives on fake news. *Reuters Institute*. Retrieved from https://reutersinstitute.politics.ox.ac.uk/our-research/news-you-dont-believe-audience-perspectives-fake-news

Pennycook, G., & Rand, D. G. (2018). Who falls for fake news? The roles of bullshit receptivity, over-claiming, familiarity, and analytic thinking. *Journal of Personality*. https://doi-org.proxy.library.upenn.edu/10.1111/jopy.12476

Poindexter, P., & Harp, D. (2008). The softer side of news. In P. Poindexter, S. Meraz, & A. S. Weiss (Eds.), *Women, men and news: Divided and disconnected in the news media landscape* (pp. 85–96). New York, NY: Routledge.

Renau, V., Carbonell, X., & Oberst, U. (2012). Redes sociales on-line, género y construcción del self [On-line social networks, gender and construction of the self]. *Revista de Psicologia, Ciències de l'Educació i de l'Esport*, *30*(2), 97–107.

Rosentiel, T. (2008). Where men and women differ in following the news. *Pew Research Center*. Retrieved from http://www.pewresearch.org/2008/02/06/where-men-and-women-differ-in-following-the-news

Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I., & Erviti, M. C. (2020). Desinformación en tiempos de pandemia: Tipología de los bulos sobre la Covid-19 [Disinformation in times of pandemic: Typology of hoaxes on Covid-19]. *El Profesional de la Información*, *29*(3).

Schulz, A., Werner, W., & Müller, P. (2018). We are the people and you are fake news: A social identity approach to populist individuals' false consensus and hostile media perceptions. *Communication Research*, *12*, 1–26.

Scolari, C. (2012). Media ecology: Exploring the metaphor to expand the theory. *Communication Theory*, *22*(2), 204–225.

Stabile, B., Grant, A., Purohit, H., & Harris, K. (2019). Sex, lies, and stereotypes: Gendered implications of fake news for women in politics. *Public Integrity*, *21*(5), 491–502.

Ştefăniţă, O., Corbu, N., & Buturoiu, R. (2018). Fake news and the third-person effect: They are more influenced than me and you. *Journal of Media Research*, *11*(3), 5–23.

Tandoc, E. C., Lim, D., & Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, *21*(3), 381–398.

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining 'fake news.' *Digital Journalism*, *6*(2), 137–153.

Toff, B., & Palmer, R. A. (2019). Explaining the gender gap in news avoidance: 'News-is-for-men' perceptions and the burdens of caretaking. *Journalism Studies*, *20*(11), 1563–1579.

Tucker, J. A., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., . . . Nyhan, B. (2018). *Social media,*

political polarization, and political disinformation: A review of the scientific literature. Menlo Park, CA: Hewlett Foundation. Retrieved from https://www.hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

Wagner, M. C., & Boczkowski, P. J. (2019). The reception of fake news: The interpretations and practices that shape the consumption of perceived misinformation. *Digital Journalism*, *7*(7), 870–885.

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Towards and interdisciplinary framework for research and policy making* (Report No. 09). Stras-

bourg: Council of Europe. Retrieved from https://rm.coe.int/information-disorder-toward-an-inter disciplinary-framework-for-researc/168076277c

We are social. (2020). Digital: 2020 Spain. *We are Social*. Retrieved from https://wearesocial.com/es/digital-2020-espana

Williams, B. (2002). *Truth and truthfulness*. Princeton, NJ: Princeton University Press.

Xiao, C., & McCright, A. M. (2012). Explaining gender differences in concern about environmental problems in the United States. *Society & Natural Resource*s, *25*(11), 1067–1084.

About the Authors

Ester Almenar is a Lecturer at the School of Communication and International Relations of the Ramon Llull University. She holds a PhD degree in Communication. Her areas of expertise are marketing and business strategy. She is a Member of the Digilab research group and her research interests are the impact of uncertainty on customers, news' consumption and trust on news.

Sue Aran-Ramspott is Associated Professor at the School of Communication and International Relations at the Ramon Llull University. Her areas of expertise are the relationship between public and media narratives and ethics of the audiovisual fictional representations. Currently, as a Member of Digilab, she leads a research project funded by the Spanish Ministry of Science, Innovation and Universities 'TeensMediaLife' She is the PhD Communication Program Coordinator.

Jaume Suau has a PhD in Communication and is a lecturer at Ramon Llull University, based in Barcelona. As a member of Digilab research group has been Project Manager of the MedMedia, being a member also of the Media Pluralism Monitor, all projects funded by the European Commission. His research interests are audience studies, news' consumption and trust on news, as well as media regulation and media development.

Pere Masip is Associate Professor at the School of Communication and International Relations Blanquerna at the Ramon Llull University (Barcelona, Catalonia). His research interests center on the impact of technology on journalistic and communication practices. He has participated in several national and international projects. He is currently coordinating a research project funded by the Spanish Ministry of Science, Innovation and Universities 'News Consumption, Social Networks and Pluralism in the Hybrid Media System.'

Article

Cross-Media Alliances to Stop Disinformation: A Real Solution?

Bella Palomo * and Jon Sedano

Department of Journalism, University of Malaga, 29018 Malaga, Spain; E-Mails: bellapalomo@uma.es (B.P.),
jon@uma.es (J.S.)

* Corresponding author

Abstract
Social networks have surpassed their intermediary role and become gatekeepers of online content and traffic. This trans-
formation has favored the spread of information disorders. The situation is especially alarming in Spain, where 57% of
Spaniards have at some moment believed false news. Since 2016, First Draft has promoted several collaborative verification
projects that brought together newsrooms to fact-check false, misleading and confusing claims circulating online during
presidential elections in several countries. The main objective of this article is to study the collaboration forged between
newsrooms in Spain in order to debunk disinformation contents in 2019 under the name of Comprobado (Verified) and
the impact of this initiative. Applying a methodological approach based on non-participant observation, interviews, con-
tent analysis of reports, scientific articles, books and media archives, we examine the internal uses of this platform, how
journalists verified public discourse, the strategies and internal agreements implemented, and the degree of participation
of the 16 media involved. Results show that only half of the initiatives begun were transformed into published reports, and
the media impact achieved was limited. Finally, we note that the principal reasons for the frustration of the project were its
improvised implementation, due to the date of the election being brought forward, and the scant culture of collaboration
in the sector. In Spain at least, cross-media alliances are still an exception.

Issue
This article is part of the issue "Disinformation and Democracy: Media Strategies and Audience Attitudes" edited by
Pere Masip (University Ramon Llull, Spain), Bella Palomo (University of Málaga, Spain) and Guillermo López (University
of Valencia, Spain).

1. Introduction

The proximity between media companies and active
audiences has been the focus of numerous professional
projects and investigations over the last decade (Engelke,
2019), which have analyzed questions from basic inter-
action (Domingo et al., 2008) to user involvement as
co-creators (Sixto-García, López-García, & Toural-Bran,
2020) of news under the label of user-generated content
(Palomo, Teruel, & Blanco, 2019), or the tensions derived
from that relationship (Lewis, 2012). This practice of inte-
gration coexists with another participatory model involv-
ing collaboration between competing media, which has
ceased to be a utopian aspiration following the success

achieved by some of these initiatives. One of the most
relevant is the Panama Papers macro-project, which
involved teamwork by more than 400 journalists from
107 media organizations in 80 countries, and was rec-
ognized with a Pulitzer Prize. This type of partnership
is central to *ProPublica*, and has been applied in stories
like Unheard, Documenting Hate or Electionland (Eads,
2018). These cases are not something alien to local jour-
nalism, where according to the Center for Cooperative
Media they have doubled in two years (Wiltshire, 2019).

These joint efforts with colleagues form part of the
regular routines of investigative units, and have made it
possible to overcome censorship and strengthen secu-
rity measures in scenarios such as those found in Latin

America (Cueva Chacón & Saldaña, 2020). There are also other relevant benefits accompanying these practices, such as: sharing of costs, information and the time required for analyzing big datasets; increasing the scope of content; tackling complex reporting on a global scale; and renewing the news agenda (Carson & Farhall, 2018).

In this networked media ecosystem, formulas for collaboration between media companies are very diverse and have been taking place for decades. Media like *The Seattle Times* and *The Seattle Post-Intelligencer* carried out a joint operating agreement between 1983 and 2009 that consisted in centralizing advertising, production and circulation, and the publication of a combined Sunday edition that allowed editorial competition to be preserved (Picard, 2015). Following this same line, cross-promotion, the design and the adoption of productive protocols ("Trusted News Initiative (TNI) steps up," 2020) have given rise to proximity between some media companies.

The expansion of the culture of innovation has also transformed some newsrooms into laboratories that are open to experimentation, such as hiring developers and data journalists or creating hybrid work teams that boost creativity in order to compete in changing environments (Gade & Perry, 2003), which could solve some of the basic problems of legacy media (Klaß, 2020). Media and professional congresses have had recourse to creative spaces like hackathons to construct prototypes, explore new concepts or seek solutions to diverse challenges that fall outside the typical workday in a brief period of time, combining their energies through multidisciplinary teams (Boyles, 2017) of coders, designers and journalists, whose members on occasion had never worked together before (Toporoff, 2016).

This networking proves to be especially necessary for fact-checkers. Conscious of their reduced impact (Masip et al., 2020) and the financing problems involved, fact-checkers have forged alliances amongst themselves to strengthen their activity and, in parallel, have collaborated with media of record to reach wider audiences (Singer, 2020). These synergies have been especially evident during the global infodemic derived from the Covid-19 crisis, when the diffusion of fake news increased notably and citizens promoted this without being aware of it (Destiny & Omar, 2020). Outstanding in this respect is the union of 91 verification units from 70 countries to feed the database The CoronaVirusFacts/DatosCoronaVirus Alliance, supported by the International Fact-Checking Network. This network also inspired the creation of LatamChequea in Latin America, a collaborative platform funded by Google in which 35 organizations coordinated by Chequeado registered more than 2,000 rumors related to the Coronavirus, content that they offered on open access for its reuse to speed up the work of journalists.

But prior to the arrival of the world health crisis, numerous studies had demonstrated that the information disorder was becoming especially acute in the polit-

ical context (Freelon & Wells, 2020). In the United States many local and national news outlets featured political fact-checking for the 2012 and 2016 elections (Graves, 2016a). Concern about the impact that the circulation of false content on social media might have on elections (Aral & Eckles, 2019) is also favoring collaboration amongst major news and global tech organizations. An example of this is the emergence of The Trusted News Initiative (TNI), an international network founded in 2019 with the aim of protecting audiences and users from disinformation. The partners working together to identify and stop the spread of dangerous content are AFP, AP, BBC, CBC/Radio Canada, European Broadcasting Union, *Financial Times*, Reuters, *The Wall Street Journal*, *The Washington Post*, The Hindu, Facebook, Google, Microsoft, Twitter, First Draft, and the Reuters Institute for the Study of Journalism. These companies share an early warning system where organizations co-operate, sharing alerts rapidly when they discover disinformation which threatens human life or disrupts democracy during elections, and avoiding the republication of falsehoods. This was put into practice during the UK 2019 General Election, the Taiwan 2020 General Election, and the Coronavirus crisis ("Trusted News Initiative (TNI) steps up," 2020).

That dynamic of activity was implanted by First Draft in earlier projects with a double aim: to defend the media from those accusing them of constructing fake news by increasing the credibility of journalistic production, thus revitalizing the tradition of truth-seeking in the field (Graves, 2016a); and to reduce the social confusion caused by the intense diffusion of false or malicious content. To achieve this end, since 2016 First Draft has promoted collaborative reporting and cross-checking experiences around elections in the United States, France, United Kingdom, Germany, Brazil, Nigeria, Spain, Uruguay, Argentina, Australia and the European Union, based on the understanding that competing newsrooms can work together for more effective, efficient and responsible news coverage. Similarly, this collaboration prevents the duplication of newsrooms debunking the same content and ensures that quality information reaches large audiences. According to Claire Wardle, First Draft Director, "to crosscheck a report means reviewing and approving the verification steps taken by another newsroom, adding the logo of your organization alongside other contributing partners, and then amplifying the report to existing audiences" (Wardle, Pimenta, Conter, Días, & Burgos, 2019, p. 4). One of the initiatives that generated less media interest and produced a smaller number of reports was the project developed in Spain, as can be seen in Table 1.

This research focuses on the Spanish case because of the particularities concentrated in this scenario. For many years social media have been the preferred starting point for the consumption of daily news in Spain (Gottfried & Shearer, 2016). They have surpassed their intermediary role and become gatekeepers of

Table 1. List of First Draft election cross-checking projects

Project	Partners	Reports	Year
CrossCheck France	AFP, BuzzFeed News, France 24, The Observers, *La Voix du Nord*, Bellingcat, *Rue89 Bordeaux*, *Les Echos*, *Rue89 Strasbourg*, *Libération*, France Télévisions, Les Décodeurs, Storyful, ScienciesPo, Global Voices, *Street Press*, *Ouest France*, *La Provence.com*, *Rue89 Lyon*, Facto Scope 2017, *L'express*, *Nice-Matin*, *Le Journal du Dimanche*, Explicite, Centre France, LCI, *Le Monde*, *Sud Ouest*, Euronews, Meedan, *Le Télégramme*, *L'avenir.net*, Euractiv, *Saphir News*, London School of Economics, 4 News, Bloomberg, EPFT, Google News Lab, Facebook, CFJ, Crowdtangle, Hearken, NewsWhip, SAM, First Draft.	67	2017
Comprova Brazil	Brazilian Association of Investigative Journalists, NSC Comunicação, UOL, *Folha de S.Paulo*, *O Estado de S. Paulo*, *Correio do Povo*, *Jornal do Commercio*, *O Povo*, *Metro Brasil*, *Exame*, Nova Escola, Piauí, *Veja*, Band News, Band TV, Canal Futura, SBT, Band News FM, Bandeirantes, AFP, *Nexo Jornal*, *Poder360*, *Gazeta Online*, *GaúchaZH*, *Gazeta do Povo9*, Projor, Google News Initiative, Facebook Journalism Project, National Newspaper Association (ANJ), David Rockefeller Center for Latin American Studies at Harvard University, Aos Fatos Brazilian fact-checking agency, Armando Alvares Penteado Foundation (FAAP), RBMDF Advogados Brazilian law firm, CrowdTangle, NewsWhip, Torabit, Twitter, WhatsApp, First Draft.	146	2018–actual
Comprobado Spain	AFP, *Ara*, Datadista, *Diario de Navarra*, EFE, *El Confidencial*, *eldiario.es*, *El Faradio*, Europa Press, *La Marea*, Newtral, Politibot, *Público*, RTVE, Servimedia, Maldita.es, First Draft.	42	2019
CrossCheck Nigeria	*Premium Times*, *Daily Trust*, News Agency of Nigeria, AFP, *The Nation*, *Tribune*, Africa Check, *The Guardian*, *Punch*, *The Sun*, Channels Television, The Cable, *The Niche*, Sahara Reporters, Freedom Radio, University of Lagos Mass Communication Department, CODE, International Centre for Investigative Reporting, Google News Initiative, Open Society Foundations, MacArthur Foundation, Ford Foundation, WhatsApp, Facebook Journalism Project, First Draft.	60	2019
CrossCheck Australia	The Centre for Media Transition at the University of Technology Sydney, Crikey, SBS News, NZME, UTS, RMIT University, AUT, HKU, QUT, First Draft.	n.a.	2019
Reverso Argentina	0223, 12noticias TV, *7Corrientes*, A24, ADN Sur, AM750, Ámbito, Anfibia, BAE Negocios, *BBC Mundo*, *Bumerang News*, C5N, Cable a Tierra, Canal 10 Mar del Plata, Canal 10 Tucumán, Canal 13, Canal4 (Posadas), Canal 7, *Clarín*, *Crónica*, *Cuarto Poder Diario*, *Diario Andino*, *Diario Huarpe*, *Diario Jornada*, *Diario La Mañana* (Formosa), *Diario Popular*, *Diario UNO*, *Diarios Bonaerenses*, Economis, *El Cronista*, *El Día*, *El Diario de Miramar*, *El Liberal*, *El Litoral*, *El Nueve*, *El Territorio*, *El Tribuno*, *Filo.News*, FM 89.3 Santa María de las Misiones, FM Cielo 103.5, FM El aire de integración, FM La Redonda, FM Milenium 106.7, *Hoy Día Córdoba*, *Infobae*, *Infocielo*, Infopico, Iprofesional, *La Capital*, *La Capital de Mar del Plata*, *La Gaceta*, *La Izquierda Diario*, *La Nación*, *La Nota*, *La Nueva*, *La Voz del Interior*, *LatFem*, *Los Andes*, LT7 Radio Provincia de Corrientes AM 900, LT17 Radio Provincia Misiones, LU2, Marcha, MDZ online, MDZ Radio, Mega 98.3, *Mendoza Post*, *Meridiano55.com*, Milénico, *Minutouno.com*, *Misiones Online*, *Misiones Opina*, Nodal, *Noticias Jesús María*, *Noticias Mercedinas*, *OPI*, *Página 12*, *Pausa*, Periódicas, POP Radio, Portal Misiones, Presentes, *Primera Edición*, *QUÉ Digital*, Radio Andina 90.1, Radio con Vos, Radio Mitre, Radio Nihuil, Radio10, RedAcción, Revista Hamartia, *Río Negro*, Rosario3, *Sitio Andino*, SL24, Taringa, Telefe, Telégrafo, *Tiempo de San Juan*, *Tiempo del Este*, TKM, TN, TVA, TV Pública, Unidiversdad, UNO, vaga fiebre, Verte, Vía País, Chequeado, AFP Factual, First Draft, Pop-up Newsroom.	180	2019

Table 1. (Cont.) List of First Draft election cross-checking projects

Project	Partners	Reports	Year
Verificado Uruguay	*La Diaria*, *Búsqueda*, *El País*, *Brecha*, En Perspectiva, Cori, AFP, Océano FM, Radio Carve, Radio Monte Carlo, Radio Sarandí, Radio Universal, RadioMundo, Radio41, *Crónicas*, TV Ciudad, Efe, Sudestada, UY Press, Martes, *Maldonado Noticias*, Organización de Prensa del Uruguay, Asociación de Periodistas del Uruguay, Nadie Chequea Nada, UYCheck.com, Universidad de la República, Universidad de Montevideo, Universidad Católica del Uruguay, Universidad ORT, Facebook, Google News Initiative, Fundación Avina, First Draft.	76	2019

online content and traffic. Parallel to this, according to the Reuters Report 2020 (Newman, Fletcher, Schultz, Andi, & Nielsen, 2020), the country ranks first in terms of news consumption via mobile devices (73%) and tenth via WhatsApp (34%), which is where the greatest number of hoaxes are propagated (Salaverría et al., 2020). These transformations have favored the spread of information disorders. The forecasts show that in 2022 citizens of developed countries will consume more disinformation than true news, because false content is 70% more likely to viralize and be retweeted compared to truthful information (Vosoughi, Roy, & Aral, 2018). The situation is especially alarming in Spain where 57% of Spaniards have at some moment believed false news. The Eurobarometer *Fake News and Online Disinformation* (Directorate-General for Communications Networks, Content and Technology, 2018) also identifies the Spaniards as the European citizens who have the most difficulties in detecting this type of content and as being amongst those most concerned about this issue in the world (Nicholls, Fletcher, Kalogeropoulos, & Nielsen, 2019).

These antecedents explain why in the spring of 2019 a coalition called Comprobado (Verified) was formed in Spain. This was coordinated by First Draft and Maldita and was initially made up of 16 media with the goal of putting a brake on disinformation during an intense electoral period, since in April and May Spanish citizens were participating in national, municipal and European elections.

The novelty of this article lies in setting out the internal work dynamic developed in this platform and analyzing the media visibility the project achieved.

2. Literature Review

The professional and academic debate on false content has intensified since 2016. During an initial conceptual phase the definition of fake news was clarified (Tandoc, Lim, & Ling, 2018) and use of the term was even discouraged (Wardle, 2017); the three types of information disorder were defined (misinformation, disinformation and malinformation); and it was established that satire and parody, false connection, misleading content, false context, imposter content, manipulated content

and fabricated content constitute the seven categories of information disorder (Wardle, 2018). Due to the electoral context that determined this investigation, our article prioritizes the concept of disinformation because it "refers to situations where actors driven by political and/or economic interests, produce and distribute information intended to disinform for their own ends" (Westlund & Hermida, in press).

More recently, numerous investigations have focused on the process and impact of verification, reflecting the exponential growth of this activity. According to Duke Reporters Lab, in 2014 there were 44 registered fact-checkers, while in 2020 the figure has risen to 290 (Stencel & Luther, 2020). Although by continent the greatest number of projects is concentrated in Europe (85), it is worth stressing the efforts made in Africa, Asia and South America, where the offer has doubled. The complexity of disinformation has also required the involvement of European institutions, governments and technological companies to contribute to its reduction. France and Germany were pioneers in passing laws to identify, stop and penalize those intending to propagate fake news (López-García, Vizoso, & Pérez-Seijo, 2019), while Facebook has allied itself with third-party fact-checkers in order to help improve quality (Saurwein & Spencer-Smith, 2020).

The peculiarities of this international multi-localization have been analyzed by Palau-Sampio (2018), who shows that no homogeneity exists in the model for evaluating the content that is subject to verification by the fact-checkers, and she suggests resolving this lack of uniformity by establishing a universal procedure.

Investigations focusing on the collaboration/fact-checking relationship prove to be scarce, and tend to provide a positive perspective on its implementation. Thus, earlier studies have analyzed cross-checking experiences as a legitimization strategy of the journalism field in response to disinformation, and have discovered how the description of the verification process has developed into new and successful narratives and formats (Young, Jamieson, Poulsen, & Goldring, 2018), where false content is also transformed into news (Trevisan Fossá & Müller, 2019). The transparency of such procedures is a basic requisite for transmitting journalistic truth (Humprecht, 2020; Kovach & Rosenstiel, 2007).

Nonetheless, this emergent practice has not been free of criticism. Fact-checkers have been accused of being partisan (Stencel, 2015), and it has even been suggested that fact-checking in general is unsuccessful in reducing misperceptions. Brandtzæg and Følstad (2017) analyzed users' perceptions of three fact-checking services and concluded that users with negative perceptions were trapped in a perpetual state of informational disbelief. Moreover, considered from the professional perspective, journalists regard fact-checking services with caution and skepticism, as they need to have a high degree of control over the process, which can prove difficult with third-party services (Brandtzæg, Følstad, & Chaparro, 2018).

Graves and Cherubini (2016) distinguish two models of fact-checkers: the newsroom model; and the NGO model. First Draft has opted for a hybrid model that integrates media companies and independent fact-checkers or professional associations and is free of editorial and business constrains in order to obtain a wide reach and ensure that nonpartisan information is diffused. To gain a better understanding of how this experience was managed in the Spanish scenario, the following research questions were posed:

RQ1: What was the internal work dynamic in Comprobado and how did the journalists adopt these routines?

RQ2: What repercussion did this media collaboration have?

3. Methods

This article presents an explorative approach to a case study, the Spanish project Comprobado, through the application of mixed methods. The integration of qualitative and quantitative data helped in obtaining a more complete analysis, which is also necessary due to the novelty of the object of study. Antecedents in the analysis of the fact-checking movement in journalism also had recourse to ethnography in order to demonstrate that mechanical testing to verify facts is a complex matter, and that fact-checkers base themselves on multiple pieces of evidence, not always conclusive, until they achieve factual coherence or a triangulation of the truth (Graves, 2016b). The present investigation enquires more deeply into this type of routine. In our case data are derived from more than 50 hours of observing procedures, training sessions and interviews, reports and a systematic analysis of the activity developed in CrossCheck, the work platform used by the media coalition. The benefits derived from the complementarity of online observations, interviews and content analysis for data-gathering in Internet-based research have been defended previously (Sade-Beck, 2004).

The University of Malaga was given the opportunity to partner with this experimental project, applying non-participant observation during its realization and participating in all the meetings as it formed part of the deliberative council, including the final one where the post-mortem review occurred (Collier, DeMarco, & Fearey, 1996). Over the course of three months, the period for which the project was active, data referring to the production routines was collected, a methodology that has also been applied in studies concerned with the same question (Henderson, 2020). From the first meeting in February 2019 until the final one in May, a diary was kept containing field notes that registered the perceptions of the fact-checkers concerning the project and how they discuss their methods and proposed solutions. For ethical reasons the identities of the participating journalists have been withheld.

The main advantage of obtaining ethnographic information from the virtual social sphere is that it provides accessibility to subjects who are physically remote from the researcher. That is why, in the second place, all of the 82 investigations housed on the CrossCheck platform and the 250 comments by 29 jounalists derived from them between 10 April and 25 May 2019, were subjected to a content analysis. This analysis was carried out using a datasheet designed to identify: 1) investigations started; 2) investigations published; 3) labels applied, 4) journalist and medium that started the investigation; 5) journalists and media participating in the verification; 6) internal comments and debates generated; 7) theme and political parties mentioned; 8) origin of content; 9) verification strategies; 10) evidence used; 11) media that validate the investigation. This analysis occurred over the lifetime of the project, making it possible to produce results that invited reflection during the project meetings.

In the third place, to determine the reach of both the project and the verifications in the collaborating media, an analysis was made of the word 'comprobado' in the press archives of the media, which was complemented by a Boolean search in Google (example: 'comprobado' site:afp.com) and the published tweets. Obtaining these quantitative data made a parallel approach with other crosscheck initiatives possible. The period of analysis was established from 1 March to 30 June in order to detect prior and post-project news coverage.

Once the project had concluded, the qualitative research was completed with in-depth interviews as a data collection method (Alshenqeeti, 2014) in order to try to understand the initative from the subject's point of view. Semi-structured interviews were conducted by phone with four journalists to learn their evaluation of the experience, highlighting the strengths and weakness of cross-checking in Spain. The profiles of these interviewees were as follows: one of the project coordinators; one of the journalists who had been most active during the project; and two legacy media journalists who attended the first meetings but whose media finally refused to participate. None of the journalists who had been less active on the project agreed to be interviewed. All the interviews together with the notes and comments

were uploaded to the Atlas.ti application to categorize the content and enrich the qualitative results.

4. Results

4.1. Internal Involvement to Build Trust

In 2018 Maldita and First Draft developed a collaborative fact-checking project in Spain, which replicated the successes achieved in France and Brazil. According to the observations, the rescheduling of the elections upset the planned chronogram and made it necessary to design the initiative rapidly. In February 2019 representatives of 40 Spanish newsrooms were brought together in Madrid to convince them that uniting experiences would favor truth production, and that collective vigilance would make it easier for journalism to perform its function as a counter-power. Concerns about professional value and status were the main motivations for journalists to practice fact-checking (Graves, Nyhan, & Reifler, 2016). During the conference the project was set out, the social and journalistic benefits of this working formula were addressed and several experts held verification workshops. The meeting was held in an atmosphere of 'friendly competitors,' but there was also prudence and skepticism.

The initial requisite established by First Draft to be able to launch the project was securing the participation of more than ten media that would represent the national, regional and local spheres. Having obtained the initial collaboration of 16 media (Table 2), it was consensually decided to work under the name of Comprobado. This decision was influenced by their Brazilian predecessor, Comprova, and by an integrative intention, since the name CrossCheck Spain might discourage participation by pro-independentist media. This macro-project was a pioneer in Spain and commenced its activities on 10 April

with a pre-test, while its final verification took place on 25 May. At that moment the CrossCheck platform had a register of 55 users distributed across the following organizations: First Draft/CrossCheck (10), Maldita (8), *Ara* (5), RTVE (5), Newtral (4), Efe (4), *Eldiario.es* (3), AFP (2), Datadista (2), *El Confidencial* (2), *El Faradio* (2), *La Marea* (2), *Público* (2), *Diario de Navarra* (1), Europa Press (1), Politibot (1), and University of Malaga (1).

None of the national newspapers of record joined the project. The interviews we conducted showed that this was a foreseeable handicap. As one of the reporters attending the initial meeting observed, "We journalists were better prepared than the media themselves. The people in charge of the latter were distrustful of sharing part of their work with competitors." The journalists shared various hypotheses for this investigation. While some alluded to the arrogance of the mainstream media that believe they don't need help in fulfilling their daily work or to the impossibility of their making an editor available in the middle of an electoral campaign, others indicated that the real motive was the effort and difficulties involved in setting an initiative underway that did not form part of their normal routine. But what emerged as the common nexus of all these suppositions was the lack of a collaborative culture. One of the founders of Maldita noted that "a considerable amount of social engineering [was done] to convince them, but more time was needed to achieve it."

The prior training of participants was not analyzed, and this could be the key to the success of cross-checking projects. We were informed by Maldita that a fact-checker requires some 25–30 hours training, and that daily practice for months and internal debates are needed for resolving a verification as best as possible. In this project, following an initial pre-test proposed by a member of First Draft, the journalists began their activity on the CrossCheck platform. Each reporter could create

Table 2. Typology of the media participating in Comprobado.

Media	Typology	Founded
AFP	International news agency	1835
Ara.cat	Catalan newspaper with pro-independence ideology	2010
Datadista	Online native medium focused on data journalism	2016
Diario de Navarra	Conservative newspaper	1903
EFE	National news agency	1939
El Confidencial	Online native newspaper with liberal ideology	2001
Eldiario.es	Online native newspaper with left-wing ideology	2012
El Faradio	Online native newspaper focused on local journalism	2012
Europa Press	National news agency	1953
La Marea	Magazine and online newspaper with republican ideology and co-operative model	2012
Maldita	Fact checker	2018
Newtral	Fact checker	2018
Politibot	Telegram bot focused on political content	2016
Público	Online newspaper with left-wing ideology	2007
RTVE	National public broadcasting service	2007
Servimedia	National news agency	1988

an investigation, publishing a post, inserting images and identifying tags, and sending the alert to the rest of the coalition representatives to evaluate the verification. When at least three media confirmed the same result they contributed their logo and each collaborating medium could publish the report-card—designed to facilitate its distribution and consumption in mobile format—and the procedure applied to reach the conclusion of whether the content was true or false.

Of the 16 participating media, 9 were digital native media. Preliminary findings suggest that digital native media were the most active when it comes to collaboration, which coincides with other prior studies that also concluded that the big challenge lay in persuading the legacy media (Hatcher & Thayer, 2016; Méndez, Palomo, & Rivera, 2020). Also noteworthy was the participation of three news agencies, which was justified by their traditional commitment to verifying content (Wishart, 2018). The project coincided with the birth of specialized units, such as Verifica RTVE and Efe Verifica.

Collaboration was open to all the news companies that respected the code of principles (accuracy, ethical responsibility, fairness and impartiality, independence, transparency) and whose request to join was approved by the deliberative council of Comprobado. During the period of the project's realization only one request was received, that of EMA-RTV (Asociación de Emisoras Municipales y Comunitarias de Andalucía de Radio y Televisión).

For six weeks they monitored content that was basically circulating on social media. The source that gave rise to most disinformation was Twitter, which accounted for half of the verifications. 20 percent of the investigations originated in public statements published in the mass media. The results proceeding from WhatsApp, Facebook or Instagram were not very representative since the project lacked accounts that would have favoured interaction on those channels, and the necessary infrastructure for automating the process.

Every time that a member detected possible disinformation, they indicated this on the platform and each medium received an alert, usually with a question as its title, inviting it to participate in the investigation. In more complex cases it was necessary to undertake and explain eight actions to conclude the report. This was due to the need to make searches for several statements and denials on social media, use translators, or even make a request to the Traffic Department for a record of incidents involving the number plate of a vehicle attributed to a political party.

The participating media not only verified content; the essential element in the collaboration consisted in filling out a report-card on which they explained clearly and objectively the steps taken during the verification process, as this narrative consistency strengthened the credibility of the activity and enabled any interested citizen to reproduce the verification. With respect to evidence, the platform made it possible to attach files, photos or links, the latter being the most employed. In spite of the fact that diagnostic tools for verifying content have become increasingly sophisticated, what predominated in the routines was monitoring online conversations in real-time and contacting official sources. During an initial phase they looked for evidence by locating already existing resources (recordings of interviews, electoral programs of parties, reverse image searches, official reports) and if these proved insufficient, they contacted the primary sources. The investigations were centered on about 50 different individuals, and this variety made it necessary to contact some thirty official sources.

In that month and a half, 82 investigations were started and at the end of the project 46 of them met the requirement of having a report-card and being verified by at least three media. The content analysis carried out showed that of the 55 journalists registered on the platform, only 29 collaborated in a verification process during the six weeks that the project was active, and five of the media in the alliance did not validate any investigation. The verification process involving the greatest number of media received support from eight. The most regular contribution proceeded from the team at the public broadcaster RTVE and the fact-checking platform Maldita, which also made the most proposals as they initiated 75% of the investigations. In this sense, Maldita's initiative was essential to maintaining the project active to the end. This unequal level of participation explains why the main tension and frustration detected in this alliance arose when comparisons were made of the time invested by the media in verifications, a situation that replicated what had happened in previous experiences (Singer, 2004). For one of the most active journalists, the lack of coordination and the unequal participation were the main weaknesses of Comprobado: "We didn't have the sensation that there was real collaboration. What's more, at some moments we had the sensation that there was some competition between the teams."

The circulation of conspiracy theories, malicious readings of political statements and electoral programs, or accusations of electoral fraud and its influence on destabilizing the democratic system, especially during the week leading up to the elections, did not alter the inscrutable attitude of the majority of the media.

Concerning the prominence of political parties in these investigations, an imbalance could also be observed in favor of the more extreme parties. Thus, Unidas Podemos was present in three out of every ten investigations, and Vox, an extreme right-wing party, in two out of every ten. The journalists consulted denied any tendentiousness in the selection, and explained that the choice was determined by the most viral and journalistically relevant content.

The 46 final reports prepared for the Comprobado website reflected four principal categories: unreliable (80.4%), inconclusive (8.7%), be aware (6.5%), and trustworthy (4.4%). Each category appeared in the design of the report-card, influencing its color so that its visual

appearance would facilitate comprehension and thus avoid resending any unevaluated false content. As many as 14 different subcategories were used to determine the type of unreliable content, with a predominance of content that had been digitally altered, manipulated or proceeded from deliberate fabrication (32.4%), followed by those for which there was no supporting evidence (18.9%) and wrong context (16.2%).

During the project three coordination meetings were held between March and May, at which agreement was reached on patterns and recommendations for action, such as the need for each collaborating medium to allow one editor to dedicate between 5 and 10 hours per week to Comprobado. On 6 May, following the general election of 28 April, another videoconference was held at which a commitment was made to continue the collective fact-checking project during the regional elections, and it was agreed to adopt routines such as the obligation to enter the platform on a daily basis for at least five minutes to crosscheck. It was understood that although not all the teams had the possibility of investigating, they could at least check the verifications of colleagues. On that occasion it was also agreed to create the Telegram group to speed up communications between the 11 journalists who volunteered to participate in that space. The themes of the messages were related to technical problems and alerts to solve errors, although in general communication on this channel was not very fluid, with only 19 messages registered in three weeks. This figure contrasts with the Brazilian case, where approximately 50 journalists who participated in Comprova shared 18,154 messages in their private WhatsApp Group (Wardle et al., 2019, p. 7).

4.2. The Visibility of Comprobado

On 11th April the first announcement was made of the birth of this pioneering media alliance in Spain. Maldita and AFP gave the widest diffusion to the initiative, and the practice of a self-reference strategy was also detected when the project started, which enabled it to exceed 5,000 followers on the official Twitter account @ComprobadoES on the first day. Nonetheless, the content analysis confirms that no textual mention of Comprobado at all could be found in five of the sixteen media involved. An analysis of the press

archives of the media involved and their social media accounts showed very little self-promotion was carried out during the development of the project, only about 50 news items were localized and the majority concerned the launching of the initiative. This silence was also practiced with the verifications, although the media were under no obligation to publish them. Of the 46 report-cards created, 38 appeared on the official Twitter account @ComprobadoES. This account achieved 124 tweets, although one-third were self-promotions, one-third were new investigations and the remaining third were retweets of verifications. A strong contrast can therefore be noted with respect to the production and impact obtained in other cross-checking projects, although the period of existence of these accounts also differs (Table 3). While @ComprobadoES was active for only two months, @Comprova was still active at the moment when this investigaiton was closed.

The experience in Brazil was one of the most significant since, according to the Comprova Report, 40.4% of those surveyed said this initiative helped them to decide their vote. In this case, the 146 verifications developed on the Comprova platform were transformed into 1,750 items of content distributed in television news programs, newspapers, magazines and online media (Rinehart, 2019). In Argentina, Reverso exceeded 20,000 followers on Twitter, where 1,447 tweets were published, although that figure was doubled in Mexico, where the highest number of followers was obtained. In Spain, the scant support received from the legacy media lowered expectations and the visibility of the initiative, but other questions also influenced this result, such as the early elections, which made it necessary to precipitate the birth of Comprobado and made it impossible to reproduce the recommended periods. In the case of Comprova, five meetings were held prior to its start-up, including workshops and a bootcamp; grants were offered for participating newsrooms; the project had an open-access website; Facebook and Twitter published adverts supporting its activity and ad campaign groups were created with the result that social media were responsible for almost half of all Comprova's web traffic. Audience proximity was essential for spreading knowledge about Comprova and its usefulness. Proof of this proximity is that there were 350,567 interactions between users and journalists on WhatsApp during the 12 weeks of the

Table 3. Cross-checking projects on Twitter.

Project	Tuits	Followers	Founded in	Last tweet
@VerificadoMX	3208	177,9k	02.2018	10.07.2018
@Comprova	1956	28,2k	06.2018	Active
@ReversoAR	1447	20,1k	05.2018	11.12.2019
@verificadouy	556	16,6k	06.2019	04.12.2019
@crosscheckNG	410	02,1k	11.2018	12.09.2019
@CrossCheckFR	186	04,6k	01.2017	01.03.2019
@ComprobadoES	124	10,2k	04.2019	11.06.2019

project, with one-third of the messages proceeding from the audience, which included suspicious claims, images, video or audio messages for the professional team to debunk (Wardle et al., 2019, p. 7). The strategies followed in France (Smyrnaios, Chauvet, & Marty, 2017) and Argentina were similar.

Although Comprobado's quantitative data were unsuccessful and demonstrate that the model of horizontal coordination to facilitate reaching consensus did not work, at Maldita they consider that the qualitative reading was positive. This was because the experience made it possible to bring together very different newsrooms, share routines between journalists that did not know each other before, and spread knowledge about fact-checking and its implementation with new formulas via outsourcing (Méndez et al., 2020). The system's open character was also one of the project's main weaknesses according to Maldita, and one of the lessons learned. If a similar initiative is attempted in the future, it will only involve the more committed companies.

5. Conclusions

The media industry is adapting to a post-truth age, developing innovative approaches in order to produce quality journalism and repair its damaged credibility (Carson & Farhall, 2018). Both competition and collaboration are considered relevant actions for surviving in the temporalities of the media ecology (Dodds, 2019; Eads, 2018). The actions implemented in this respect in newsrooms must also consider the expectations of audiences, who are demanding objective reporting, analysis, explanation and transparency from the journalist (Loosen, Reimer, & Hölig, 2020). The initiatives promoted by First Draft have taken all these characteristics into account, designing places of truth production (Trevisan Fossá & Müller, 2019), a new model where the watchdog role is shared and the function of a gatekeeper working for the public interest is strengthened. This is an old-new journalism that is more open and responsible.

This article has analyzed the internal activity of the media that formed an alliance to put a brake on the circulation of false news stories during the 2019 elections in Spain and the visibility received by that initiative. In answer to the first question posed, this investigation describes a systematic strategy whose implementation depended on the willingness of each journalist. In theory, the combination of effort frees more time for developing other stories in depth, and audiences profit from these strategies. However, not all the media believed that working together rather than separately would be more valuable and efficient; only a dozen journalists from five media were active and carried out a forensic content analysis, with half of the initiatives left unfinished. There was a shortage of time for convincing, trust and commitment.

That is why in spite of the success achieved in previous cross-checking initiatives (Wardle et al., 2019),

this collaboration in Spain had less impact. The project revealed the existence of a collision of cultures (Dailey, Demo, & Spillman, 2005) that provoked the resistance of the big national news outlets to participating in this coalition and an uneven involvement by the collaborating journalists. While it is possible to identify practices related to the convergence continuum model, and especially the coopetition stage (Dailey et al., 2005), the abovementioned factors reduced the initiative's impact. This was also influenced by the fact it did not reproduce the same model that was implanted in countries like Brazil and Argentina, where coordination and contacts were designed and established during the six months prior to the elections, while in the Spanish case Comprobado's development was precipitated. Similarly, accessibility to Comprobado's activity was articulated around two axes: the diffusion that each medium decided to give to a verification, and a Twitter account.

With respect to the study's limitations, an analysis of the Spanish case cannot be generalized to the rest of the CrossCheck initiatives, and the data from the Brazilian case shows that there can be clear differences between countries. As Humprecht (2020) observes, professional fact-checking depends on the country in which one lives and on occasions fact-checking plays a less important role in non-Anglophone countries. This preliminary exploration of a political and media context as polarized as the Spanish case makes it possible to propose a comparative study with the rest of the countries, and even a longitudinal analysis if the initiative is repeated in future elections. Additionally, we suggest pursuing the qualitative approaches that make it possible to determine how journalists and audiences perceive the work developed by cross-checking experiences, discover the editorial motivations invoked by media so as not to participate in such alliances, and whether a program of incentives would alter this reluctance or even favor the long-term sustainability of such collaborations.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

References

Alshenqeeti, H. (2014). Interviewing as a data collection method: A critical review. *English Linguistic Research*, *3*(1), 39–45.

Aral, S., & Eckles, D. (2019). Protecting elections from social media manipulation. *Science*, *365*(6456), 858–861.

Boyles, J. L. (2017). Laboratories for news? Experimenting with journalism hackatons. *Journalism*, *21*(9), 1338–1354. https://doi.org/10.1177/14648849177 37213

Brandtzæg, P. B., & Følstad, A. (2017). Trust and distrust in online fact-checking services. *Communications of the ACM*, *60*(9). https://doi.org/10.1145/3122803

Brandtzæg, P. B., Følstad, A., & Chaparro, M. A. (2018). How journalists and social media users perceive online fact-checking and verification services. *Journalism Practice*, *12*(9), 1109–1129.

Carson, A., & Farhall, K. (2018). Understanding collaborative investigative journalism in a "post-truth" age. *Journalism Studies*, *19*(13), 1899–1911.

Collier, B., DeMarco, T., & Fearey, P. (1996). A defined process for Project post mortem review. *IEEE Software*, *13*(4), 65–72.

Cueva Chacón, L., & Saldaña, M. (2020). Stronger and safer together: Motivations for and challenges of (trans)national collaboration in investigative reporting in Latin America. *Digital Journalism.* Advance online publication. https://doi.org/10.1080/ 21670811.2020.1775103

Dailey, L., Demo, L., & Spillman, M. (2005). The convergence continuum: A model for studying collaboration between media newsrooms. *Atlantic Journal of Communication*, *13*(3), 150–168.

Destiny, O., & Omar, B. (2020). Fake news and Covid-19: Modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, *56*, 101475. https://doi.org/10.1016/j.tele.2020.101475

Directorate-General for Communications Networks, Content and Technology. (2018). Flash Eurobarometer 464: Fake news and disinformation online. Brussels: European Commission. Retrieved from https:// op.europa.eu/en/publication-detail/-/publication/ 2d79b85a-4cea-11e8-be1d-01aa75ed71a1/ language-en

Dodds, T. (2019). Reporting with WhatsApp: Mobile chat applications' impact on journalistic practices. *Digital Journalism*, *7*(6), 725–745.

Domingo, D., Quandt, T., Heinonen, A., Paulussen, S., Singer, J. B., & Vujnovic, M. (2008). Participatory journalism practices in the media and beyond. *Journalism Practice*, *2*(3), 326–342.

Eads, D. (2018, June 19). When do reporters collaborate, and when do they compete? *ProPublica*. Retrieved from https://www.propublica.org/article/ask-propublica-illinois-reporters-journalism-collaboration-between-newsrooms

Engelke, K. (2019). Online participatory journalism: A systematic literature review. *Media and Communication*, *7*(4), 31–44.

Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, *37*(2), 145–156.

Gade, P. J., & Perry, E. L. (2003). Changing the newsroom culture: A four-year case study of organizational development at the St. Louis Post-Dispatch. *Journalism & Mass Communication Quarterley*, *80*(2), 327–347.

Gottfried, J., & Shearer, E. (2016, May 26). News use across social media platforms 2016. *Pew Research Center*. Retrieved from https://www.journalism.org/ 2016/05/26/news-use-across-social-media-platforms-2016

Graves, L. (2016a). *Deciding what's true: The rise of political fact-checking in American journalism*. New York, NY: Columbia University Press.

Graves, L. (2016b). Anatomy of a fact check: Objective practice and the contested epistemology of fact ckecking. *Communication, Culture & Critique*, *10*(3), 518–537.

Graves, L., & Cherubini, F. (2016). *The rise of fact-checking sites in Europe*. Oxford: Reuters Institute for the Study of Journalism.

Graves, L., Nyhan, B., & Reifler, J. (2016). Understanding innovations in journalism practice: A field experiment examining motivations for fact-checking. *Journal of Communication*, *66*(1), 102–138.

Hatcher, J. A., & Thayer, D. (2016). Assessing collaboration in one media ecosystem. *Journalism Practice*, *11*(10), 1283–1301.

Henderson, K. (2020). Marketing the construction of reality: Multiplatform production routines and the renegotiation of journalistic role identities in a legacy local televisión newsroom. *Journalism*. Advance online publication. https://doi.org/10.1177/1464 884920940940

Humprecht, E. (2020). How do they debunk "fake news"? A cross-national comparison of transparency in fact checks. *Digital Journalism*, *8*(3), 310–327.

Klaß, N. (2020). Open innovation in media innovation research: A systematic literature review. *Journal of Media Business Studies*, *17*(2), 190–218.

Kovach, B., & Rosenstiel, T. (2007). *The elements of journalism: What newspeople should know and the public should expect*. New York, NY: Three Rivers Press.

Lewis, S. C. (2012). The tension between professional control and open participation. *Information, Communication & Society*, *5*(6), 836–866.

Loosen, W., Reimer, J., & Hölig, S. (2020). What journalists want and what they ought to do (in)conguences between journalists' role conceptions and audiences.' *Journalism Studies*, *21*(12), 1744–1774. https://doi.org/10.1080/1461670X.2020.1790026

López-García, X., Vizoso, A., & Pérez-Seijo, S. (2019). Verification initiatives in the scenario of misinformation: Actants for integrated plans with multi-level strategies. *Brazilian Journalism Research*, *15*(3). https:// doi.org/10.25200/BJR.v15n3.2019.1215

Masip, P., Almenar, E., Aran, S., Capilla, P., Puertas, D., Ruiz, C., . . . Zilles, K. (2020). *El consumo de información durante el confinamiento por el coronavirus: Medios, desinformación y memes* [Information consumption during coronavirus confinement:

Media, misinformation and memes]. Barcelona: Ramon Llull University. Retrieved from https://www.researchgate.net/publication/340685890_El_consumo_de_informacion_durante_el_confinamiento_por_el_coronavirus_medios_desinformacion_y_memes

Méndez, A., Palomo, B., & Rivera, A. (2020). Managing social networks in online-native newsrooms: When less means more. *Media and Communication*, *8*(2), 124–134.

Newman, N., Fletcher, R., Schultz, A., Andi, S., & Nielsen, K. (Eds.). (2020). *Digital news report 2020*. Oxford: Reuters Institute for the Study of Journalism.

Nicholls, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, K. (Eds.). (2019). *Digital news report 2019*. Oxford: Reuters Institute for the Study of Journalism.

Palau-Sampio, D. (2018). Fact-checking and scrutinity of power: Supervision of public discourses in new media platforms from Latin America. *Communication & Society*, *31*(3), 347–365.

Palomo, B., Teruel, L., & Blanco, E. (2019). Data journalism projects based on user-generated content. How La Nacion Data transforms active audience into staff. *Digital Journalism*, *7*(9), 1270–1288.

Picard, R. G. (2015). Natural death, eutanasia, and suicide: The demise of joint operating agreements. *Journal of Media Business*, *4*(2), 41–64.

Rinehart, A. (2019, June 26). What we learned in Brazil about building trust and collaborations. *First Draft*. Retrieved from https://firstdraftnews.org/latest/what-we-learned-in-brazil-about-building-trust-and-collaborations

Sade-Beck, L. (2004). Internet ethnography: Online and offline. *International Journal of Qualitative Methods*, *3*(2), 45–51.

Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I., & Erviti, M. C. (2020). Desinformación en tiempos de pandemia: Tipología de los bulos sobre la Covid-19 [Disinformation in times of pandemic: Typology of hoaxes on Covid-19]. *El Profesional de la Información*, *29*(3). https://doi.org/10.3145/epi.2020.may.15

Saurwein, F., & Spencer-Smith, C. (2020). Combating disinformation on socialo media: Multilevel governance and distributed accountability in Europe. *Digital Journalism*, *8*(6), 820–841.

Singer, J. (2004). Strange bedfellows? The diffusion of convergence in four news organizations. *Journalism Studies*, *5*(1), 3–18.

Singer, J. (2020). Border patrol: The rise and role of fact-checkers and their challenge to journalists' normative. *Journalism*. Advance online publication. https://doi.org/10.1177/1464884920933137

Sixto-García, J., López-García, X., & Toural-Bran, C. (2020). Oportunidades para la cocreación de contenidos en los diarios nativos digotales [Opportunities for content co-creation in digital native newspapers]. *El Profesional de la Información*, *29*(4). https://doi.org/

10.3145/epi.2020.jul.26

Smyrnaios, N., Chauvet, S., & Marty, E. (2017). *L'impact de CrossCheck sur les journalistes & les publics* [The impact of CrossCheck on the journalists & the publics]. London: First Draft. Retrieved from https://fr.firstdraftnews.org/wp-content/uploads/sites/3/2017/12/Crosscheck_rapport_FR_1202.pdf

Stencel, M. (2015). *"Fact check this": How U.S. politics adapts to media scrutiny*. Reston, VA: American Press Institute. Retrieved from http://www.americanpressinstitute.org/fact-checking-project/fact-checking-research/u-s-politics-adapts-media-scrutiny

Stencel, M., & Luther, J. (2020, June 22). Annual census finds nearly 300 fact-checking projects around the world. *Duke Reporters Lab*. Retrieved from https://reporterslab.org/annual-census-finds-nearly-300-fact-checking-projects-around-the-world

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153.

Toporoff, S. (2016, December 8). The journey of a news product: From hackathon to the newsroom. *GEN Editors Lab*. Retrieved from https://medium.com/editors-lab-impact/the-case-for-hackathons-3283edbd9f82

Trevisan Fossá, M. I., & Müller, K. A. (2019). Crosscheck as a legitimization strategy of the journalism field in response to fake news. *Brazilian Journalism Review*, *15*(3), 430–451.

Trusted News Initiative (TNI) steps up global fight against disinformation with new focus on US presidential election. (2020, July 13). *BBC*. Retrieved from https://www.bbc.co.uk/mediacentre/latestnews/2020/trusted-news-initiative

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

Wardle, C. (2017, February 16). Fake news: It's complicated. *First Draft*. Retrieved from https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79

Wardle, C. (2018). The need for smarter definitions and practical, timely empirical research on information disorder. *Digital Journalism*, *6*(8), 951–963.

Wardle, C., Pimenta, A., Conter, G., Días, N., & Burgos, P. (2019). An evaluation of the impact of a collaborative journalism Project on Brazilian journalists and audiences. *First Draft*. Retrieved from https://firstdraftnews.org/wp-content/uploads/2019/06/Comprova-Summary-Report-English.pdf?x91514

Westlund, O., & Hermida, A. (in press). Data journalism and misinformation. In H. Tumber & S. Waisbord (Eds.), *Handbook on media misinformation and populism*. London: Routledge.

Wiltshire, V. (2019). Keeping it local: Can collaborations help save local public interest journalism? (Reuters Institute Fellowship Paper). Oxford: Reuters Insti-

tute for the Study of Journalism. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-09/Vanessa_Wiltshire_JF_paper_keeping_it_local.pdf

Wishart, E. (2018). A wire agency journalist's perspective on 'fake news.' *Media Asia*, *44*(2), 102–106.

Young, D. G., Jamieson, K. H., Poulsen, S., & Goldring, A. (2018). Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck.org and FlackCheck.org. *Journalism and Mass Communication Quarterly*, *95*(1), 49–75.

About the Authors

Bella Palomo is Full Professor at the University of Malaga (Spain). Palomo has focused her line of research on digital journalism during the last two decades. She has been Visiting Scholar at the Universities of Washington, Rutgers, Miami (US), and Federal de Bahia (Brazil). She is leading the project 'The Impact of Disinformation in Journalism: Contents, Professional Routines and Audiences.'

Jon Sedano is Researcher at the School of Communication of the University of Malaga (Spain). His main research interests are journalistic innovations and fact-checking. His teaching is related to photojournalism. Expert in multimedia journalism and new narratives, Sedano's professional projects have been published in Vocento newspaper group, as well as in *El País*.

Article

Fact-Checking Interventions as Counteroffensives to Disinformation Growth: Standards, Values, and Practices in Latin America and Spain

Victoria Moreno-Gil [1], Xavier Ramon [2,]* and Ruth Rodríguez-Martínez [2]

[1] Nebrija University, 28015 Madrid, Spain; E-Mail: vmorenog@nebrija.es
[2] Department of Communication, Pompeu Fabra University, 08018 Barcelona, Spain; E-Mails: xavier.ramon@upf.edu (X.R.), ruth.rodriguez@upf.edu (R.R.-M.)

* Corresponding author

Abstract

As democracy-building tools, fact-checking platforms serve as critical interventions in the fight against disinformation and polarization in the public sphere. The Duke Reporters' Lab notes that there are 290 active fact-checking sites in 83 countries, including a wide range of initiatives in Latin America and Spain. These regions share major challenges such as limited journalistic autonomy, difficulties of accessing public data, politicization of the media, and the growing impact of disinformation. This research expands upon the findings presented in previous literature to gain further insight into the standards, values, and underlying practices embedded in Spanish and Latin American projects while identifying the specific challenges that these organizations face. In-depth interviews were conducted with decision-makers of the following independent platforms: *Chequeado* (Argentina), *UYCheck* (Uruguay), *Maldita.es* and *Newtral* (Spain), *Fact Checking* (Chile), *Agência Lupa* (Brazil), *Ecuador Chequea* (Ecuador), and *ColombiaCheck* (Colombia). This qualitative approach offers nuanced data on the volume and frequency of checks, procedures, dissemination tactics, and the perceived role of the public. Despite relying on small teams, the examined outlets' capacity to verify facts is noteworthy. Inspired by best practices in the US and Europe and the model established by *Chequeado*, all the sites considered employ robust methodologies while leveraging the power of digital tools and audience participation. Interviewees identified three core challenges in fact-checking practice: difficulties in accessing public data, limited resources, and the need to reach wider audiences. Starting from these results, the article discusses the ways in which fact-checking operations could be strengthened.

1. Introduction

The normative theories of media highlight that socially responsible journalism should provide truthful, complete, and accurate information to help citizens understand public affairs and, having been informed, to participate in the community (Christians, Glasser, McQuail, Nordenstreng, & White, 2009). Nevertheless, journalism's ability to comply with this normative goal is increas-

ingly threatened by a combination of profound challenges and provocations. Digital transformations have come at a cost for legacy media, which "have suffered from the collapse of the traditional advertising-funding model combined with dwindling circulation numbers" (Ramon & Tulloch, 2019, p. 2). In addition, the expansion of 'ASAP journalism' (Usher, 2018) news culture has diminished opportunities for thorough investigations, source-checking, and verification, thus limiting

the proper operation of journalism's watchdog function (Zelizer, 2018). Furthermore, the growth of practices linked to commodification, such as click-baiting, have challenged journalism's traditional norms and values. The fixation on metrics has also led news organizations and other actors to stimulate "polarized over moderate views" (McCluskey & Kim, 2012, p. 566). In a context characterized by political instability and the rise of populism, increasing partisanship is arguably problematic, since polarized media systems can also lead to polarized societies (Allcott & Gentzkow, 2017).

The disturbing spread of disinformation unmistakably poses the latest "existential challenge to journalists dealing with an audience losing its faith in what journalism does" (Richardson, 2017, p. 1). The rise of disinformation can be attributed, among other factors, to the politicization of the media, citizens' distrust towards institutions, and the psychological biases and social rewards that drive individuals to share fake news (García-Marín, 2020). Given that citizens are increasingly employing mobile devices and social media to access content (Newman, 2020), the spread of disinformation through digital platforms "can eventually lead to false beliefs or factual misperceptions, posing vexing problems on democratic decision-making" (Hameleers & van der Meer, 2020, p. 230).

Graves and Anderson (2020) consider fact-checking as a strand of contemporary journalism while Amazeen (2020, p. 98) refers to fact-checking platforms as interventions that appear "when a threat is perceived." Precisely, as democracy-building tools, fact-checking operations serve as critical interventions in the fight against the expansion of false and/or misleading news. As part of a global movement, these entities scrutinize the claims of public representatives and alert citizens to online disinformation. In the process, they attempt to "revitalize the 'truth seeking' tradition in journalism" (Graves, 2016, p. 6).

Internal fact-checking originated in the US during the first decades of the twentieth century. However, the rise and consolidation of modern fact-checking platforms is much more recent (Amazeen, 2020). The creation in 2003 of *FactCheck.org* at the Annenberg Public Policy Center at the University of Pennsylvania paved the way for the appearance in 2007 of *The Washington Post*'s *Fact Checker* and *PolitiFact.com*, a non-profit project developed by the *Tampa Bay Times* (then, the *St. Petersburg Times*) which is now operated by the Poynter Institute for Media Studies. Fact-checking initiatives rapidly made inroads into other countries, being introduced by organizations such as *Channel 4* (2005), *Libération* (2008), and *Le Monde* (2009). The growth of disinformation related to political elections and events such as the Brexit referendum facilitated the expansion of fact-checking operations, which have gained visibility and legitimacy over the years (Lowrey, 2017).

Graves (2018) emphasizes that fact-checking is a global movement. In 2015, the Poynter Institute estab-

lished the International Fact-Checking Network (IFCN) to bring together fact-checkers from around the world. Besides organizing an annual Global Fact-Checking Summit, the IFCN promoted a code of principles signalling its signatories' commitment to: (1) non-partisanship and fairness; (2) standards and transparency of sources; (3) transparency of funding and organization; (4) standards and transparency of methodology; and (5) an open and honest corrections policy. The Reporters' Lab at the Sanford School of Public Policy at Duke University notes that there are currently 290 active fact-checking sites in 83 countries (Stencel & Luther, 2020). These organizations are remarkably diverse but can be classified under the two models described by Graves and Cherubini (2016): the 'newsroom model' and the 'NGO' model. The first refers to fact-checking units operating within established news organizations; the second to initiatives that do not belong to newsrooms. Over the last few years, digital technologies have lowered production barriers (Singer, 2018), enabling new figures to enter the fact-checking scene. This includes non-profit independent projects, platforms linked to NGOs or developed by or in collaboration with universities, such as the *RMIT ABC Fact Check* in Australia (Farrer, 2017).

In spite of this diversity, the movement is characterized by its "shared discourse and overlapping practices" (Graves, 2018, p. 614). An essential trait of fact-checking is the embodiment of scientific objectivity to overcome the 'he-said/she-said' reporting style that has pervaded contemporary journalism practice. The systematic fact-checking process involves the following steps: (1) selecting statements of public interest; (2) identifying evidence and context to scrutinize the accuracy of those claims; and (3) writing and publicizing assessments (UNESCO, 2018). Singer's (2020) interviews with fact-checkers on four continents revealed that fact-checkers consider accuracy, impartiality, accountability, objectivity, independence, transparency, and completeness as essential cornerstones of their work. Fact-checkers also perceive that their task is "not only a complement but also a corrective for mainstream media," especially in territories where media are "relatively weak," "servile" and "strong on spreading fake news and spin" (Singer, 2020, p. 9).

Fact-checking projects are run by small teams of journalists who leverage technology to develop and disseminate their work (Graves, 2018). Notably, professionals embrace verification tools such as TinEye, Google Reverse Search, and FotoForensics, which assist them in the "process of authenticating online content items such as text, images, and videos" (Brandtzaeg, Følstad, & Chaparro Domínguez, 2018, p. 1110). Online options have also allowed fact-checking projects to disseminate verifications through textual and multimedia elements (Vázquez-Herrero, Vizoso, & López-García, 2019). Digital platforms have also allowed these projects to cultivate close relationships with their audiences (Singer, 2018). According to fact-checkers, serving audiences not only

implies reaching them but also educating them through media literacy (Singer, 2020).

Previous studies have indicated the potential of fact-checking to counter disinformation and political polarization (Hameleers & van der Meer, 2020). Researchers have also highlighted the challenges that hinder the development of fact-checking, such as limited visibility, resource and time constraints, and the incipient development of machine learning (Humprecht, 2020; Lowrey, 2017; Molina, Sundar, Le, & Lee, 2019). Research on pioneering platforms in the US context paved the way for subsequent analyses focused on other territories such as the UK, Germany, Austria, Ukraine, and sub-Saharan Africa (Cheruiyot & Ferrer-Conill, 2018; Haigh, Haigh, & Kozak, 2018; Humprecht, 2019; Singer, 2020).

2. Fact-Checking in the Latin American and Spanish contexts

Following the US model, Latin American and Spanish fact-checking platforms started operating in 2010 with the launch of *Chequeado* (Argentina), which set an example for many subsequent outlets regarding methodology and workflow. Historically, Spain and Latin America have had remarkable similarities, such as commentary-oriented journalism, low levels of newspaper circulation, instrumentalization, and limited journalistic autonomy (Hallin & Papathanassopoulos, 2002). Traditionally, the Spanish landscape has been characterized by a high level of polarization in the public sphere and the political classes' strong influence over journalism (Baumgartner & Bonafont, 2015; Masip, Ruiz, & Suau, 2018). Along with other problems such as job precariousness, Spanish journalists remain deeply concerned by the politicization of the media, its dependence on institutional sources, and pressure from government (Luengo, Maciá-Barber, & Requejo-Alemán, 2017; Mauri-Rios, López-Meri, & Perales-García, 2020).

In the same vein, Latin America is a region that "has traditionally faced serious obstacles to achieving genuine media democracy" (Palau-Sampio, 2018, p. 352). First, journalists' access to public data has been widely restricted. According to Saldaña and Mourão (2018, p. 319), "despite the approval of laws granting access to information in most countries, practice is still limited by uneven implementation and bureaucratic delay tactics." Second, journalistic autonomy has been threatened by the "close ties between government and elite-owned media organizations" (Saldaña & Mourão, 2018, p. 311). According to Freedom House and Reporters Without Borders, high media concentration and close links between the media's ownership and the political class in the region results in self-censorship and the slackening of media pluralism. Intimidation, harassment, and violence against journalists who cover sensitive topics are not limited to countries such as Colombia since it also occurs in other territories such as Argentina, Uruguay, and Chile. In Ecuador, media freedom has improved since

Lenin Moreno became president in 2017, but local officials and authorities "are still responsible for implementing assaults on the press through legislative, judicial, and administrative means" (Freedom House, 2020). In Brazil, attacks have intensified on journalists and outlets critical of Jair Bolsonaro, who was elected in 2018 "after a campaign marked by hate speech, disinformation, violence against journalists and contempt for human rights" (Reporters Without Borders, 2020).

Disinformation is a growing concern both in Latin America and Spain. As the Reuters Institute Digital News Report highlights, disinformation is a major threat in those countries "where social media use is high and traditional institutions are often weaker" (Newman, 2020, p. 17). According to the Reuters Institute survey, politicians are seen as most responsible for spreading disinformation, followed by political activists and journalists. Political strife and the high levels of polarization in countries such as Argentina, Brazil, Chile, or Spain have facilitated the circulation of false and misleading news, mostly throughout social media platforms and messaging apps.

In the face of such shared challenges, independent fact-checking platforms have flourished across Latin America and Spain (Stencel & Luther, 2020). Several exploratory studies have offered an overview of the organizational structures and methods deployed by Spanish-language fact-checking sites (Bernal-Triviño & Clares-Gavilán, 2019; López-Pan & Rodríguez-Rodríguez, 2020; Palomo & Sedano, 2018; Rodríguez-Pérez, 2020; Vizoso & Vázquez-Herrero, 2019). Other contributions have focused on the task carried out by independent fact-checkers during political campaigns (Chaves & Braga, 2019; Magallón-Rosa, 2019; Vizoso & López-García, 2019). Combining quantitative and qualitative approaches, Palau-Sampio (2018) examined nine fact-checking projects in six Latin American countries to analyse their workflows as well as the topics and actors that are on their agendas. The study revealed that those sites prioritize political and social issues, especially those that intersect with education, health, and the economy. Recently, Ufarte-Ruiz, Anzera, and Murcia-Verdú (2020) compared fact-checking outlets in Spain (*Maldita.es* and *Newtral*) and Italy (*Pagella Politica*). The study revealed that these fact-checking platforms verify political statements and viral content circulating on social media while continuously personalizing content in a bid to adapt to market demand and consumers' tastes. Their business model completely differs from legacy media in terms of revenue sources.

Through in-depth interviews, this article expands upon the findings presented in previous literature to gain further insight into the standards, values, and underlying practices embedded in Spanish and Latin American projects while identifying the specific challenges currently facing these organizations. This qualitative study is informed by the perspectives of decision-makers from eight active organizations in seven countries: (*Chequeado*, *UYCheck*, *Maldita.es*,

Newtral, *Fact Checking*, *Agência Lupa*, *Ecuador Chequea*, and *ColombiaCheck*).

3. Method

The objective of this research was to expand both the theoretical and practical understanding of how fact-checking is performed by the following independent platforms across seven different countries in Latin America and Spain: *Chequeado* (Argentina), *UYCheck* (Uruguay), *Maldita.es* and *Newtral* (Spain), *Fact Checking* (Chile), *Agência Lupa* (Brazil), *Ecuador Chequea* (Ecuador), and *ColombiaCheck* (Colombia).

The aforementioned independent non-profit organizations are integrated into the Reporters' Lab database at Duke University and were purposively selected due to their trajectory and visibility within their home countries. Seven of the platforms under consideration are members of the *LATAM Chequea* network (https://chequeado.com/proyectos/latam-chequea) and, with the exception of *UYCheck*, are verified signatories of the IFCN code of principles (Table 1). Following the criteria established by Humprecht (2020), *Agência Lupa* was also considered as a project within the 'NGO model' (Graves & Cherubini, 2016). The platform is now hosted on the *Piauí* magazine website but was established and operates as an independent project (Palau-Sampio, 2018). To allow space for a broader range of perspectives within this model, a leading project launched in the higher education environment (*Fact Checking*, created by the Pontifical Catholic University of Chile) was also incorporated into the sample. Three research questions guided this study:

RQ1: How many people work in these projects? What is their agenda and volume of publication?

RQ2: Which procedures and resources are employed by these organizations? How do they disseminate fact-checks and what role does the audience play in the process?

RQ3: What are the major challenges faced by these platforms and how do editors evaluate the impact of fact-checking to counter disinformation?

Eight in-depth interviews were conducted with the decision-makers representing these platforms: Laura Zommer (*Chequeado*), Clara Jiménez (*Maldita.es*), Eliana Álvarez (*UYCheck*), Enrique Núñez (*Fact Checking*), Gabriel Narváez (*Ecuador Chequea*), Pablo Medina (*ColombiaCheck*), and Natália Leal (*Agência Lupa*), along with the editorial team of *Newtral*. The first six interviews, ranging between 60 and 90 minutes in length, were conducted between June and July 2019 via Skype. Conversations were audio recorded for subsequent transcription and analysis. The last two interviews were conducted via email. Questionnaires were sent to *Agência Lupa* and *Newtral* during the aforementioned period but were returned to the authors in February 2020. *Newtral* answered the questionnaire considering the viewpoints of the editorial team as a whole. Therefore, quotes are not attributed to a single editor or decision-maker, as is the case with the rest of interviewees.

Interviews with decision-makers have been employed in recent studies on fact-checking (Graves & Anderson, 2020; Palomo & Sedano, 2018; Singer, 2018, 2020). In-depth interviews allow researchers to get a closer perspective towards a part of the object of study that cannot be approached through content analysis, providing them with breadth and depth of nuances and details arising from first-hand descriptions. As a tool for qualitative interviews, Skype allows researchers "to transcend geographical boundaries" (Lo Iacono, Symonds, & Brown, 2016, p. 3), thus enabling them to broaden the range and diversity of initiatives that can be examined. In addition, "with the use of VoIP technologies for interviews, time can be used more flexibly, around the needs of participants, while retaining synchronicity with the interviewer" (Lo Iacono et al., 2016, p. 5).

The conversational scripts contained a total of 28 questions, which pivoted around the following areas arising from the research questions: (1) description of

Table 1. Characteristics of the fact-checking platforms examined in the study.

Project	Website	Country	Creation	Staff	LATAM Chequea	IFCN signatory
Chequeado	https://chequeado.com	Argentina	2010	30	Yes	Yes
UYCheck	http://uycheck.com	Uruguay	2014	8–9	Yes	No
Fact Checking	https://factchecking.cl	Chile	2013	8	No	No
Agência Lupa	https://piaui.folha.uol.com.br/lupa	Brazil	2015	10	Yes	Yes
Ecuador Chequea	http://www.ecuadorchequea.com	Ecuador	2016	3	Yes	Yes
ColombiaCheck	https://colombiacheck.com	Colombia	2016	8	Yes	Yes
Maldita.es	https://maldita.es	Spain	2014	15	Yes	Yes
Newtral	https://www.newtral.es	Spain	2018	8	Yes	Yes

projects and personnel; (2) agenda, volume, and frequency of checks; (3) fact-checking procedures and routines employed; (4) dissemination of fact-checks; (5) role of the public; and (6) editors' opinions on the challenges they face and the impact of fact-checking platforms as counteroffensives to disinformation. In order to properly design the interview guides (Kvale & Brinkmann, 2009, p. 97), we collected information offered by these organizations on their websites. This task aimed to design a solid script that would include key elements regarding the verification process and the resources employed as well as lesser-known aspects that would provide crucial information regarding fact-checking practice, its perceived impact, and the main concerns according to fact-checkers.

The interviews were transcribed and analysed employing the constant comparison technique (Wimmer & Dominick, 2013). Qualitative data arising from interview transcripts were assigned to the six specific themes of the study, allowing for the comparison and contrast of the material. Interview transcription and coding were assessed by all the authors to ensure the completeness and trustworthiness of data (Janesick, 2015).

4. Results

4.1. Description of Projects and Personnel

As a pioneering project in the Latin American landscape, *Chequeado* was founded in 2010 as a reaction to the absence of a space for factual information in mainstream media. The willingness to contribute to democratic wellbeing is a fundamental value shared by all the projects analysed. Some of these initiatives specifically emerged with the aim of monitoring political speech during episodes of great importance, such as the 2014 Uruguayan presidential elections (*UY Check*), the vote on the peace agreement with the FARC in 2016 (*ColombiaCheck*), or the 2017 presidential and legislative election campaigns in Ecuador (*Ecuador Chequea*). In the case of *Maldita.es*, the proliferation of disinformation following the Catalan independence referendum on October 1, 2017, was decisive in reinvigorating *Maldita Hemeroteca*, an initiative launched in 2014 by journalists Julio Montes and Clara Jiménez.

Fact-checking organizations in the Latin American and Spanish sphere are inspired by best practice in the US and Europe. The editors agreed that they consider *PolitiFact.com* as a role model, but also mentioned *The Washington Post's Fact Checker* and *FactCheck.org* (US); *Channel 4* and *FullFact* (UK); and *Le Monde* and *Liberation* (France). Notably, decision-makers at *Colombia Check*, *Ecuador Chequea*, *Maldita.es*, *UYCheck*, and *Agência Lupa* considered *Chequeado* as a fundamental reference: "they very kindly shared their method with us and gave us the initial training and we have been adapting the method," Gabriel Narváez pointed out.

These organizations maintain close contact with other projects, illustrating the way in which fact-checking

is understood as an international community of practice. Six projects are signatories of the IFCN Code of Principles (see Table 1). Interestingly, in 2014, *Chequeado* promoted the LATAM Chequea network. In their first meeting, journalists from 17 organizations were able to share their experiences and the tools used to verify information. Today, this network consists of 22 participating organizations from 15 countries, including 6 of the initiatives examined. Members of this alliance have recently developed a platform to combat disinformation on Covid-19 (https://www.chequeado.com/latamcoronavirus).

The teams that make up these projects are generally small (see Table 1). As regards to *Fact Checking* (Chile), the journalism students conducting the verifications, which range between 20 and 55 per group, are supervised by two editors, two lecturers, and four teaching assistants. Most teams rely mainly on journalists. At *Maldita.es*, all members have undergone training in journalism, while other projects present a more multidisciplinary team. *Newtral* has journalists with diverse profiles, with training in economics or even in law. *UYCheck*'s team is formed by professionals with backgrounds in political science, sociology, psychology, journalism, economic development, and graphic design. For its part, *Chequeado* is composed of journalists, economists, engineers, political scientists, and sociologists.

4.2. Agenda, Volume, and Frequency of Checks

According to the decision-makers interviewed, the projects examined address a wide range of topics, among which politics, economics, education, science, and health prevail. The particular political context of each country determines the agenda: For example, in *Colombia Check*, the verifications on the peace agreement have a significant weight in their output while the Catalan independence referendum in 2017 was crucial for the launch of *Maldita.es* in Spain.

Despite operating with small teams, the capacity of these projects to verify facts is noteworthy (see Table 2). Although the volume of publication can be quite different depending on the platform, most of them complete an average of, at least seven verifications per week, except *UYCheck*, which publishes two or three per week, and the Spanish platforms (*Maldita.es* and *Newtral*), which produce and average of 25 weekly verifications. The shared criteria employed to select content is news relevance: of the person, of the topic, or of the repercussion/virality that the issue may have. The statements from institutional sources have a fundamental weight in the agenda. As Gabriel Narváez emphasizes:

> When public representatives work with incorrect figures that do not reveal factual realities, what ends up happening is that the public policies that are built around these discourses are also incorrect…it is essential for us to carry out surveillance work.

Table 2. Volume of publication.

Project	Number of fact-checks per week
Chequeado	10–15
UYCheck	2–3
Fact Checking	60–70 (per semester)
Agência Lupa	10–15
Ecuador Chequea	7
ColombiaCheck	10–13
Maldita.es	25
Newtral	25

4.3. Procedures and Resources for Debunking Disinformation

The fact-checking projects examined employ transparent and robust methodologies to guarantee a consistent verification process. The methodology established by *Chequeado* is shared by other initiatives, which follow the same workflow with small nuances (Figure 1).

The number of categories used to classify fact-checks ranges between four and nine (see Table 3). According to Gabriel Narváez, having a limited range of categories allows fact-checking platforms "to be specific and not open the scope too much in order to adequately convey the result to users." This framework is adapted to address disinformation on the Internet and social media. Since in many cases it is not possible to contact the original source, online tools and other methodologies such as interviews are used to verify data.

To carry out political fact-checks, all projects consult official and alternative sources: international organizations, civil society organizations, consultants, foundations, experts and researchers, and scientific literature. Depending on the topic or access difficulties, these consultations are made at the same time "as you don't know who will answer you first" (Laura Zommer). Organizations

have clearly identified their most widely used sources (see Table 4), with the exception of *Newtral* and *Agência Lupa*, which does not have such a list because they use different sources depending on the type of information they are aiming to verify.

Interviewees admitted that their relationships with the sources they check is generally respectful. As Pablo Medina stresses, "some of them do not like it very much, but most political sources are used to it and end up answering us." A good practice shared by these projects is to guarantee the right of reply. According to Laura Zommer, "if they answer us with data and arguments that prove them right, we obviously correct. And if not, we ignore it." Clara Jiménez points out that "nuances have been added to certain fact-checks from talking to the politician."

In most initiatives, it is common for one person to carry out each verification, with the exception of more complex issues. Before publishing, verifications go through various control filters (Table 5), ensuring the rigour and trustworthiness of the process. As a general rule, these projects do not establish a maximum time for verification: some fact-checks appear even three months after claims have been made. Although some topics may begin to be researched but eventually not be finalized,

1. Selecting a claim from the public sphere and weighing its relevance

2. Contacting the original source

3. Consulting official and alternative sources

4. Situating the claim in a broader context

5. Rating the claim according to the established categories and publishing it
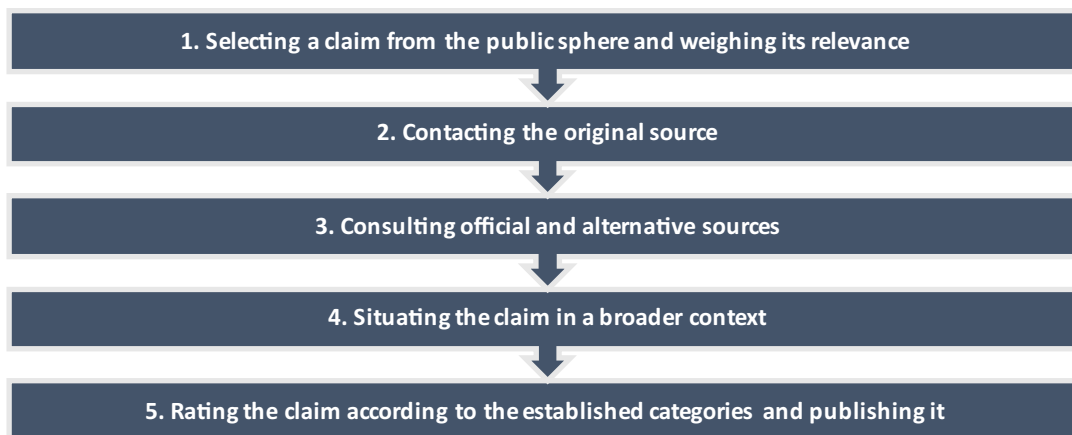
Figure 1. Fact-checking workflow. Source: Authors' elaboration with information from *Chequeado*'s website (Chequeado, 2020) and interviewees' responses.

COGITATIO

Table 3. Categories employed by fact-checking platforms.

Project	Categories	Definition of categories
Chequeado	9	*Inchequeable* (unverifiable), *verdadero** (true), *verdadero…pero* (true…but), *discutible* (disputable), *apresurado* (hasty), *exagerado* (exaggerated), *engañoso* (deceptive), *insostenible* (untenable), *falso** (false).
UYCheck	7	*Verdadero** (true), *verdad a medias* (half true), *inflado* (inflated), *ni ni* (neither nor), *engañoso* (deceptive), *falso** (false), *ridículo* (ridiculous).
Fact Checking	6	*Creíble** (credible), *creíble pero** (credible, but), *sería creíble pero* (would be credible, but), *se puso creativ@* (creative), no es creíble (not credible), *ciencia ficción* (science fiction).
Agência Lupa	9	*Verdadeiro* (true), *verdadeiro mas* (true, but), *ainda é cedo para dizer* (early to say), *exagerado** (exaggerated), *contraditório* (contradictory), *subestimado* (understated), *insustentável* (untenable), *falso** (false), *de olho* ("we are watching").
Ecuador Chequea	4	*Cierto* (true), *sí pero** (yes, but), *insostenible* (untenable), *falso** (false)
ColombiaCheck	5	*Verdadero* (true), *verdadero pero* (true, but), *cuestionable** (questionable), *falso** (false), *inchequeable* (unverifiable).
Maldita.es	6	Political fact-checking: *Falso** (false), *verdadero pero* (true, but), *falso pero* (false, but); Disinformation: *Bulo* (hoax), *qué sabemos* (what we know) *y no hay pruebas* (no proof).
Newtral	4	*Verdadero* (true), *verdad a medias* (half true), *engañoso** (deceptive), *falso** (false).

Note: The most-used categories by each fact-checking platform are marked with an asterisk.

there is widespread awareness among editors that "it is very possible that they will return to the agenda" (Laura Zommer).

Regarding technological resources, fact-checking projects make intensive use of open-access digital tools: search engines (reverse image search in Google, Yandex, and Tineye), tools aimed at verifying videos and photos (InVID, FotoForensics) and map verification applications (Google Maps, Wikimapia, OpenStreetMaps, Yandex Maps, Baidu Maps, Naver Maps). Other platforms such as *Chequeado*, *Agência Lupa*, and *Colombia Check* use CrowdTangle, a Facebook verification tool that helps monitor the virality of fake content on social networks. This 'arsenal' of possibilities is completed, in the case of *Chequeado*, with 'Chequeador,' an in-house bot that "reads 30 Argentinian newspapers every day, the speeches of the president, and everything that happens in the congress" (Laura Zommer). In all cases,

Table 4. Most-employed sources.

Project	Most-employed sources
Chequeado	National Institute of Statistics and Censuses (INDEC); Ministry of Labour; CIPPEC (independent nonprofit organization on public policies); Ministry of Economy; National Social Security Administration (ANSES).
UYCheck	National Institute of Statistics (INE); Ministry of the Interior; Central Bank of Uruguay; The World Bank; Ministry of Social Development.
Fact Checking	Governments' and sub-secretaries' websites; National Congress of Chile website; Transparency portal.
Agência Lupa	Varies according to the type of information that is being checked.
Ecuador Chequea	Central Bank of Ecuador; National Institute of Statistics and Censuses (INEC); Ministries' websites; Twitter accounts from the National Assembly and the Constitutional Court of Ecuador; Transparency websites created by the government.
ColombiaCheck	National Civil Registry; Misión de Observación Electoral (NGO); National Department of Statistics (DANE); Ministry of the Interior; Office of the Attorney General of Colombia; KROC Institute for International Peace Studies.
Maldita.es	National Institute of Statistics (INE); Eurostat; Ministry of Labour.
Newtral	Varies according to the type of information that is being checked.

Table 5. Number and description of filters employed by fact-checking platforms.

Project	Filters	Description of filters
Chequeado	4	Reporter, additional journalist, two editors
UYCheck	3	Reporter, two editors
Fact Checking	3	Reporter, two editors
Agência Lupa	2	Reporter, editor
Ecuador Chequea	3	Reporter, additional journalist, editor
ColombiaCheck	2	Reporter, editor
Maldita.es	4	Reporter, three editors
Newtral	4	Reporter, additional journalist, two editors

the fact-checkers themselves employ these tools, as no special training or background in computer science is required to use them.

4.4. Dissemination of Fact-Checks

To ensure the broadest dissemination possible, fact-checks are published both on the projects' websites and social media accounts. Some verifications also appear in other outlets, such as *Eldiario.es*, which has a section for *Maldita.es*, or the *Teletrece* programme, which welcomes verifications conducted by *Fact Checking*. For its part, *UYCheck* collaborates with the VTV programme "En la Mira" and has also worked with other Uruguayan newspapers such as *El País*. These experiences demonstrate the mainstream media's growing interest in fact-checking, a task that many outlets cannot assume due to their structures, professional routines, or political ties. As Eliana Álvarez explains:

> Media find it interesting and they wish they could do this job, but many times they cannot do it. They show us confidence in our work: We can be wrong in the data, but nobody distrusts that we are playing the political game on anyone.

Independence is not always well-received, as evidenced by the case of *Ecuador Chequea*. The project has agreements with two radio stations and with the *Criterios* magazine of the Quito Chamber of Commerce. However, its critical distance from power has limited its opportunities in other spaces, as Gabriel Narváez explains:

> Between November 2018 and January 2019, we had an agreement with the Ecuadorian public media and our fact-checks were published in the newspaper *El Telégrafo*—one of the newspapers with the largest circulation nationwide—*Radio Pública FM*, and *Ecuador TV*. We are trying to renew that contract, I believe it generated impact…however, not everyone agrees that the media should openly say that the president is a liar.

4.5. Role of the Public

Audiences' participation is essential to these projects. Myriad opportunities are offered to foster close relationships with citizens, including email, postal mail, forms on websites, and social media (WhatsApp, Twitter, Facebook, and Instagram). Through these channels, audiences send comments and materials and suggest verifications. They also provide criticism or corrections to published fact-checks. Clara Jiménez highlights that new channels facilitate citizens' involvement in the fight against disinformation: "We have solved a lot of things using our own community. If citizens get involved in the fight against lies, they contribute to the viralization of truth."

According to the interviewees, fact-checks proposed by audiences play a crucial role in these projects: two to three weekly checks at *Chequeado* are proposed by the public. In *Agência Lupa*, approximately 10% of the monthly production starts from public suggestions, while in *Ecuador Chequea* these topics represent between 20–30% of its production. For its part, *Newtral* developed a Verification on Demand application for WhatsApp, where the team replied to 6,100 messages between July 2018 and the end of 2019. In some cases, such as in *Colombia Check*, limited structures complicate participation management, as Pablo Medina explains:

> We have a WhatsApp number and Twitter and Facebook accounts and we check them periodically. At the moment we do not make much publicity of these channels because we do not have someone who can constantly review and organize all the requests that come to us.

4.6. Editors' Opinions on the Challenges and the Impact of Fact-Checking Platforms

These organizations acknowledge they face several problems and challenges. The first challenge shared across projects refers to accessing information. As Laura Zommer mentions:

Sometimes there is no updated data for all phenomena, or sometimes data is not public. We only work with data that we can see or contrast, not with what somebody tells us. Our contract with audiences does not necessarily imply that citizens should believe us, they should be able to check data by themselves.

The second core challenge faced by fact-checking projects refers to the scarcity of resources. As Natália Leal points out, "there is a financial problem common to all journalistic initiatives and for companies like ours, it is something that weighs heavily." Most teams are made up of part-time employees or volunteers, which limits the volume and range of issues that can be addressed. A third crucial challenge, in the eyes of Enrique Núñez, is delivering fact-checking to mass audiences:

Fact-checking is an elite product. I am especially concerned about this in Latin American countries, which are countries of oral culture, of late literacy, and where television is still the most consumed medium. To be massive in these countries, fact-checking must think about that oral culture. Fact-checking must be on television and radio programmes.

Going forward, interviewees defended the crucial role of fact-checking as a counter-offensive to disinformation and as a tool that "helps to create a civic spirit on how information should be consumed" (Eliana Álvarez). Despite the aforementioned difficulties, the interviewees were optimistic about the future of fact-checking. As Pablo Medina highlights:

There are going to be fewer mainstream media and more small media. Fact-checking platforms are going to be part of that universe as a very particular way of doing journalism. Fact-checking is what journalism should be doing but has stopped doing for many reasons: because of the desire to get clicks and because there is not enough budget to hire people to verify content.

5. Discussion and Conclusion

In a time characterized by profound challenges, fact-checking interventions represent an alternative destination in journalism that seeks to combat the spread of disinformation, educate citizens, and contribute to restoring the credibility of journalism. This article aims to contribute to the existing literature on fact-checking by paying attention to key aspects regarding this phenomenon through the analysis of eight verification initiatives in Latin America and Spain.

The research builds on previous works focused in those territories (Bernal-Triviño & Clares-Gavilán, 2019; Chaves & Braga, 2019; López-Pan & Rodríguez-Rodríguez, 2020; Magallón-Rosa, 2019; Palau-Sampio, 2018; Palomo & Sedano, 2018; Rodríguez-Pérez, 2020;

Ufarte-Ruiz et al., 2020; Vizoso & López-García, 2019; Vizoso & Vázquez-Herrero, 2019) by offering a deep understanding of the standards, values, and practices while identifying the specific challenges that eight independent organizations in Argentina, Brazil, Chile, Colombia, Ecuador, Spain, and Uruguay currently face. Through qualitative interviews, this work contributes to the fact-checking literature by offering decision-makers' insight on the volume and frequency of checks, procedures, and the perceived role of the public. This contribution also sheds light on fact-checkers' opinions on the impact of their work and the difficulties they have to overcome. The novelty of this work is that it combines the study of several fact-checking projects in two regions which, despite their differences, have shown remarkable similarities in terms of their media landscapes and ongoing journalistic challenges.

Findings indicate that the outlets examined in Latin American and Spanish contexts carry out an intensive task. The projects analysed were founded as a reaction to the absence of space for factual information in the mainstream media and with the aim of monitoring political speech during important political episodes (such as presidential elections, a peace process, or a referendum) when disinformation is more likely to arise, circulate, and potentially harm democratic decision-making.

Indeed, despite relying on small teams, their capacity to verify facts is noteworthy: Most of the Latin American sites examined offer a minimum of seven fact-checks per week while the Spanish ones complete up to 25 weekly verifications. In addition to this, the majority of the interviewees assure that they do not establish a maximum time for verification. This implies that, even when the latest news prevails, fact-checking differs from mainstream journalism practice since it breaks the traditional concept of urgency associated with news reporting (Zelizer, 2018).

All platforms shared fixed criteria regarding content selection: news relevance due to the person, the topic or the repercussion, and/or virality on mainstream and social media. Decision-makers point out that they address different topics in which the analysis of public representatives' claims has a fundamental weight in the agenda. Inspired by best practice in the US and Europe and the model established by *Chequeado*, all sites considered employ innovative, robust, and transparent methodologies that guarantee a consistent verification process. These organizations consult a broad range of official and alternative sources to check data while leveraging the power of digital tools. Notably, audience participation is considered critical to help locate and count disinformation circulating through myriad platforms.

Overall, the fact-checking platforms examined fulfil a function that many organizations in Latin America and Spain have halted due to constrained resources, hyper-accelerated cycles, and the impact of commodification on news production and distribution (Usher, 2018). They occupy a distinctive domain within journalism

by reclaiming the importance of accuracy, independence, rigour, contextualization, completeness, and transparency, identified as normative values of journalism (Christians et al., 2009). Professional values highlighted by the interviewees are consistent with those identified as essential by other fact-checkers across different continents (Singer, 2020). With their approach, they also seek to transcend some of the specific limitations and constraints shared across the Spanish and Latin American media systems, such as limited journalistic autonomy, high levels of polarization, and the politicization of the media (Baumgartner & Bonafont, 2015; Freedom House, 2020; Hallin & Papathanassopoulos, 2002; Luengo et al., 2017; Masip et al., 2018; Mauri-Rios et al., 2020; Saldaña & Mourão, 2018).

Interviewees identified three core challenges in fact-checking practice: difficulties in accessing public data, limited resources, and the need to reach wider audiences. These major concerns are widely shared across different journalistic cultures (Amazeen, 2020; Humprecht, 2020; Lowrey, 2017; Singer, 2020). There are different ways in which the outlets analysed try to respond to these challenges. First, they expose the lack of transparency or problems in accessing public data in the territories they operate (Saldaña & Mourão, 2018) by doing the exact opposite: offering data to readers in the most transparent way. Secondly, they overcome the lack of resources by creating multidisciplinary and highly qualified teams while sharing knowledge between platforms. Finally, they constantly work to broaden their audience through social media and by establishing collaborations with mainstream outlets.

Looking forward, fact-checking operations could be strengthened and expanded by fostering cooperation. The fact-checking movement is making steady progress thanks to initiatives such as the IFCN and LATAM Chequea networks. Further collaboration between organizations can help fact-checkers research more complex topics. Capitalizing on journalists' increasing interest in participating in national and transnational collaborative projects (Cueva-Chacón & Saldaña, 2020), fact-checkers can materialize partnerships that help to reduce production costs, share content across platforms and facilitate the dissemination of verifications in different countries. Enhancing a continuous dialogue between fact-checking and mainstream media organizations is of utmost importance. The experiences developed by sites such as *Agência Lupa* or *Newtral* reveal that agreements with news outlets help give broader visibility to fact-checkers' output while providing them with additional income. Partnerships with universities represent a supplementary source of revenue. Initiatives such as verification courses offered by *Newtral* and *Maldita.es* indicate that there is a growing need for alliances between newsrooms and academia. Fact-checking organizations can help universities by assisting them in designing and introducing verification modules to curricula, thus contributing to strengthening future journalists' verification skills.

Moreover, fact-checking institutions can contribute to society at large by encouraging media literacy activities. The organization of meetings, workshops, and training sessions can promote critical thinking among citizens so that they can make informed decisions about the content they consume. According to Ana Pastor, founder of *Newtral*, "citizens should be able to fact-check by themselves and also to fact-check us" (Tardáguila, 2019).

The results of this qualitative study should be seen in light of its limitations. The findings cannot be generalized to all organizations from the independent/NGO model. Yet, information from seven different countries helps to advance true understandings of fact-checking practice in Latin America and Spain. To broaden the scope of this analysis, future research should monitor the evolution of fact-checking interventions in the Latin American and Spanish contexts while trying to understand the practices used by other outlets. The incipient collaboration between fact-checking platforms and major companies such as Google and Facebook (Graves & Anderson, 2020) should also be closely followed. Those actors can contribute to funding fact-checking platforms, assist them in the development of tools for automated fact-checking and help them reach broader audiences. Fact-checking platforms can also leverage their expertise to verify the accuracy of content on those platforms. Recent developments in this field include the participation of *Chequeado*, *Newtral* and *Maldita.es* in Facebook's data verification programme. In addition, future studies should investigate how fact-checkers operate under particularly challenging contexts, such as during the Covid-19 crisis.

Another avenue for future research revolves around gender. In contrast to other areas in journalism, gender does not seem to be a setback in fact-checking practice. In fact, many platforms at the international level are driven by women. While this field "can serve as an example for other areas" (Eliana Álvarez), the extent to which women's entrepreneurship in fact-checking is a response to the structural gender divide that permeates the sector should be interrogated.

Finally, the involvement of audiences also lends opportunities for future examination. Focus groups and interviews would allow a much more nuanced understanding of citizens' experiences and expectations with fact-checking platforms. Audiences' perspectives will also allow researchers to identify additional ways in which those sites could be improved to continue contributing to democratic wellbeing.

Acknowledgments

by the Spanish Ministry of Science and Innovation and the State Research Agency (FACCTMedia, PID2019—106367GB-I00/AEI/10.13039/501100011033).

Conflict of Interests

The authors declare no conflict of interests.

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Amazeen, M. A. (2020). Journalistic interventions: The structural factors affecting the global emergence of fact-checking. *Journalism*, *21*(1), 95–111. https://doi.org/10.1177/1464884917730217

Baumgartner, F. R., & Bonafont, L. C. (2015). All news is bad news: Newspaper coverage of political parties in Spain. *Political Communication*, *32*(2), 268–291. https://doi.org/10.1080/10584609.2014.919974

Bernal-Triviño, A., & Clares-Gavilán, J. (2019). Uso del móvil y las redes sociales como canales de verificación de fake news: El caso de Maldita.es [Using mobile devices and social networks as checking channels of fake news: The case of Maldita.es]. *El Profesional de La Información*, *28*(3), 1–8. https://doi.org/10.3145/epi.2019.may.12

Brandtzaeg, P. B., Følstad, A., & Chaparro Domínguez, M. Á. (2018). How journalists and social media users perceive online fact-checking and verification services. *Journalism Practice*, *12*(9), 1109–1129. https://doi.org/10.1080/17512786.2017.1363657

Chaves, M., & Braga, A. (2019). The agenda of disinformation: "Fake news" and membership categorization analysis in the 2018 Brazilian presidential elections. *Brazilian Journalism Research*, *15*(3), 474–495. https://doi.org/10.25200/BJR.v15n3.2019.1187

Chequeado. (2020). Método [Method]. *Chequeado*. Retrieved from https://chequeado.com/metodo

Cheruiyot, D., & Ferrer-Conill, R. (2018). "Fact-checking Africa": Epistemologies, data and the expansion of journalistic discourse. *Digital Journalism*, *6*(8), 964–975. https://doi.org/10.1080/21670811.2018.1493940

Christians, C. G., Glasser, T. L., McQuail, D., Nordenstreng, K., & White, R. A. (2009). *Normative theories of the media: Journalism in democratic societies*. Urbana, IL: University of Illinois Press.

Cueva-Chacón, L. M., & Saldaña, M. (2020). Stronger and safer together: Motivations for and challenges of (trans)national collaboration in investigative reporting in Latin America. *Digital Journalism*. Advance online publication. https://doi.org/10.1080/21670811.2020.1775103

Farrer, G. (2017). Fact check unit an opportunity for a new kind of J. education. *Asia Pacific Media Educator*, *27*(1), 10–16. https://doi.org/10.1177/1326365X17702269

Freedom House. (2020). *Freedom in the world 2020: A leaderless struggle for democracy*. Washington, DC: Freedom House. Retrieved from https://freedomhouse.org/report/freedom-world/2020/leaderless-struggle-democracy

García-Marín, D. (2020). Infodemia global: Desórdenes informativos, narrativas fake y fact-checking en la crisis de la Covid-19 [Global infodemic: Information disorders, false narratives, and fact checking during the Covid-19 crisis]. *El Profesional de la Información*, *29*(4), e290411. https://doi.org/10.3145/epi.2020.jul.11

Graves, L. (2016). *Deciding what's true: The rise of political fact-checking in American journalism*. New York, NY: Columbia University Press.

Graves, L. (2018). Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, *19*(5), 613–631. https://doi.org/10.1080/1461670X.2016.1196602

Graves, L., & Anderson, C. W. (2020). Discipline and promote: Building infrastructure and managing algorithms in a "structured journalism" project by professional fact-checking groups. *New Media and Society*, *22*(2), 342–360. https://doi.org/10.1177/1461444819856916

Graves, L., & Cherubini, F. (2016). *The rise of fact-checking sites in Europe*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/research/files/The%2520Rise%2520of%2520Fact-Checking%2520Sites%2520in%2520Europe.pdf

Haigh, M., Haigh, T., & Kozak, N. I. (2018). Stopping fake news: The work practices of peer-to-peer counter propaganda. *Journalism Studies*, *19*(14), 2062–2087. https://doi.org/10.1080/1461670X.2017.1316681

Hallin, D., & Papathanassopoulos, S. (2002). Political clientelism and the media: Southern Europe and Latin America in comparative perspective. *Media, Culture and Society*, *24*(2), 175–195. https://doi.org/10.1177/016344370202400202

Hameleers, M., & van der Meer, T. G. L. A. (2020). Misinformation and polarization in a high-choice media environment: How effective are political fact-checkers? *Communication Research*, *47*(2), 227–250. https://doi.org/10.1177/0093650218819671

Humprecht, E. (2019). Where "fake news" flourishes: A comparison across four Western democracies. Information *Communication and Society*, *22*(13), 1973–1988. https://doi.org/10.1080/1369118X.2018.1474241

Humprecht, E. (2020). How do they debunk "fake news"? A cross-national comparison of transparency in fact checks. *Digital Journalism*, *8*(3), 310–327. https://doi.org/10.1080/21670811.2019.1691031

Janesick, V. J. (2015). Peer debriefing. In G. Ritzer (Ed.), *The Blackwell encyclopedia of sociology*). Lon-

don: John Wiley & Sons. https://doi.org/10.1002/9781405165518.wbeosp014.pub2

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Thousand Oaks, CA: Sage.

Lo Iacono, V., Symonds, P., & Brown, D. H. K. (2016). Skype as a tool for qualitative research interviews. *Sociological Research Online*, *21*(2), 103–117. https://doi.org/10.5153/sro.3952

López-Pan, F., & Rodríguez-Rodríguez, J. (2020). El fact checking en España: Plataformas, prácticas y rasgos distintivos [Fact checking in Spain: Journalistic projects, practices and distinctive features]. *Estudios sobre el Mensaje Periodístico*, *26*(3), 1045–1065. https://doi.org/10.5209/esmp.65246

Lowrey, W. (2017). The emergence and development of news fact-checking sites: Institutional logics and population ecology. *Journalism Studies*, *18*(3), 376–394. https://doi.org/10.1080/1461670X.2015.1052537

Luengo, M., Maciá-Barber, C., & Requejo-Alemán, J. L. (2017). Evaluating organisational ethics in Spanish news media. *Journalism*, *18*(9), 1142–1162. https://doi.org/10.1177/1464884916643682

Magallón-Rosa, R. (2019). Verificado México 2018: Desinformación y fact-checking en campaña electoral [Verificado Mexico 2018: Disinformation and fact-checking on electoral campaign]. *Revista de Comunicación*, *18*(1), 234–258. https://doi.org/10.26441/RC18.1-2019-A12

Masip, P., Ruiz, C., & Suau, J. (2018). *Monitoring media pluralism in Europe: Application of the Media Pluralism Monitor 2017 in the European Union, FYROM, Serbia & Turkey. Country Report: Spain*. Florence: Centre for Media Pluralism and Media Freedom. Retrieved from https://cmpf.eui.eu/wp-content/uploads/2018/11/Spain_MPM2017_country-report.pdf

Mauri-Rios, M., López-Meri, A., & Perales-García, C. (2020). The professional ethics of journalists in the face of governments and politicians: Perceptions of professionals and citizens in Spain. *Revista Latina de Comunicación Social*, *77*, 295–308. https://doi.org/10.4185/RLCS-2020-1459

McCluskey, M., & Kim, Y. M. (2012). Moderatism or polarization? Representation of advocacy groups' ideology in newspapers. *Journalism & Mass Communication Quarterly*, *89*(4), 565–584. https://doi.org/10.1177/1077699012455385

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2019). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*. Advance online publication. https://doi.org/10.1177/0002764219878224

Newman, N. (2020). *Executive summary and key findings of the 2019 report*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from http://www.digitalnewsreport.org/survey/2019/overview-key-findings-2019/

Palau-Sampio, D. (2018). Fact-checking and scrutiny of power: Supervision of public discourses in new media platforms from Latin America. *Communication and Society*, *31*(3), 347–365. https://doi.org/10.15581/003.31.3.347-365

Palomo, B., & Sedano, J. (2018). WhatsApp como herramienta de verificación de fake news: El caso de B de Bulo [WhatsApp as a verification tool for fake news: The case of B de Bulo]. *Revista Latina de Comunicación Social*, *2018*(73), 1384–1397. https://doi.org/10.4185/RLCS-2018-1312

Ramon, X., & Tulloch, C. (2019). Life beyond clickbait journalism: A transnational study of the independent football magazine market. *Communication & Sport*. Advance online publication. https://doi.org/10.1177/2167479519878674

Reporters Without Borders. (2020). *World press freedom index 2020*. Paris: Reporters Without Borders. Retrieved from https://rsf.org/en/ranking

Richardson, N. (2017). Fake news and journalism education. *Asia Pacific Media Educator*, *27*(1), 1–9. https://doi.org/10.1177/1326365X17702268

Rodríguez-Pérez, C. (2020). Una reflexión sobre la epistemología del fact-checking journalism: retos y dilemas [A reflection on the epistemology of fact-checking journalism: challenges and dilemmas]. *Revista de Comunicación*, *19*(1), 243–258. https://doi.org/10.26441/rc19.1-2020-a14

Saldaña, M., & Mourão, R. R. (2018). Reporting in Latin America: Issues and perspectives on investigative journalism in the region. *The International Journal of Press/Politics*, *23*(3), 299–323. https://doi.org/10.1177/1940161218782397

Singer, J. B. (2018). Fact-checkers as entrepreneurs: Scalability and sustainability for a new form of watchdog journalism. *Journalism Practice*, *12*(8), 1070–1080. https://doi.org/10.1080/17512786.2018.1493946

Singer, J. B. (2020). Border patrol: The rise and role of fact-checkers and their challenge to journalists' normative boundaries. *Journalism*. Advance online publication. https://doi.org/10.1177/1464884920933137

Stencel, M., & Luther, J. (2020). Annual census finds nearly 300 fact-checking projects around the world. *Duke Reporters' Lab*. Retrieved from https://reporterslab.org/annual-census-finds-nearly-300-fact-checking-projects-around-the-world

Tardáguila, C. (2019). Ana Pastor, in Spain: 'Citizens should be able to fact-check by themselves and also to fact-check us.' *Poynter.org*. Retrieved from https://www.poynter.org/fact-checking/2019/ana-pastor-in-spain-citizens-should-be-able-to-fact-check-by-themselves-and-also-to-fact-check-us

Ufarte-Ruiz, M. J., Anzera, G., & Murcia-Verdú, F. J. (2020). Plataformas independientes de fact-checking en España e Italia: Características, organización y método [Independent fact-checking platforms in Spain and Italy: Features, organisation and method]. *Revista Mediterránea de Comunicación*, *11*(2), 23–39.

https://doi.org/10.14198/MEDCOM2020.11.2.3

UNESCO. (2018). *Journalism, fake news & disinformation: Handbook for journalism education and training*. Paris: UNESCO. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000265552

Usher, N. (2018). Breaking news production processes in US metropolitan newspapers: Immediacy and journalistic authority. *Journalism*, 19(1), 21–36. https://doi.org/10.1177/1464884916689151

Vázquez-Herrero, J., Vizoso, Á., & López-García, X. (2019). Innovación tecnológica y comunicativa para combatir la desinformación: 135 Experiencias para un cambio de rumbo [Technological and communicative innovation to fight misinformation: 135 Experiences for a change of direction]. *El Profesional de La Información*, 28(3), e280301. https://doi.org/10.3145/epi.2019.may.01

Vizoso, Á., & López-García, X. (2019). Newtral y Comprobado: Experiencias de fact-checking durante la campaña electoral de las Elecciones Generales en España [Newtral and Comprobado: Fact-checking experiences during the electoral campaign of the Spanish General Elections]. *Cuadernos Artesanos de Comunicación*, 171, 77–98. https://doi.org/10.4185/cac171

Vizoso, Á., & Vázquez-Herrero, J. (2019). Fact-checking platforms in Spanish: Features, organisation and method. *Communication and Society*, 32(1), 127–142. https://doi.org/10.15581/003.32.1.127-142

Wimmer, R. D., & Dominick, J. R. (2013). Mass media research: An introduction. Belmont, CA: Cengage Learning.

Zelizer, B. (2018). Epilogue: Timing the study of news temporality. *Journalism*, 19(1), 111–121. https://doi.org/10.1177/1464884916688964

About the Authors

Victoria Moreno-Gil (PhD) is a Lecturer at the School of Communication and Arts of Nebrija University. She holds a PhD in Communication from the Carlos III University and a MA degree in Ethics and Law in journalism from the UNED. She has been a Visiting Researcher at Pompeu Fabra University. Her research focuses on journalism ethics, media accountability, and media literacy. She combines a research career with professional journalism. She currently works for the European Pressphoto Agency (EPA).

Xavier Ramon (PhD) is a Lecturer at the Department of Communication of Pompeu Fabra University. Member of Journalism Research Group (GRP). His research and teaching focus on journalism ethics, media accountability, and sports communication. He has been a Visiting Researcher at the University of Stirling and the University of Glasgow (UK), the University of Alabama (US), and the IOC Olympic Studies Centre (Switzerland). He has participated in three competitive projects on media accountability and fact-checking.

Ruth Rodríguez-Martínez (PhD) is a Lecturer at the Department of Communication at Pompeu Fabra University. A member of the Journalism Research Group (GRP), her research focuses on media ethics, fact-checking, and cultural journalism. She is one of the principal investigators of the funded projects MediaACES and FACCTMedia. She has carried out postdoctoral stays at the Missouri University School of Journalism and the Columbia University in New York (US) and the Laboratoire du Comunicacion et Politique at CNRS (France).

Article

Debunking Political Disinformation through Journalists' Perceptions: An Analysis of Colombia's Fact-Checking News Practices

Carlos Rodríguez-Pérez [1,*], Francisco J. Paniagua-Rojano [2] and Raúl Magallón-Rosa [3]

[1] Department of Journalism, University of Ibagué, 730002 Ibagué, Colombia; E-Mail: carlos.rodriguez@unibague.edu.co
[2] Department of Journalism, University of Malaga, 29071 Malaga, Spain; E-Mail: fjpaniagua@uma.es
[3] Department of Communication, Carlos III University of Madrid, 28903 Getafe, Spain; E-Mail: raul.magallon@uc3m.es

* Corresponding author

Abstract

Fact-checking alliances emerged worldwide to debunk political disinformation in electoral contexts because of social concerns related to information authenticity. This study, thus, included the Latin American context in fact-checking journalism studies as a journalistic practice to fight political disinformation. Through analyzing RedCheq, the first fact-checking journalism alliance in an electoral regional context led by Colombiacheck, 11 in-depth interviews were conducted to identify the perceptions of regional fact-checkers regarding the usefulness of this journalistic practice, its achievements, and the key aspects for incorporating fact-checking into the regional media ecosystem. The study results revealed that RedCheq achieved the goal of fighting disinformation, and that fact-checking developed as transformational leverage for the regional media. Regional journalists perceived fact-checking as an element that restores credibility and social trust in regional media as the epistemology of this journalistic practice neglects the power pressure and dissemination of official narratives. Finally, this study highlighted how fact-checking journalism contributes to the democratic quality and civic empowerment in silenced and polarized environments. In addition, it discussed the need to expand fact-checking journalism's coverage to new geographical areas and improve journalists' professional competencies and training, thereby enabling them to function as using verification tools based on regional journalists' requirements.

Issue

This article is part of the issue "Disinformation and Democracy: Media Strategies and Audience Attitudes" edited by Pere Masip (University Ramon Llull, Spain), Bella Palomo (University of Málaga, Spain) and Guillermo López (University of Valencia, Spain).

1. Introduction

This study analyzes the fact-checking RedCheq collaborative project, which is the first alliance of checkers in Colombia that was established to fight disinformation prevailing in the Colombian regional elections in 2019. Online disinformation campaigns threaten fair elections, which are core processes for achieving the democratic good of accountable political representation (Tenove, 2020). The initiative was developed in August 2019 with the participation of journalists, media, universities, and civil society organizations from across the country. These participants attended training workshops to be able to verify content and information during the electoral campaigns; limit the collateral effects of fake and deceitful information related to candidates, political parties, and processes and party platforms; and thoroughly monitor and examine regionally coded electoral public discourses.

The RedCheq collaborative initiative is based on other existing initiatives launched by First Draft, such as CrossCheck in France, Electionland in the US, Checkpoint

in India, and Comprobado in Spain. Similar initiatives at the state level have already been established in Latin America in countries such as Mexico (Verificado; Magallón-Rosa, 2019), Brazil (Comprova), Argentina (Reverso), and Uruguay (Verificado). RedCheq was headed by Colombiacheck, a digital, open, and collaborative fact-checking platform signatory of the Code of Principles of the International Fact-Checking Network and Consejo de Redacción, a nonprofit organization that counts among its ranks more than 100 associated journalists in Colombia to foster research journalism and contribute to the democratization of information.

In Colombia, the exercise of the journalist profession suffered the consequences of the armed conflict. Cartografías de la información, conducted by Fundación para la Libertad de Prensa (2019), a nonprofit organization that defends press freedom and the practice of journalism, presents the following alarming conclusion: 578 out of the 994 analyzed municipalities correspond to silenced zones, i.e., places where there is no media outlet to produce local journalism, and 29% of the population resides in such places. Gentzkow, Shapiro, and Sinkinson (2011) stated that media outlets positively influence political participation, particularly when the first outlet is established.

A recent study on the working conditions in Colombian regional journalism (Valencia-Nieto, Pereira-González, & Rodríguez-Leuro, 2019) that was based on 139 interviews depicts regional journalism as a poor trade where journalists are subjected to power pressures and disseminate official narratives; therefore, regional journalism is at its critical levels of credibility because it is unable to examine power thoroughly. Consequently, its contribution to democratic quality is insignificant.

Colombia had already experienced a disinformation context during the October 2nd, 2016 referendum on peace agreements. As Pablo Medina-Uribe (2018) recalls:

> During the autumn of 2016, a prominent politician admitted having deliberately deceived the audience before one of the most important elections in the history of the country. By doing so, he confirmed what many citizens already knew: WhatsApp, a digital communication platform, which is now widely used, is a fertile ground for group manipulation.

Further, Colombiacheck and Consejo de Redacción guaranteed the regional journalists' training by organizing 10 workshops in Casanare, Santander, Cesar, Tolima, Caldas, and Valle del Cauca, among others, which ended in a Chequeatón in Bogota (August 13th and 14th, 2019), and established RedCheq. In addition, the fact-checking process was supervised, edited, and sponsored using the Colombiacheck methodology.

The methodological process implemented by RedCheq comprised six steps, namely, to select a phrase (content) from the public sphere due to its relevance and impact, to consult with the author of the phrase, to compare the information and data with official and reliable sources and then with alternative and expert sources, to contextualize the checking in terms of time and socioeconomic or political aspects, and to qualify the statement (true, slightly true, questionable, fake, or uncheckable).

Further, this study incorporates the Latin American context to fact-checking journalism studies through in-depth interviews of journalists within the network. As academic attention is mainly focused on the English-speaking countries, it is imperative for new approaches to be developed in order to define and compare the fact-checking and disinformation systems (Hallin & Mancini, 2004). Vázquez-Herrero, Vizoso, and López-García (2019) analyzed 135 fact-checking initiatives from across the world and concluded that most of the fact-checking platforms are European and American native digital media; they found that their regional coverage scope is only 9%. At the beginning of May 2020, it was estimated that 237 teams devoted to fact-checking were active worldwide, with 26 teams in South America: Argentina (2), Bolivia (1), Brazil (10), Chile (2), Colombia (3), Ecuador (1), Peru (2), Uruguay (3), and Venezuela (2; Duke Reporters' Lab, 2020). Only 78 of these fact-checking journalistic organizations were signatories of the Code of Principles of the International Fact-checking Network.

1.1. Colombian Political Context

The first regional elections since the peace agreements with the Revolutionary Armed Forces of Colombia in 2016 were held on October 27th, 2019. According to the National Civil Registry Office's (2019) data, Colombians elected a total of 1,101 mayors, 32 governors, 12,063 city councilors, 418 representatives, and 6,814 elected officials. To this end, 3,306 elections were held to vote for local authorities. The main participation data revealed that 60.65%, which equals 22,1 million of Colombians based on the census of 36,5 million people, exercised their right to vote in order to nominate their mayor. Participation in the election of the departmental governors was 61%, with 18,8 million of votes being cast on the basis of a census of 30,8 million of Colombians.

The Electoral Observation Mission (2019) witnessed the progress made in the exercise of the right to vote—thanks to the "conclusive progress in terms of safety related to noninterference by the armed groups in the elections…civic engagement, electoral debate, and the results of authorities in chasing electoral crimes." However, this aspect does not conceal the irregularities detected during the previous months such as electoral transhumance or fraud committed on the election day.

Colombia conducted the regional elections in a political context marked by the tribulation and a crisis between the government and the governed, with the latter motivated by the lack of integrity and efficiency of public management. The political culture sur-

vey conducted between July and August in 2019 by the National Statistics Office revealed that 80.7% of Colombians consider living in a democracy to be very important, but 47.1% exhibited profound dissatisfaction with its functioning. In the Colombian political context, corruption is one of the main reasons for strife in the country. Municipal governments are perceived to be corrupt or very corrupt by 61.2%, and a very similar percentage is obtained for the departmental governments (60.4%). However, both these percentages are less than the percentage obtained for the national government (64%). These percentages reflect a major concern regarding the political system because a setback can be observed between 2017 and 2019 in the urban population, which demonstrates robust confidence in city halls (−4.5%), governance (−8.8%), and departmental assemblies (−2%; National Statistics Office, 2019).

2. Literature Review

After the Brexit electoral processes, the Colombian referendum, and the US presidential elections in 2016, there is no doubt that disinformation strategies are effective and may imperil the prospects of democracy in any country of the world. Moreover, fake news is not a new phenomenon; it has always existed. However, the current situation fosters a higher global scope and impact because the media, through which the fake news is disseminated, and the speed of dissemination also favor it (Burkhardt, 2017). Focused on the political disinformation, Valenzuela, Halpern, Katz, and Miranda (2019) stated that political commitment is an important consequence of news consumption through social networks and that it entails a key precedent when it comes to sharing fake news. Therefore, disinformation emerges as a problem for democracies, particularly during the electoral times and referendums (European Commission, 2018; Newman, Fletcher, Kalogeropoulos, & Kleis-Nielsen, 2019).

Fregoso (2018) states that, with the peak of independent digital media, several readers have begun to doubt the quality and impartiality of the information being published through newspapers; however, for millions of new young readers, the printed information seems to be no longer relevant. New technologies have modified the social behaviors related to producing and consuming information within the field of journalism because the changes related to the digitalization of communication outlets and the dissemination of erroneous information challenge the social role and authority of journalism (Ekström, Lewis, & Westlund, 2020). These changes and the information obesity in social media need to be examined to understand the relationship between the use of news and trust in news (Kalogeropoulos, Suiter, & Eisenegger, 2019). This aspect broadens the horizon to determine the factors that aid or inhibit resilience in the society regarding the disinformation phenomenon (Humprecht, Esser, & Van Aelst, 2020).

The academic literature has offered several approaches and classifications for disinformation and related concepts such as misinformation and fake news. Fake news has become a catch-all, buzzword, or mainstream word to include all misleading information. Tandoc, Lim, and Ling (2018), based on a review of how previous studies had operationalized the term fake news, identified two dimensions (facticity and deception) and they established six types of fake news: news satire, news parody, fabrication, manipulation, advertising, and propaganda. In the same line, but through contrasting the concept of fake news with real news, Molina, Sundar, Le, and Lee (2019) identified seven types of fake news: false news, polarized content, satire, misreporting, commentary, persuasive information, and citizen journalism. There is a debate about the inclusion or exclusion of the satire content as a typology of disinformation due to a satire content which does not have a deceitful intention to mislead, although may exert the function of misleading (Fallis, 2015; Molina et al., 2019; Tandoc et al., 2018).

However, some authors are avoiding the usage of the term fake news to define the information pollution phenomenon. The contention lies on the epistemology and the better accuracy of disinformation to refer to the information disorder. Fake news does not hold an ordinary public meaning. For instance, its linguistic deficiency, its media delegitimizing meaning and the political propaganda that surrounds it, and the variety of different informative content that not always imitating the aesthetic of news content (Estrada-Cuzcano, Alfaro-Mendives, & Saavedra-Vásquez, 2020; Habgood-Coote, 2019; Rodríguez-Pérez, 2019; Wardle & Derakhshan, 2017; among others). Specifically, fake news is currently a weaponizing term that favors polarization when politicians mention news content that refutes a partisan position, eroding the credibility and trust in news media. In the recent research within the journalism practice, Egelhofer, Aaldering, Eberl, Galyga, and Lecheler (2020) found in Austrian news reporting that the buzzword fake news is associated more often with attacks on legitimate news media and less with the threat of online disinformation.

On the other hand, Wardle and Derakhshan (2017) explained the information disorder which is based on falseness and harmful factors. There are three environments that are distinguished, namely: misinformation, disinformation, and mal-information. Misinformation refers to false information without a harmful intention; disinformation is false information shared to cause harm; and mal-information is genuine information shared to cause harm. The difference between misinformation and disinformation states if there is an active and intentional attempt to mislead (Fallis, 2015). For instance, Fetzer (2004, p. 231) defines misinformation as "false, mistaken, or misleading information" and disinformation as "distribution, assertion, or dissemination of false, mistaken, or misleading information in an intentional, deliberate, or purposeful effort to mislead, deceive, or confuse."

In this context, fact-checking journalism is shaped as a new practice to fight disinformation or as an emerging journalistic genre (Graves, 2018). Fact-checking objectives is not to address electoral-voting behavior because a political actor has deliberately and repeatedly lied; instead, one of its objectives is to enable citizens to exercise their right to vote with the most verified information possible and within a public sphere that is uncontaminated by propaganda and disinformation strategies, which are extremely sensitive to emotional appeal and polarized scenarios.

Researchers investigating fact-checking have analyzed the need for news checkers to consolidate within the media ecosystem and gain the audience's trust (Humprecht, 2019; Lowrey, 2017), the journalists' perception and practices (Cheruiyot & Ferrer-Conill, 2018; Graves, 2017; Mena, 2019), and the effectiveness and social effects of fact-checks (Walter, Cohen, Holbert, & Morag, 2019). Moreover, journalists concur on pointing out the normative aspects of fact-checking and the need for clear methodological guidelines and boundaries between fact-checking and activism because fact-checking should be unbiased (Mena, 2019). For these purposes, an epistemological debate prevails between the preconceived political judgments of fact-checkers and their common access to the objective truth (Uscinski, 2015). However, supported by empirical quantification, Amazeen (2015) exhibited a high level of consistency in fact-verification and concluded that fact-checking journalism has a crucial role to play in fighting disinformation.

3. Objectives and Method

Fact-checking journalism is emerging as a journalistic practice to fight disinformation. Centered around Colombia, RedCheq was established as the first cross-regional alliance to fight disinformation related to regional elections. Therefore, factors such as RedCheq's achievements, its journalistic works, and competencies that are considered necessary for fact-checking by the regional journalists should be explored. Given below are the three specific objectives of this study: (1) Through the regional journalists' viewpoints, analyze how fact-checking journalism is useful in the regional political contexts; (2) highlight the achievements by RedCheq as an alliance to fight political disinformation in the regional elections and its contribution to Colombian society and journalism; and (3) identify the requirements and necessary competencies in regional journalism for the practice of fact-checking journalism.

Further, in-depth interviews provide a complete analytical understanding in order to acquire a specific viewpoint as well as providing a reasoned opinion of the individual work of the regional fact-checking journalists, the operations of RedCheq, and the characterization of disinformation that was disseminated in Colombia during the regional elections. First, we obtained the local journalist sample from the database of Colombiacheck,

a fact-checking media and leader of the project. Then, we assessed the available information provided by classifying journalists according to those who worked freelance and those who were affiliated with regional media outlets. In these latter cases, we interviewed the team leader, who is the designated journalist in charge of fact-checking news developed by the regional media outlet. We interviewed all freelance journalists and team leaders affiliated to local media that participated in RedCheq based on data provided. In all, 11 in-depth interviews, using structured questionnaires through digital communication applications such as Google Meet and WhatsApp, were conducted to the regional journalists who were a part of RedCheq in April 2020 (see Table 1). 7 out of 11 interviewees worked in local media outlets such as print, digital-native, television and radio station located in Antioquia, Casanare, La Guajira, Risaralda, Santander, and Valle del Cauca. 4 out of 11 worked as freelance journalists or were associated with Universities of Caldas, Caquetá, Tolima, and Valle del Cauca. Consequently, and supported by Colombiacheck's headquarter in Bogotá, the regional fact-checking network was in nine Colombian' departments listed as follows: Antioquia, Caldas, Caquetá, Casanare, La Guajira, Risaralda, Santander, Tolima, and Valle del Cauca. This instrument enables to gather, in orderly fashion, the different individual characteristics emphasized by the interviewees and identify patterns through which to deduce the most noticeable trends based on the compared analysis of each one of them can be deduced.

The questionnaire included some control variables (i.e., geographical location, years of professional experience, and first experience as a fact-checker) plus 13 questions divided into three main parts according to our specific research objectives. The first row of questions was about political disinformation in Colombia, its characteristics, and the influence in the electoral context. The second row of questions asked about methods, purposes, and skills to become a fact-checker. Lastly, we asked for a general assessment and further steps for RedCheq. We recorded the interviews using artificial intelligence applications and a proof-reading review was carried out to match the transcription with the statements. We collected all the answer into a single document from which we made comparisons, identified the patterns, and deduced the most relevant arguments according to our specific research objectives.

Questions are listed as follows: (1) How would you describe the disinformation disseminated during the Colombian regional elections? (2) How particularly would you describe your fact-checking work in your regional area? (3) Which reasons motivated you to develop fact-checking journalism? (4) Working as a fact-checker, what was your purpose as a RedCheq journalist? (5) Do you consider, after participating in RedCheq, that your work as a fact-checker allowed obtaining a higher recognition from peers? (6) Considering all disinforma-

Table 1. List of interviewees.

Interviewee ID	Date of interview	Years of professional experience	First experience as a fact-checker
A	April 1st, 2020	19	Yes
B	April 1st, 2020	3	Yes
C	April 4th, 2020	12	Yes
D	April 6th, 2020	12	Yes
E	April 7,th 2020	2	No, previously associated with Colombiacheck during the national elections (2018)
F	April 8th, 2020	18	Yes
G	April 10th, 2020	5	No, previously associated with Colombiacheck during the national elections (2018)
H	April 13th, 2020	4	Yes
I	April 20th, 2020	21	No, previously involved in the fact-checking practices within the nonaccredited fact-checking media
J	April 20th, 2020	6	No, previously associated with Colombiacheck during the national elections (2018)
K	April 25th, 2020	10	Yes

tion contents, what do you consider that has the worst effects on citizens? (7) From your point of view, does the fact-checker prioritize the scrutiny of public discourse, debunks viral disinformation contents or a balance of both? (8) Did you notice your work as a fact-checker at RedCheq contributed to achieving a better public debate? (9) Which criteria did you follow to decide which contents would select for verification? (10) Do you consider your work at RedCheq was free of political interests, social activism or economic pressures? (11) How do you evaluate your skills for developing fact-checking? (12) Do you think RedCheq achieved its goal of debunking disinformation associated with the regional election that took place in October 2019? (13) If you had to repeat the cross-regional network, what would you change? What would you improve?

4. Results

4.1. Achievements

Regional journalists unanimously state that RedCheq fulfilled the objective of fighting disinformation and monitoring the public discourse and malicious use of political information. Their reasons are based on the impact caused by the work on political discourse and public debate. The regional fact-checkers linked the prominent political disinformation to misleading and manipulating strategies, which then polarized and attacked other political candidates. They explained the danger caused by disinformation when these misleading contents influenced or modified citizens' political behavior. Their arguments mentioned the refusal against citizens involved in peace agreements, panic scenarios, and the delegitimization of politicians, among others. This also increased the risk of

WhatsApp chains and how much people trusted in the news they had received on their smartphones.

They perceived that their work was based on the exercise of media literacy related to fact-checking processes in order to develop critical thinking competence in citizens. The study finding demonstrates that regional journalists emphasize that RedCheq contributed toward revealing topics silenced in the media outlets because "several contents were extracted from areas where traditional media was reluctant to enter" (Interviewee J). Furthermore, "RedCheq caused an impact at the national level; even the controlled organizations themselves began to replicate and worked with much input from the network" (Interviewee K). They perceived that the fact-checking exercise contributed toward improving the status and credibility of regional media because it informed about the political and electoral affairs by exercising an oversight counterpower regarding election process guarantees and producing electoral costs when the associated contents were verified to be inaccurate:

> This contributed to the extent that public figures or candidates were obliged to be more careful with their discourse when we performed the fact-checks. Toward the end of the campaign, the discourse was more moderate and that is a part of the contribution made by RedCheq. (Interviewee A)

In addition, the exercise of fact-checking fostered the critical spirit of citizens by providing arguments and data for debates: "Readers used to send e-mails or notifications for us to verify. The debate and desire to debunk fake news were present in that minority audience" (Interviewee B). Articulation and collaboration by fact-checkers and citizens create a renewed process

of citizen journalism (Rodríguez-Pérez, 2020), which, in addition, enables "people to become aware of the seriousness of replicating information that is not verified" (Interviewee E). Consequently, this aspect generates increased prudence and skepticism regarding the information received through social media or instant messaging applications: "People tried to be frugal when receiving those speeches and not to believe everything they heard or read, and even this was important as it ratified that what they were saying was real" (Interviewee F).

Therefore, the interviewees' opinions highlight the importance of fact-checking within the citizens' social concerns about distinguishing between what is real and what is fake and answering this need by empowering them through training and arguments:

The exercise not only involves checking but also examining the manner in which the media contributes to public education and the process by which the public can begin to learn to identify this type of publication themselves so as not to share it, thereby reducing such dissemination. (Interviewee C)

Checking not only involves debunking or revealing the truth but also showing the citizens the elements that are potentially available to them to build autonomy and critical thinking about each leader and public figure, thus providing them with the tools required in order to be involved in the public discourse. (Interviewee G)

For such purposes, the regional journalists consider that fact-checking should be addressed as a majority bet on the Colombian media system. Moreover, they underscore the precedent of this network as a seedbed that will have a national impact:

I believe that all media outlets should focus on how they train journalists in fact-checking. I think that people, i.e., public opinion, are deliberately asking for this. The media cannot turn a blind eye to this issue. (Interviewee C)

RedCheq set a precedent for this type of journalism in our country. It is a seedbed, and if we foster journalists' training in news verification further, this will positively establish the journalism dynamic at the national level. (Interviewee K)

The in-depth interviews exhibit a significant relationship between the exercises developed with media literacy and university training. In particular, one of the regional media that are associated with RedCheq operated as a teaching media outlet and capitalized on the fact-checking practices to instill social values among young people. Another regional journalist instilled these values from the university sphere:

Our news program is a school. In a context where everyone is a content producer, we also forget a very important value, which is important not only for a journalist but also for a citizen: rigor. (Interviewee K)

I wanted to pass that passion for truth to students and wanted them to understand that this is not just about picking a piece of information and saying whether it is true or false but about an amazing passion for truth that is, basically, journalism. (Interviewee G)

4.2. Skills and Training

Most of the interviewees rate their competencies in fact-checking as sufficient or good, although all of them associate the verification process with the ongoing learning that requires constant training. One of the interviewees defined RedCheq as a "network to train journalists'' (Interviewee K). The interviewees emphasize that fact-checking journalism, as a new practice, requires constant training so that journalists can be better equipped to fight disinformation and enhance their knowledge through their experiences:

Platforms change at the same pace as tools to produce fake news. So, what was clearly required to be checked today or yesterday can change tomorrow in such a disguised manner that it may even go unnoticed. (Interviewee J)

This is an ongoing learning exercise. We are capable and willing to do it, but, perhaps, we need to know ways to access information to continue our training and the contexts in which we have to keep on applying this. I would like to have more space for fact-checking training because I think all journalists should be able to do it. (Interviewee K)

The main limitation that the interviewees expressed is related to the competencies in the use of digital tools for checking videos or photographs (multimedia formats):

There are technical aspects that we would like to know more about; for example, ways to identify whether a video is fake or whether the photographs have been manipulated. On a 1-to-10 scale, we score 6–7. (Interviewee C)

We have a long way to go. New tools have been launched, and we have to keep ourselves updated. I have been using 70% of the capacity I have until now. However, we also need the remainder percentage to consider the integrity and protection of journalists. (Interviewee A)

The interviewees state that knowledge acquired through daily practice is another aspect to conduct a successful check; they emphasize that it is an aspect that

is key to develop a journalistic instinct in order to detect disinformation:

> Fact-checking is something new, and because it is a relatively new trend, it is something that we have no experience in at the time of performing it. Besides, this requires that for every refuted piece of news something must be learned on the subject. (Interviewee B)

> Competencies are good, but there is a possibility to improve them continuously. There are some things in which I think I am more competent in than others are (altered images, context of publications, and data and hiring documents). There are other qualifications that I think I can continue to improve such as developing a journalistic instinct so as not to take some things for granted. (Interviewee E)

After the training workshops and the fact-checking practice by RedCheq, most of the interviewees stated a perceived professional recognition from their peers and citizens. The journalistic value of fact-checkers becomes a "lighthouse" (Interviewee A) against the concern on disinformation. It continued after the regional elections when peers and citizens asked for checking contents disseminated in social media. However, and most importantly, they pointed out that fact-checking changed their professional practice, making them more prone to innovation in their newsrooms. They commented on these improvements regarding the use of digital tools, new procedures, and narrative recourses to produce journalism as well as the prevailing ethical standards to assure the accuracy of their own news production.

4.3. Method

RedCheq's methodology was the same implemented by Colombiacheck as interviewees recognized it. The content selection had to verify the political information in its core. From the answers, we identified three main selection criteria elements: (1) The monitoring of leading candidates according to polls; (2) electoral debates and the public discourse; and (3) the category of viral in social media contents to avoid the Streisand effect because of the fact-checking process. Furthermore, the interviewees reckon mostly (10 out of 11) the fact-checker should have to balance the scrutiny of public discourse with the verification of viral contents:

> A political discourse that states against certain democratic value is as dangerous as fake content that becomes viral. The viral content requires to be checked faster, but both should be balanced. (Interviewee H)

> Both bot farming and fake users are a common practice among political parties that use the social networks for amplifying a political message. (Interviewee A)

The dilemma emerged when the interviewees decided about the political statement to be checked. Their answers allow identifying a fear to be perceived as partisan, reported by readers. The political polarization context, as well as the strength or weakness of certain political candidates emerged as the causes that tipped the balance on one political side:

> I felt like it was a problem that others thought that I had political bias, especially if the fact-check made a candidate looks better. (Interview F)

> We wanted to have an equilibrium producing fact-checks in the political spectrum, but we understood that was something we could not force. If a certain candidate did not say anything checkable or there was not anything about a certain candidate because it was not a leading candidate, we could not produce a fact-check. (Interviewee C)

> If we produced ten fact-checks, probably five assessed a misleading against the ex-president Álvaro Uribe's political leaning and the rest distributed among the other parties. Political sphere surrounding Uribe's political party generated disinformation; it was also one of the most mediated and robust electoral campaigns. (Interviewee A)

Besides, one of the boundaries established around fact-checking is the non-partisan practice and the clear distinction between fact-checking and activism. Asking for political and economic pressure as well as social activism, some interviewees declared Colombiacheck editorial board promoted press freedom assuring the non-partisan filter. The economic pressure did not reveal, although one of the interviewees said when they assessed a claim as inaccuracy the advertising from that campaign was reduced. Political pressure happened when candidates were checked:

> There were pressures when the content was close to elites. When a journalist began to ask, the answer suggested not disseminating it….The process is threatened. That also occurs from activism when after producing a fact-check, activists encouraged you finding more false claims. (Interviewee D)

> We felt the higher pressure came from political campaigns. For instance, they said: "You are publishing a lot about us, why do you not assess what others say against us?" (Interviewee C)

On the other hand, other interviewees with less professional experience (Interviewees B, E, and H) admitted what they call "social activism" during their reporting.

They associated this commitment with improving the public debate, understanding that debunking disinformation protected the democratic good of fair elections.

4.4. A Way Forward

The interviewees qualified the coordination and supplementary works of Colombacheck during the whole process as "good" (Interviewee J), "wonderful" (Interview K), or "titanic" (Interviewee C). On the basis of these perceptions, most of the interviewees agree that training is a factor that is particularly vital toward improving an initiative such as this one. To this end, two alternatives are suggested: "Monitoring" (Interviewee K) or "customized accompaniment" (Interviewee J) to improve the quality and quantity of checks, respectively, and "to continue with the trainings on different tools" (Interviewee E).

In addition, and in accordance with the abovementioned viewpoints, journalists agree that the lack of experience and competencies in using fact-checking digital tools led to the delays in the checking process: "Colombiacheck has the experience and training, but issues still remain with regard to optimizing the process and having trained people in the regions so as not to centralize everything in people with technical abilities" (Interviewee C).

RedCheq's internal process included the first proposal shared by the regions with Colombiacheck's national coordination unit or Colombiacheck with regional journalists. After receiving feedback regarding the suggested drafts and specific requests for verification of multimedia content submitted by the regions to Colombiacheck, the fact-check was written down and then reviewed by the Colombiacheck editors to make adjustments or proceed with its posting on the Web. Given that Colombiacheck headed RedCheq, the regional verification network included the seal of the International Fact-checking Network, making it obligatory for them to check on and comply with the international fact-checking methodological process standard.

Therefore, the interviewees perceive the "flow" (Interviewee D), the "internal process speed" (Interviewee F), or the articulation of "the production chain between the journalist performing the check, the editor, and the final piece of news" (Interviewee H) as areas of improvement. Moreover, the only suggestion provided to make up for this aspect is to improve the regional journalists' training.

One of the challenges faced by fact-checking is to try and balance the reach and dissemination of disinformation (Vosoughi, Roy, & Aral, 2018) by answering rapidly and "reducing the time between the release of disinformation and the publication of the check" (Interviewee C). With regard to this, as a third aspect, the interviewees highlight the need for "dissemination improvement" (Interviewee I). They suggest innovation as a solution, along with new narratives that are based on multimedia formats and social network and Web

environment strategies:

> Disinformation could actually be mitigated, but this failed to reach a larger audience or to equal the fake news that was disseminated….To remedy this situation, I would suggest a community manager strategy or something that helps better the position of the website. (Interviewee B)

> I feel that we stick inordinately to the text. The process would be much more dynamic if journalists were trained in the creation of verification content in several formats. They began producing podcasts, but I think we need to be quicker in data visualization and the creation of audiovisual, as well as interactive content. (Interviewee E)

Correspondingly, they suggest that RedCheq should incorporate more media and journalists to cover more municipalities and departments of the country:

> Regional journalists' contribution is important. There are things happening in the other area of Colombia, and it seems like they have been taken out of a science fiction book. Here, we see things that a person living in Bogota cannot even imagine actually happen in the country. (Interviewee A)

> It would have been wonderful to have a responsible from RedCheq present in those municipalities to monitor the public debate underway in the face of elections in that territory. (Interviewee G)

Finally, only one of the interviewees stated the need to modify the verification categories by "making fact-checking more expedited" (Interviewee A) to improve citizens' understanding.

5. Discussion and Conclusions

The zenith of disinformation involves rethinking the news production routines and ways to disseminate journalistic information effectively. As distrust of media outlets and politics prevails in Colombia, finding certain evidence that associates less trust in the news and politics with a higher probability of believing in online disinformation was initiated (Zimmermann & Kohring, 2020).

RedCheq emerged as a collaborative experience, headed by Colombiacheck, for debunking the spread of false political information in the regional areas. This need cannot be addressed without considering the social, labor-related, and economic contexts of the Colombian regional media outlets. While disinformation easily reaches mobile devices, verified and checked information requires the participation of information professionals. This aspect entails their being present and procuring training and independence to conduct good journalism.

This context contrasts with the perceptions of the regional journalists who were interviewed in this study after working for RedCheq. They perceive the fact-checking practice as a wake-up call to modify the performance of regional journalism. The 11 regional journalists who were interviewed agree on underscoring the work conducted by RedCheq to oversee power and make a thorough public examination of the electoral campaign statements. A widespread perception about RedCheq exists in that its reputation precedes it regarding its contribution to improve informed public debate by offering contrasted and unbiased information. However, such improvements were limited regarding scope and regional application. RedCheq was thus shaped as a seedbed for the introduction of fact-checking as a novel journalistic practice in regional journalism.

In summary, the practice of fact-checking journalism by RedCheq has dared to address the following three cleavages. It restores credibility to regional journalism performance by contrasting facts and data in an unbiased and rigorous manner; it contributes to the democratic training of public opinion, which is empowered with contrasted information; and it encourages power auditing. In the political and electoral contexts, journalistic coverage is a key driver of electoral accountability (Snyder & Strömberg, 2008) and the political participation (Gentzkow et al., 2011). A well-informed and mobilized electorate better regulates government corruption and mismanagement practices (Adsera, Boix, & Payne, 2003). Finally, journalism is crucial for civic engagement (Shaker, 2014).

Regarding the professional recognition after completion of RedCheq, Colombian journalists' perceptions coincide with the argument of Graves, Nyhan, and Reifler (2016) that fact-checkers receive recognition and status as upholding the ideals of the profession. On the other hand, our results are slightly different from Mena's results (2019) regarding the perceptions of the regional fact-checkers towards the principal purpose of the fact-checking practice. In contrast to US fact-checkers who prioritized the assessment of statements by public figures and institutions (91.8% strongly agree rather than 57,4% strongly agree with debunking false news stories spread on social media), Colombian fact-checkers affirmed it should be balanced between politicians' statements and viral contents. Moreover, this case study reflects one of the epistemological dilemmas of fact-checking: the boundaries between fact-checking and activism.

Moreover, disinformation, as a worldwide phenomenon, and fact-checking journalism, as a journalistic answer, reveals the need to create a new map of competencies, profiles, and trends. This aspect entails curricular strengthening and an innovative development of the profession. Academicians and communication media are becoming aware of fact-checking journalism as a quality assurance practice and a duty in exercising the profession, thereby necessitating learning new competencies (Ufarte-Ruiz, Peralta-García, & Murcia-Verdú, 2018).

Therefore, it is appropriate to facilitate the acquisition of university knowledge to provide aspiring journalists with new professional opportunities available at small media outlets, which are more independent, diverse, and committed to professional, social, and ethical values. Furthermore, universities should update their curricula to integrate these competencies into digital tools. This aspect would enable aspiring journalists to learn digital forensics techniques to analyze the accuracy of each source of information—from multimedia formats to public statements (Steensen, 2018)—and debunk disinformation content, particularly at a time when an emerging media ecosystem sponsored by fact-checking journalism is envisaged in Latin America to achieve a real media democracy (Palau-Sampio, 2018).

These differential competencies may lead to the creation of new media outlets and the revitalization of the media sector, given the growing civic concern to distinguish between what is true and what is fake. Furthermore, universities may use this opportunity to strengthen their associations with the business sector and civil society by becoming involved in the verification projects, which have been very closely related to the academic environment (Graves & Cherubini, 2016).

To conclude, RedCheq is the first regional fact-checking alliance, a network of regional newsrooms and freelance journalists engaged in delivering fact-checking journalism within the electoral contexts. To this end, a new set of news practices was formulated and established to deliver accountable journalistic practices that offer high quality fact-checks to regional citizens. Based on the journalists' perceptions, we summarize the fact-checking practice experience and training that are the key determinants in making a successful practice out of fact-checking journalism. Regional journalists regard fact-checking as a demanding journalistic practice. Therefore, expertise and technical skills are required to debunk political disinformation, thus increasing the fact-checks' coverage and dissemination and reducing the gap between the disinformation content that becomes viral and the fact-check releases. Moreover, the fact-checking movement should reinforce its scope and broaden its horizons to go beyond the main cities and into new geographical areas. To this end and for the purpose of advancing and reinforcing the fact-checking processes in the regional areas, upgrading journalism and communication networks might be the most suitable way of guaranteeing the acquisition of media competencies by aspiring journalists. All of these factors are aimed at integrating the verification operations in any journalistic process irrespective of the media outlet, geographic location, or content.

Thus, in the context of social and political polarization, along with weak regional journalism that depends on the political, economic, or social powers, regional journalists perceived fact-checking as a catalyst for the awakening of the regional journalism and as a reconnection with the journalist's social responsibility to foster

civic empowerment and to subject political discourse in the electoral campaigns to rigorous examination.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

References

Adsera, A., Boix, C., & Payne, M. (2003). Are you being served? Political accountability and quality of government. *The Journal of Law, Economics, and Organization*, *19*(2), 445–490. https://doi.org/10.1093/jleo/ewg017

Amazeen, M. A. (2015). Revisiting the epistemology of fact-checking. *Critical Review*, *27*(1), 1–22. https://doi.org/10.1080/08913811.2014.993890

Burkhardt, J. M. (2017). Combating fake news in the digital age. *Library Technology Reports*, *53*(8), 1–33. https://doi.org/10.5860/ltr.53n8

Cheruiyot, D., & Ferrer-Conill, R. (2018). Fact-checking Africa. *Digital Journalism*, *6*(8), 964–975. https://doi.org/10.1080/21670811.2018.1493940

Duke Reporters' Lab. (2020). Update: 237 fact-checkers in nearly 80 countries…and counting. *Reporterslab*. Retrieved from https://reporterslab.org/category/fact-checking

Egelhofer, J. L., Aaldering, L., Eberl, J., Galyga, S., & Lecheler, S. (2020). From novelty to normalization? How journalists use the term "Fake news" in their reporting. *Journalism Studies*. https://doi.org/10.1080/1461670X.2020.1745667

Ekström, M., Lewis, S. C., & Westlund, O. (2020). Epistemologies of digital journalism and the study of misinformation. *New Media and Society*, *22*(2), 205–212. https://doi.org/10.1177/1461444819856914

Electoral Observation Mission. (2019). Informe de cierre de la misión de observación electoral MOE [Final report of Election Observation Mission (MOE)]. *MOE*. Retrieved from https://bit.ly/3jn5r9p

Estrada-Cuzcano, A., Alfaro-Mendives, K., & Saavedra-Vásquez, V. (2020). Disinformation y misinformation, posverdad y fake news: Precisiones conceptuales, diferencias, similitudes y yuxtaposiciones [Disinformation and misinformation, post-truth and fake news: Conceptual precisions, differences, similarities and juxtapositions]. *Información, Cultura Y Sociedad*, *42*, 93–106. https://doi.org/10.34096/ics.i42.7427

European Commission. (2018). *Flash Eurobarometer 464: Fake news and disinformation online*. Brussels: European Commission. Retrieved from https://bit.ly/3jrAYHb

Fallis, D. (2015). What is disinformation? *Library Trends*, *63*(3), 401–426. https://doi.org/10.1353/lib.2015.0014

Fetzer, J. H. (2004). Information: Does it have to be true? *Minds and Machines*, *14*(2), 223–229. https://doi.org/10.1023/b:mind.0000021682.61365.56

Fregoso, J. (2018). *#Mexico2018|'fake news' and social media: The new heads of the Hydra*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from https://bit.ly/3b8Oixp

Fundación para la Libertad de Prensa. (2019). Cartografías de la información [Information cartographies]. *FLIP*. Retrieved from https://bit.ly/34FcvKq

Gentzkow, M., Shapiro, J., & Sinkinson, M. (2011). The effect of newspaper entry and exit on electoral politics. *American Economic Review*, 101(7), 2980–3018. https://doi.org/10.1257/aer.101.7.2980

Graves, L. (2017). Anatomy of a fact check: Objective practice and the contested epistemology of fact checking. *Communication, Culture & Critique*, *10*(3), 518–537. https://doi.org/10.1111/cccr.12163

Graves, L. (2018). Boundaries not drawn. *Journalism Studies*, *19*(5), 613–631. https://doi.org/10.1080/1461670x.2016.1196602

Graves, L., & Cherubini, F. (2016). *The rise of fact-checking sites in Europe*. Oxford: Reuters Institute for the Study of Journalism.

Graves, L., Nyhan, B., & Reifler, J. (2016). Understanding innovations in journalistic practice: A field experiment examining motivations for fact-checking. *Journal of Communication*, *66*(1), 102–138. https://doi.org/10.1111/jcom.12198

Habgood-Coote, J. (2019). Stop talking about fake news! *Inquiry*, 62(9/10), 1033–1065. https://doi.org/10.1080/0020174X.2018.1508363

Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge: Cambridge University Press.

Humprecht, E. (2019). How do they debunk "fake news"? A cross-national comparison of transparency in fact checks. *Digital Journalism*, *8*(3), 310–327. https://doi.org/10.1080/21670811.2019.1691031

Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, *25*(3), 493–516. https://doi.org/10.1177/1940161219900126

Kalogeropoulos, A., Suiter, J., & Eisenegger, M. (2019). News media trust and news consumption: Factors related to trust in news in 35 countries. *International Journal of Communication*, *13*, 3672–3693.

Lowrey, W. (2017). The emergence and development of news factchecking sites. *Journalism Studies*, *18*(3), 376–394. https://doi.org/10.1080/1461670X.2015.1052537

Magallón-Rosa, R. (2019). Verificado México 2018: Desinformación y fact-checking en campaña electoral

[Verificado Mexico 201: Disinformation and fact-checking on electoral campaign]. *Revista De Comunicación*, *18*(1), 234–258. http://doi.org/10.26441/RC18.1-2019-A12

Medina-Uribe, P. (2018). In Colombia, a WhatsApp campaign against Posverdad. *The Wilson Quarterly*. Retrieved from https://www.wilsonquarterly.com/quarterly/the-disinformation-age/in-colombia-a-whatsapp-campaign-against-posverdad

Mena, P. (2019). Principles and boundaries of fact-checking: Journalists' perceptions. *Journalism Practice*, *13*(6), 657–672. https://doi.org/10.1080/17512786.2018.1547655

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2019). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*. https://doi.org/10.1177/0002764219878224

National Civil Registry Office. (2019). Autoridades territoriales 2019 [Regional/local authorities 2019]. *National Civil Registry Office*. Retrieved from https://bit.ly/34L4LX6

National Statistics Office. (2019). Encuesta de cultura política (ECP) [Political culture survey (ECP)]. *DANE*. Retrieved from https://www.dane.gov.co/index.php/comunicados-y-boletines/estadisticas-sociales/cultura-politica

Newman, N., Fletcher, R., Kalogeropoulos, A., & Kleis-Nielsen, R. (2019). *Digital news report 2019*. Oxford: Reuters Institute for the Study of Journalism.

Palau-Sampio, D. (2018). Fact-checking and scrutiny of power: Supervision of public discourses in new media platforms from Latin America. *Communication & Society*, *31*(3), 347–365. https://doi.org/10.15581/003.31.3.347-363

Rodríguez-Pérez, C. (2019). No diga fake news, di desinformación: Una revisión sobre el fenómeno de las noticias falsas y sus implicaciones [Do not say fake news, say disinformation: A review article about the fake news phenomenon and its implications]. *Comunicación*, *40*, 65–74. https://doi.org/10.18566/comunica.n40.a05

Rodríguez-Pérez, C. (2020). Una reflexión sobre la epistemología del fact-checking journalism: Retos y dilemas [A reflection on the epistemology of fact-checking journalism: Challenges and dilemmas]. *Revista de Comunicación*, *19*(1), 243–258. https://doi.org/10.26441/rc19.1-2020-a14

Shaker, L. (2014). Dead newspapers and citizens' civic engagement. *Political Communication*, *31*(1), 131–148. https://doi.org/10.1080/10584609.2012.762817

Snyder, J., & Strömberg, D. (2008). Press coverage and political accountability. *Journal of Political Economy*, *118*(2), 355–408. https://doi.org/10.1086/652903

Steensen, S. (2018). Journalism's epistemic crisis and its solution: Disinformation, datafication and source criticism. *Journalism*, *20*(1), 185–189. https://doi.org/10.1177/1464884918809271

Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2018). Defining "fake news": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tenove, C. (2020). Protecting democracy from disinformation: Normative threats and policy responses. *The International Journal of Press/Politics*, *25*(3), 517–537. https://doi.org/10.1177/1940161220918740

Ufarte-Ruiz, M., Peralta-García, L., & Murcia-Verdú, F. (2018). Fact checking: Un nuevo desafío del periodismo [Fact checking: A new challenge in journalism]. *El Profesional De La Información*, *27*(4), 733–741. https://doi.org/10.3145/epi.2018.jul.02

Uscinski, J. E. (2015). The epistemology of fact checking (is still naïve): Rejoinder to amazeen. *Critical Review*, *27*(2), 243–252. https://doi.org/10.1080/08913811.2015.1055892

Valencia-Nieto, D. G., Pereira-González, J. M., & Rodríguez-Leuro, J. A. (2019). Condiciones laborales y sociales de los periodistas en las regiones de Colombia [Labor and social conditions of journalists in the regions of Colombia]. *Mediaciones Sociales*, *18*, 85–105. https://doi.org/10.5209/meso.66907

Valenzuela, S., Halpern, D., Katz, J. E., & Miranda, J. P. (2019). The paradox of participation versus misinformation: Social media, political engagement, and the spread of misinformation. *Digital Journalism*, *7*(6), 802–823. https://doi.org/10.1080/21670811.2019.1623701

Vázquez-Herrero, J., Vizoso, Á., & López-García, X. (2019). Innovación tecnológica y comunicativa para combatir la desinformación: 135 experiencias para un cambio de rumbo [Technological and communicative innovation to fight misinformation: 135 experiences for a change of direction]. *El Profesional de la Información*, *28*(3), 1–12. https://doi.org/10.3145/epi.2019.may.01

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2019). Fact-checking: A meta-analysis of what works and for whom. *Political Communication*, *37*(3), 350–375. https://doi.org/10.1080/10584609.2019.1668894

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe*. Retrieved from https://bit.ly/2ECHDj1

Zimmermann, F., & Kohring, M. (2020). Mistrust, disinforming news, and vote choice: A panel survey on the origins and consequences of believing disinformation in the 2017 German parliamentary election. *Political Communication*, *37*(2), 215–237. https://doi.org/10.1080/10584609.2019.1686095

About the Authors

Carlos Rodríguez-Pérez is Lecturer for Journalism at the Department of Journalism and Social Communication at the University of Ibagué (Colombia) where he teaches Digital Journalism. PhD Candidate at the Complutense University of Madrid (Spain). He holds a Combined Degree in Journalism and Audio-visual Communication by Carlos III University of Madrid (Spain) and a MA in Political Communication by IUIOG-Madrid (Spain). His research lines focus on online disinformation and fact-checking journalism. https://orcid.org/0000-0002-4830-5554; Twitter: @CarlosRguezPrez

Francisco J. Paniagua-Rojano is Senior Lecturer (Tenured) at the Department of Journalism, at the University of Malaga, where he teaches Corporate Communication and Strategies. His main research lines focus on credibility of media, social media and journalism. He has published more than 40 papers about Journalism, credibility and social media. He has worked as Chief of Communication in the International Andalusian University (2011–2014) and Andalusian Federation of Municipalities and Provinces (2000–2003) Since 2008, he is the secretary of the section about "Corporate Communication, Advertising and Public Relations," in the Spanish Association of Communication Researchers.

Raúl Magallón-Rosa is Senior Lecturer in Communication Theory at the Department of Communication, Universidad Carlos III de Madrid (Spain). He has an European Doctorate in Communication at Complutense University of Madrid. He is the author of *Unfaking News: Cómo Combatir la Desinformación* [Unfaking news: How to tackle disinformation] (Pirámide, 2019) and other published articles in peer-reviewed journals on the topic of fake news and fact-checking.

Article

Political Memes and Fake News Discourses on Instagram

Ahmed Al-Rawi

School of Communication, Simon Fraser University, Burnaby, V5A 1S6, Canada; E-Mail: aalrawi@sfu.ca

Abstract
Political memes have been previously studied in different contexts, but this study fills a gap in literature by employing a mixed method to provide insight into the discourses of fake news on Instagram. The author collected more than 550,000 Instagram posts sent by over 198,000 unique users from 24 February 2012 to 21 December 2018, using the hashtag #fakenews as a search term. The study uses topic modelling to identify the most recurrent topics that are dominant on the platform, while the most active users are identified to understand the nature of the online communities that discuss fake news. In addition, the study offers an analysis of visual metadata that accompanies Instagram images. The findings indicate that Instagram has become a weaponized toxic platform, and the largest community of active users are supporters of the US President Donald Trump and the Republican Party, mostly trolling liberal mainstream media especially CNN, while often aligning themselves with the far-right. On the other hand, a much smaller online community attempts to troll Trump and the Republicans. Theoretically, the study relies on political memes literature and argues that Instagram has become weaponized through an ongoing 'Meme War,' for many members in the two main online communities troll and attack each other to exert power on the platform.

Keywords
fake news; Instagram; polarization; political memes; social media

1. Introduction

This study started as an exploratory one to investigate the hashtag #fakenews on Instagram, but the researcher was surprised to find the extent of political polarization, racism, and hate on this platform around this issue which prompted him to pursue this project in more detail since it was possible to collect data referencing to a hashtag on Instagram. The research empirically examines fake news discourses on Instagram by focusing on memes and offers a unique analysis of visual metadata, a feature that has not been examined before in relation to the quantitative analysis of visual discourses on Instagram. From a theoretical point of view, the article attempts to expand on the political memes literature by arguing that Instagram has become a weaponized platform despite its reputation in popular culture as a cool space for young people to post their selfies, food, and travel pictures.

Here, political memes are designed and disseminated to troll opponents in different online communities in what appears to be a Meme War founded on ideological differences and beliefs (Al-Rawi, 2020). The continuation and enhancement of this Meme War is facilitated by the platform's positive apolitical reputation and structure that is built on sharing funny memes, yet the reality is different and far more toxic than is thought, as will be shown here.

Instagram has been largely understudied due to the difficulty of obtaining large datasets from this platform (Highfield & Leaver, 2015). Unlike the case of other social media outlets like Twitter and Facebook, the other unique aspect of this research study is that it investigates fake news discourses on Instagram as previous studies analyzed fake news on other social media platforms and mainstream media, as well as the use of bots (Al-Rawi, 2019a; Al-Rawi, Groshek, & Zhang, 2019; Lazer et al., 2018; Vosoughi, Roy, & Aral, 2018).

2. Literature Review

This article focuses on political memes and how they are used in relation to fake news discourses on social media partly following the recommendation made by Leskovec, Backstrom, and Kleinberg (2009) on the importance of studying political memes and their diffusion. In its basic form, a meme is part of today's participatory and mainstream culture (Jenkins, 2006) and is generally defined as an "image, video, piece of text, etc., typically humorous in nature, that is copied and spread rapidly by internet users, often with slight variations" (Meme, 2020). A more detailed definition is presented here:

> *Internet memes* are digital texts—verbal, visual, or audial—that share common attributes and undergo variations by multiple users. They are created with awareness of one another, and they require prior acquaintance for proper production and consumption, as the repetitive (or memetic) element of the text often incorporates coded cultural information essential for proper interpretation….This practice facilitates, on the one hand, the expression of originality and creativity, and, on the other, a sense of belonging to a cultural collective. In this sense, the memetic practice meets both the demand for individualism and the yearning to belong characteristic of participatory culture. (Gal, 2018, pp. 529–530)

The term was first introduced by Richard Dawkins, a biologist, who mentioned that memes are transmitted units of culture which are similar to genes because they are disseminated either by copying or imitation (Dawkins, 1976). Indeed, memes have become very popular today because of social media. Similar to viral content, many memes are widely shared on Social Networking Sites (SNS), and they often replicate and reappear in different formats, shapes, and contexts (Tay, 2014, p. 48). Hence, memes are understood to be "socially constructed public discourses in which different memetic variants represent diverse voices and perspectives" (Shifman, 2014, p. 7).

One of the important features of memes is their importance in bringing digital communities together (Nissenbaum & Shifman, 2017, p. 485). Here, Limor Shifman stresses that memes operate on microlevel due to their associated and engaged small online communities, but they are also vital at the macrolevel because they often shape public discourses and enhance collective identities (Shifman, 2012). Indeed, they "play an integral part in some of the defining events of the twenty-first century" (Shifman, 2014, p. 6). In other words, many digital communities share and disseminate relevant memes that reflect their ideologies through which they practice gatekeeping activities (Burgess, 2008; Miltner, 2014). In this regard, memes are understood to be cultural productions produced and/or disseminated by activists, sometimes called "warriors," in their efforts to oppose and possibly change the status quo (Lasn, 2012,

p. 147); even the dissemination of these memes is considered "a product of societal and communal coordination" (Nissenbaum & Shifman, 2017, p. 485). Shifman (2014) mentions here that memes have three functions: as forms of persuasion or political advocacy, as grassroots action, and as expression and public discussion (p. 122). Due to these functions, memes carry power which is why many online users are "fighting for the hegemony of memes" in what is known as "memes warfare" (Häkkinen & Leppänen, 2014, p. 7, 19). In this regard, Nissenbaum and Shifman (2017) emphasize that memes have three formulations as capital: subcultural knowledge, unstable equilibriums, and discursive weapons, and even the word discursive indicates that there is some kind of "repetition" or "thematic matter from within an established meme" (Wiggins & Bowers, 2015, p. 7). In addition, Seiffert-Brockmann, Diehl, and Dobusch (2018) indicate that memes contain three types of communication logic: wasteful play online, social media political expression, and cultural evolution. Finally, Shifman (2014) views memes as playing a role in the "new landscape of Web-based political participation, both in grassroots and top-down campaigns" (p. 122). In this article, we focus on the discursive weapons and features of political expression that are directly linked to fake news discourses and main functions of memes.

Due to the affordances of social media, political memes often have global outreach as they are used in different political contexts and regions in relation to resistance, activism, and "democratic subversion" (Shifman, 2014, p. 123). They often have many implied meanings that carry ideological undertones. In China, for instance, memes are used to express political views about social norms (Yang, 2014), while memes in Russia are often employed as a networked political action and protest against Vladimir Putin (Lonkila, 2017). In some cases, photoshopped images presented in the form of memes are used to protest perceived injustice regarding police acts (Bayerl & Stoynov, 2016). In the US, some ironic memes were used to make comparisons between Barack Obama and Martin Luther King such as the case 'I Have a Drone' meme challenge (Howley, 2016). In other words, memes can express "reactions to and protests against political events and figureheads" (Häkkinen & Leppänen, 2014, p. 7). In a Twitter study on the 2014 Brazilian presidential election, a total of 599 memes were analyzed following Shifman's conceptualization, categorizing them as follows: persuasive, grassroots action, and public discussion (Chagas, Freire, Rios, & Magalhães, 2019). Another study examined the use of memes on Twitter during the 2015 State of the Nation Debate in Spain, and the content analysis shows significant differences in the use of memes by a variety of Spanish political parties (Martínez-Rolán & Piñeiro-Otero, 2016). Similarly, during the 2008 US election, video memes like 'Obama Girl,' 'Wassup,' and 'Yes We Can' became very popular attracting the attention of millions of users (Shifman, 2014, p. 124), while other memes were used to troll

Barack Obama during the 2012 presidential campaign, presenting him as unpatriotic such as using the 'crotch salute,' 'left-hand salute,' and 'Veterans Day non-salute' memes (Burroughs, 2013). Finally, and as will be mentioned below, memes were aggressively used during the 2016 US election often targeting Hillary Clinton and Donald Trump (Denisova, 2019). In brief, memes are viewed as creative and often funny political productions that are associated with activism, social movements, political protests, and cultural resistance.

Yet, memes have also been lately weaponized, hijacked, and exploited by hate groups, state-run trolls, and extremists to support their political agenda. Numerous studies examined this aspect of memes' use mostly by alt-right trolls and groups (Bay, 2018; Bogerts & Fielitz, 2018; Dematagoda, 2017; Flisfeder, 2018; Hannan, 2018; Harmer & Lumsden, 2019; Heikkilä, 2017; Woolley & Guilbeault, 2017). For instance, Russian trolls actively used memes to target young people, mostly showing Western leaders as fascists, while Ukrainian soldiers were presented as Nazis with fake photoshopped images (Aro, 2016, p. 125). These types of memes are regarded as toxic because they carry different types of hate speech discourses (Coker, 2008, p. 911). However, empirical research on Instagram as well as on fake news discourses is not available despite the fact that a US Senate Intelligence Committee report indicated that Russian trolls actively used Instagram more than Facebook during the 2016 election (Frier & Dennis, 2018). Hence, this study fills another gap in literature on the issue of toxic speech on Instagram.

In general, many of the US President Donald Trump's followers on Reddit were active in trolling and distributing 'dank memes' during the 2016 US election. These types of memes are characterized as low-quality images or videos whose main purpose is the dissemination of jokes. Engaged in the so-called 'Great Meme War,' those supporters call themselves 'keyboard commandos' whose goal is to "harass Trump's detractors and flood the Internet with pro-Trump, anti-Hillary Clinton propaganda" (Schreckinger, 2017). This was part of Meme War I that was meant to get Donald Trump elected (Roose, 2017) with the help of 'meme armies' that lately announced that Meme War II has already started (Roose, 2017) and is meant to keep Trump in power. In a recent study on Trump's followers on the subreddit /r/The Donald (T D), the authors highlighted how fake news stories are widely circulated such as the one calling for action on the murder of the Democratic National Committee (DNC) employee Seth Rich, for his fatal incident was falsely linked to a conspiracy theory revolving around his connection to the DNC email leak (Flores-Saviaga, Keegan, & Savage, 2018). Aside from Reddit, other social media platforms like 4chan received some scholarly attention regarding the study of Trump's trolls and their memes such as 'Trump Train' and 'You can't stump the Trump' (Merrin, 2019). Other popular memes include the far-right 'Pepe the Frog' and the

'Deplorables' ones which present Trump himself as Pepe the Frog figure (Hine et al., 2017; Merrin, 2019, p. 208). As a matter of fact, Trump promoted the meme of Pepe the Frog on his Twitter account in 2016 (Revesz, 2016), which assisted in making some of the alt-right symbols become mainstream (Davey, Saltman, & Birdwell, 2018). In brief, the 'Great Meme Wars' are waging on different social media outlets in which political trolls from different ideological backgrounds try to exert their cultural power and control over the online discourses. Since there is a dearth of empirical research of political memes on Instagram, this study fills a few gaps in the academic literature. This study attempts to answer the following research questions:

RQ1: What is the nature of the main online communities that discuss fake news with Instagram's political memes?

RQ2: What are the dominant topics of fake news discourses in these memes?

3. Method

The data collection in this study was done in two stages. The first stage involved collecting 293,773 Instagram posts for the period between 30 September 2012 to 13 November 2018 using Netlytic academic subscription. However, when Instagram changed its Application Programming Interface rules in late 2018, the data collection stopped, so the researcher adapted a Python script that collected JSON file for each Instagram post using the search term #fakenews. The search resulted in collecting 551,402 Instagram messages posted between 24 February 2012 to 21 December 2018 by 198,684 unique users (Figure 1). The Python script collected older data dating back to 2012 which referenced #fakenews, while any deleted posts were not retrieved from the platform. In this regard, the highest number of posts (n = 10,294) were posted on October 21, 2018, which coincided with a tweet from President Donald Trump on the same day, stating: "Facebook has just stated that they are setting up a system to 'purge' themselves of Fake News. Does that mean CNN will finally be put out of business?" (Trump, 2018). The Python script has limitations, for it only collected Instagram posts that are available online, constituting about 50% of the total number of posts referencing #fakenews on the platform.

After the data collection, a number of mixed methodological measures were followed to analyze the Instagram posts. First, topic modelling was used with the use of a commercial software called QDA Miner—WordStat 8. Topic modeling is a machine learning language technique based on analyzing unstructured data. For our study, we used Non-negative matrix factorization (NNMF) approach which is a text mining method (Pauca, Shahnaz, Berry, & Plemmons, 2004), often employed in investigating a variety of issues like topics identi-
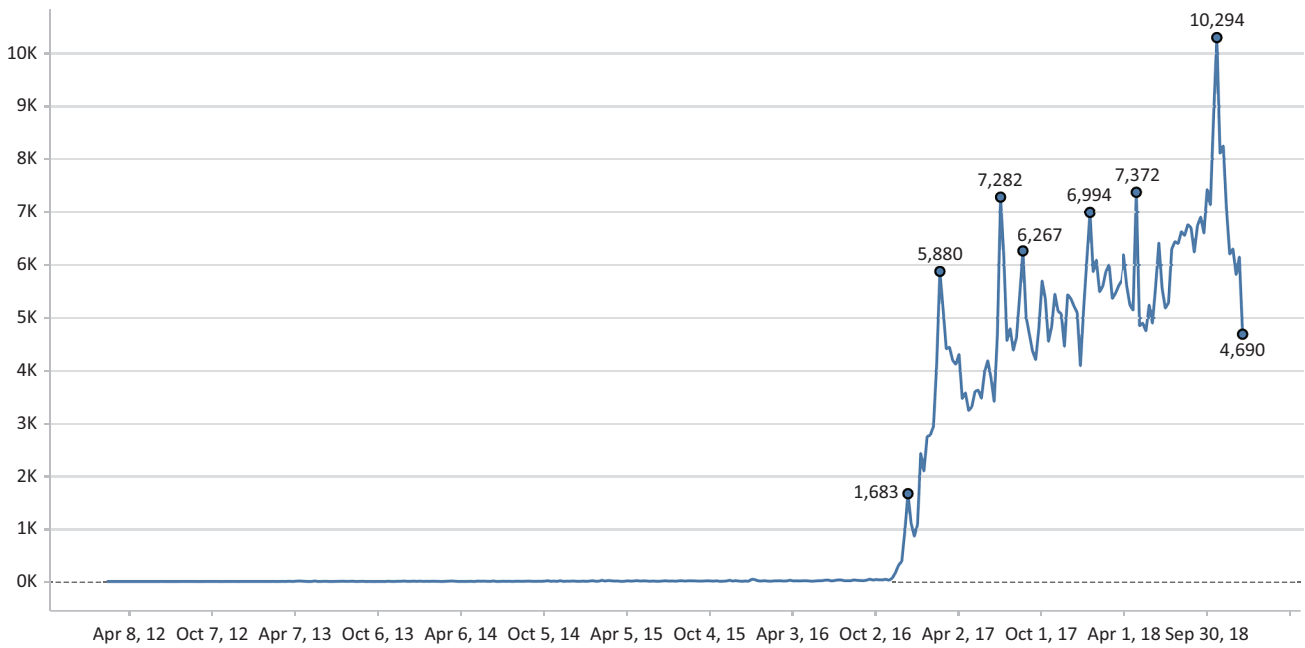
Figure 1. Timeline of Instagram posts distribution referencing #fakenews.

fication and monitoring (Vaca, Mantrach, Jaimes, & Saerens, 2014). NNMF topics were ranked based on their coherence which "refers to the overall quality and the semantic relatedness of the terms appearing in a topic descriptor'' (Belford, Mac Namee, & Greene, 2018, pp. 8–9). The same computer software was used in identifying the most recurrent hashtags, user mentions, most recurrent words and phrases as well as their co-occurrences measured through proximity plots.

Second, we used pivot tables to identify the unique users though the process was tedious because the Python script only provides individual user ID numbers in JSON format, so the author had to convert the files in bulk into Excel and individually search for the top Instagram usernames on the platform. Using inductive coding, a qualitative examination of 20 users was initially conducted by two coders to examine any emerging categories (Wimmer & Dominick, 2013), and three main online communities were identified, mostly centered around personalized discussion of famous individuals: (1) Pro-Trump, (2) Anti-Trump, (3) Bolsonaro-related, (4) Other. Intercoder reliability test was conducted using Krippendorf Alpha test ($\alpha = 1.0.$). From this representative sample, five accounts were initially suspended or deleted including three pro-Trump and two without clear affiliation. The latter category included users who have no clear political affiliation and often refer to #dankmemes, while a few others focused on general news dissemination, conspiracy theories, marketing, and creative art and photography. Some of the irrelevant conspiracy theories promoted include the illuminati, veganism, UFOs, and the notion that earth is flat. These users were excluded from the study due to the diversity of their posts and their small numbers in the dataset. The third online community included five users dealing with the

populist Brazilian President, Jair Bolsonaro, who posted in Portuguese, repeatedly adding the hashtag #fakenews to their posts. This is another minor online community that is excluded from the study because the author is not familiar with Portuguese language. Afterwards, the top 100 most active users, who collectively sent 100,507 Instagram posts, were thoroughly examined by studying their online profiles, other Instagram posts if available, and the archived Instagram posts in case their accounts were deleted.

Ultimately, the study adds a unique aspect that has not been included before in fake news research on social media. In this regard, the computational analysis of a large number of images in media and communication discipline is rare, and this study attempts to fill a gap in literature especially in relation to Instagram research. We followed here the recommendation made by Highfield and Leaver (2016) who stress that "developing approaches to track and study the visual as widespread social media form, including across platforms as content is shared and reappropriated, is a necessary undertaking for a critical understanding of social media use" (p. 58). Out of the total sample of collected Instagram posts, there were 310,649 visuals (56.3%) that were automatedly tagged by Instagram. The platform sometimes provides an automated description of images that includes thousands of categories like '1 person,' 'tree, snow, sky, table and outdoor,' or '3 people, meme, crowd and text.' These visual tags can appear if one hovers with the mouse over some but not all Instagram images. In our study, we found over 5,900 different tag categories. We textually analyzed these visual tags as well as 4,596 textual content (0.8%) in these visuals (memes that contain text) in order to add more depth into our textual analysis. In this study, we found that fake news discourses on Instagram

are more likely to include images 78.9% (n = 435,387), videos 14.3% (n = 79,004), and less frequently sidecar images 6.7% (n = 37,011), which refers to a series of photos posted by one user and dealing with a similar theme or event.

Finally, research ethics clearance was not required from the author's university because the social media data is publicly available. More importantly, informed consent from active social media users is not usually needed when conducting critical media research like the case study examined here that involves alt-right users (see Townsend & Wallace, n.d., pp. 11–12).

4. Results and Discussion

This study focuses on the examination of political memes in relation to fake news discourses on Instagram. Its goal is to understand the nature of the active users and their online communities as well as the main topics discussed on Instagram since there is a gap in literature on this research aspect. To answer the first research question on the nature of the online communities on Instagram, we found that the most dominant online community among the top 100 most active users is the Pro-Trump one (68%) followed by Anti-Trump (10%) online community. A third minor online community is formed around the Brazilian populist President, Jair Bolsonaro (5%). The Pro-Trump camp posted 69,630 messages constituting 69.2% of the total posts of the active users, while the Anti-Trump online community sent 14,061 ones which amounts to 13.9% only. Indeed, the Pro- and Anti-Trump users form the major Instagram communities which discuss fake news, and what makes them distinctive and unique is the personalized praise as well as attacks or trolling activity against individual politicians. This has been evident not only in the qualitative investigation of the most active users presented above and in the following paragraph but also the examination of the top posts, most recurrent words and phrases, and more importantly the large data analysis of the accompanying images.

Regarding the visual tags, Figure 2 provides a representation of the top 50 most popular image descriptions, and we found that the top one is related to images containing '1 person' (n = 71,876) followed by images with 'text' (n = 36,508), '1 person, text' (n = 30,650), '2 people' (n = 24,134), and '2 people, text' (10,201). As can be seen, there is a personalized type of messages in these visual discourses that mostly deal with one person, which explains the prevalence of certain individuals in the textual discourses about fake news on Instagram. In fact, the total number of images containing one person in different formats is 136,716, constituting 44% of the total tagged images. This is the highest percentage among all the other visual tags.

As for the visual analysis, we agree with Highfield and Leaver's (2016) assertion that the "visual is central to everyday life and social media practices" (p. 49). In this study, we found that Donald Trump is the main focus of the visual discussion, for he is part of the most frequent phrase (n = 338) if one takes into account the combination of the two phrases 'Trump Realdonaldtrump' (n = 173) and 'Donald Trump' (n = 165). The same emphasis is found in the most frequent words in the image texts, for 'Trump' comes first (n = 1229) if one ignores the basic tag words like 'text,' 'person,' and 'people,' and without taking into account the word 'Realdonaldtrump' (n = 220). Other popular politicians include 'Obama' (n = 165) and Clinton (n = 129), while the recurrent phrases include 'Hillary Clinton' (n = 66) and 'Ben Shapiro' (n = 53), the famous conservative commentator. Once again, this finding confirms the discussion above on the personalized type of discussion on fake news which is evident in the visual findings as well as the textual analysis presented below.

To further understand these two main online communities, we provide here more in-depth analysis of some of the top 20 most active users (Table 1). These users consist of a majority of pro-Trump Instagrammers (n = 17) and only a couple of anti-Trump ones as well as one with no clear affiliation. This table clearly indicates the dominance of the pro-Trump community who regard themselves as meme warriors, as mentioned above. Such users often use a long list of hashtags in order to spread their memes as far as possible which includes several themes like praise for their community leaders and members (#Rarepepe, #Kekistan, #Praisekek, and #kek), attacks against democrats and liberals (#liberalsSuck, #liberaltears, #draintheswamp, and #libtards), support for Donald Trump (#MAGA, #IsMyPresident, #buildthewall, #trumptrain, and #Trump2020; see Figure 3), and criticism against liberal mainstream media especially CNN (#CNNisFakenews and #fakenews). These results are actually similar to a similar investigation on Twitter that examined the hashtag #fakenews (Al-Rawi, 2019a). Regarding the last theme, the majority of users in this community express doubts and mistrust in mainstream media such as the case of the user @stoppingtheabuse who describes himself as follows: "I am a patriot. I am a proud American. I am white. I am Christian. I am a republican....I don't listen to fake news."

To examine a few top users from this community, the 7th most active Instagrammer calls himself "@Captain_Kekistan" which is a reference to the Hollywood superhero, Captain America, and it is also evident in the user's profile that refers to the same character though (s)he changed the color into green (Figure 4). As for 'Kekistan,' it is a popular far-right term which is based on a coded language used among this community who believe in conspiracy theories like QAnon (Al-Rawi, 2020; Merrin, 2019). The latter is a conspiracy theory term that implies the existence of a deep state in the US whose goal is allegedly overthrowing or undermining Donald Trump's presidency (Al-Rawi, 2020). Another example of a Pro-Trump user, conservative__americans posts similar Instagram messages with those periodically sent by Captain_Kekistan. The user's full name
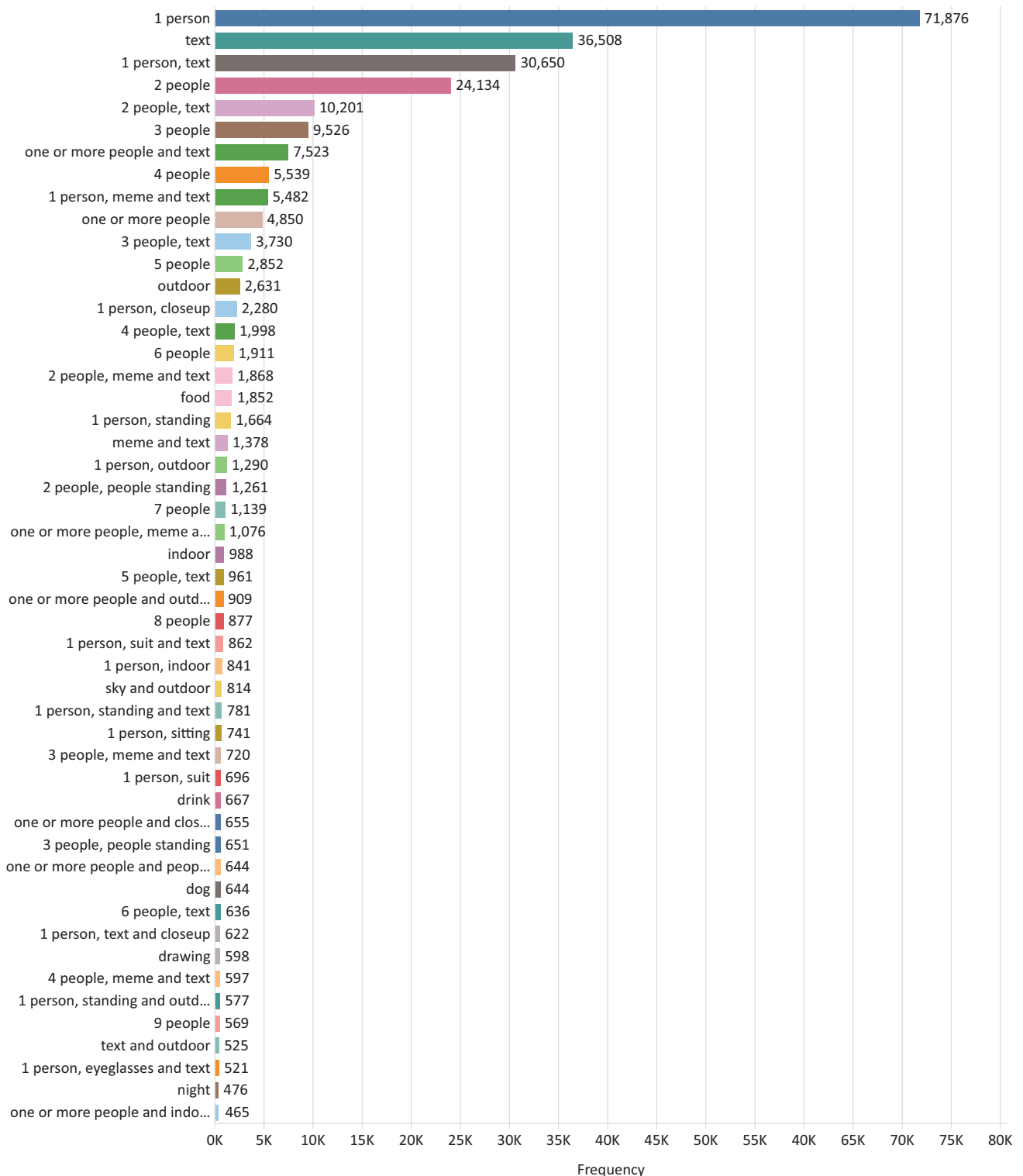
Figure 2. The top 50 most associated Instagram's visual tags.

is 'Right Wing Meme Factory,' denoting the purpose of this account. Interestingly, the user's online account shows only four messages though (s)he sent 1399 posts archived in our dataset which means that (s)he periodically deletes messages probably to avoid detection or removal by the platform. This is also a social media strategy used by drug dealers (Al-Rawi, 2019b) and state-run users (Al-Rawi & Shukla, 2020) to spam others while making sure their online presence remains active on social media platforms. I call these users 'sleeper cells' because they get activated when needed. After performing their required duties, they disguise themselves again by removing all their public messages. Another way of avoiding detection is by keeping the account private such

Table 1. The top 20 most active users on Instagram in relation to #fakenews.

No.	Username	Frequency
1.	alternative_news_media	10,426
2.	breakupartist	5,620
3.	_conservative_warrior_	4,419
4.	maxcua	2,610
5.	unfiltered.politics	2,550
6.	onyxnegus	2,314
7.	captain_kekistan	1,896
8.	only_conservative_	1,755
9.	deplorable_brandysthebomb	1,726
10.	usa_stands_in_honor__	1,612
11.	lepetersworld	1,566
12.	thetrumpphenomenon	1,561
13.	thedukesofdonaldtrump	1,504
14.	this.mom.is.right	1,485
15.	bigleaguetrumpster	1,407
16.	dmone_fat	1,404
17.	conservative__americans	1,399
18.	doomedamerica	1,325
19.	vertical_stance	1,283
20.	universal_enlightenment	1,142

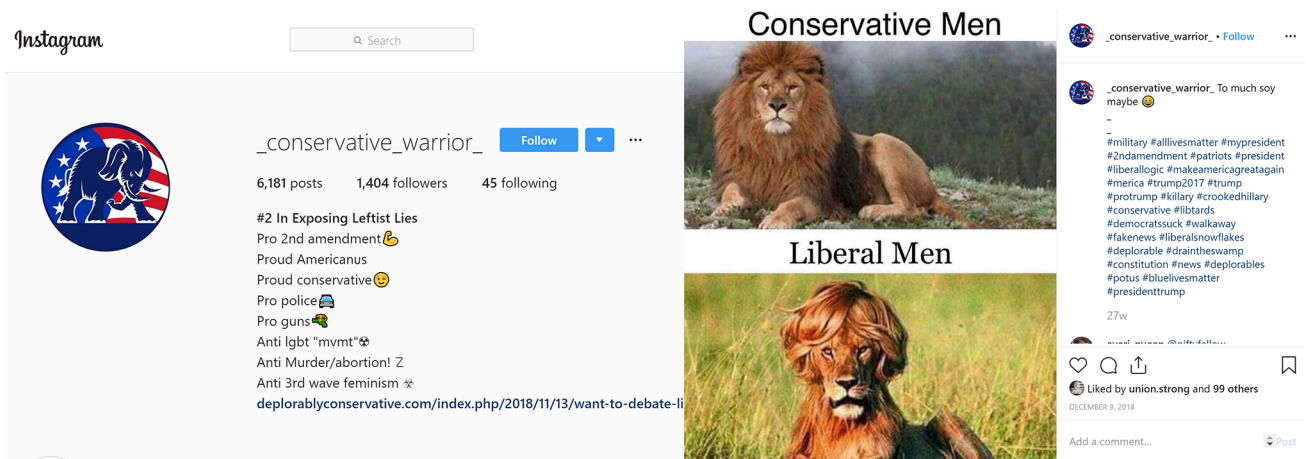**Figure 3.** A screenshot of a Pro-Trump follower using the hashtag #fakenews (https://www.instagram.com/_conservative_warrior_/?hl=en).
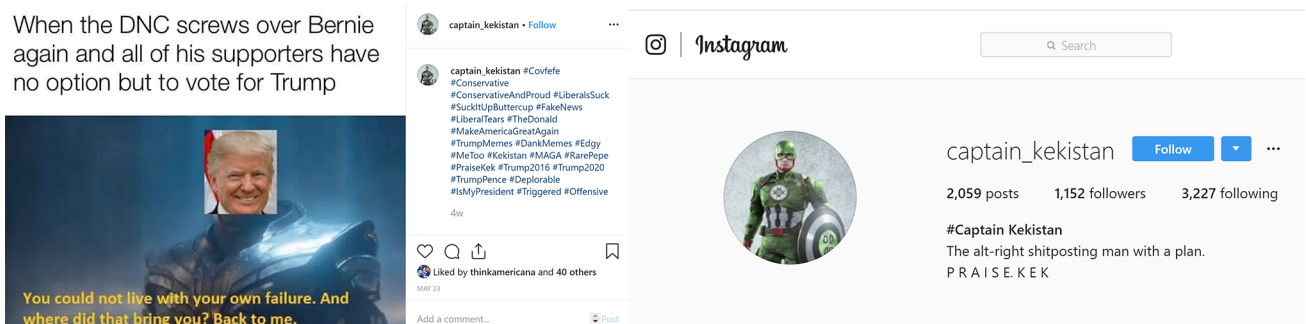

Figure 4. Screenshots of Captain_Kekistan on Instagram (https://www.instagram.com/captain_kekistan/?hl=en).

as the case of @conservative.comedy that posted 705 Instagram messages as well as @but_muh_clownshow that is discussed below. Some of the other general features of this online community are its positions on several issues such as anti-immigration, distrust of Muslims, nationalism, protectionism, and xenophobia. For instance, @pjb101263 posted 597 Instagram messages referencing #fakenews including the following one that shows the US Congresswoman, Ilhan Omar, flying an airplane towards what appears to be the twin towers (Figure 5). Another feature that some users in this online community employ is anti-Semitic discourses which are manifested in several ways including the use of far-right hashtags (see above), emojis like 🐸 that was used 794 times, and the triple parentheses or echoes (Williams, 2016). In our dataset, we found the echoes used 213 times in the public discourses on fake news. For example, @trumpmemz user mentions that Elon Musk is triggering Jewish mainstream journalists for his alleged plan to create a ranking website for journalists' credibility.

While the study is being written, two more Pro-Trump accounts were deleted by Instagram including @vertical_stance and @excuse_me_excuse_me_. It is not clear why the accounts were removed, though. Similar accounts shortly emerged like vertical_stance2.0 whose original motto is 'Exposing the mainstream media' and 'Journalism is dead.' As for @excuse_me_excuse_me_, the user acts like a troll who has been repeatedly blocked and deleted, stating: "My dear conservative friends 😂😎 I have made a second account, just in case. 😉 This will remain my main page, but feel free to follow the other one @excuse_me_excuse_me_." The user coordinates with similar trolls in attacking liberal users. One post, for instance, praises the collective efforts directed at a par-

ody account of Hillary Clinton, mentioning: "Guys, you really hit that profile! @hillarry_clintton 😂😂😂 Last post can't be commented any more, so….Good job! 😎👍👍'' or "Tag a liberal 😁😁.'' As a member of an active online community, this user repeatedly urges his friends and followers to target democratic leaders like the real account of Hillary Clinton: "😂😂😂 @hillaryclinton ——————————————'' as well as follow other active users or influential in the community, stating: "Follow my partners: 🔨@thetrump phenomena🔨/🔨 @usa_everyday 🔨/🔨 @authentic donaldjtrump🔨.'' Similarly, numerous other users like @stoppingtheabuse and @conservative.comedy encourage their friends to re-post favorable messages and follow other members in the community by listing their Instagram usernames. This technique is similar to the Swarmcast model in communication wherein affiliated users gather around their opponents to attack them (Al-Rawi, 2018, p. 743). In brief, there is a clear coordination of organized efforts to identify and troll liberals on Instagram in order to silence or limit their activities. It is important to mention here that it is not possible to know whether these coordinated trolling activities are centralized or not as they can be organically driven by conservative fans and followers.

On the other hand, the anti-Trump online community on Instagram, which is much smaller and less organized than the Pro-Trump one, is often involved in trolling Republicans and the far-right group. Some of the popular hashtags used by this anti-Trump community include #Blacklivesmatter, #Resist, #Dumptrump, #Fucktrump, #Notmypresident, #Makeamericablueagain, #Obama care, #LGBTQ, and #Progressive. To take a couple of examples, the second top user @Breakupartist (n = 5,620) almost always trolls Donald Trump and

Figure 5. A photoshopped image showing Congresswoman Ilhan Omar as a 9/11 attacker.

his administration, sometimes using offensive language similar to the discourse used by the Pro-Trump community. In Figure 6, we find that Trump was photoshopped to make him look like Hitler, while the US President logo sign was appropriated to read 'Moron of the United States' and MOTUS instead of POTUS. In addition, the image designer used the triple parentheses or echoes as a gesture of resistance against anti-Semitic trolls and in solidarity towards Jews (Williams, 2016). Another user, @ledsjam_trump, sent 450 messages on fake news, describing himself as follows in his profile: "No Drumpf....Non-Stop! I'm the 400 lb troll in my parents' basement, that you've heard about. 🔥Dedicated to Mocking Our Nation's Highest Ranking Buffoon. Homemade. RESIST!" (Figure 6).

To sum up, there are two distinct online communities that are active in relation to the discourses about fake news on Instagram. Each community trolls each other, but the pro-Trump one is more active and visible which is apparent in their organized efforts, total frequency of their posts, and prevalence in the top 100 most active users. The active members of these two communities regard themselves as activists who have a duty to counter the hegemony and attack the credibility of their opponents, and the ultimate goal is to control the discourses, gain more followers and sympathizers, and possibly influence the minds of Instagram users through the use of appealing, witty, and funny political memes. The results of this study align with previous research on the role of memes in enhancing community ties and creating a collective identity (Shifman, 2012) especially that sharing certain memes is a reflection of the ideological affiliation of some online communities and their gatekeeping practices (Burgess, 2008; Miltner, 2014). In addition, the political polarization that separates the two main communities is translated into tense and often toxic Instagram content that is often described as an ongoing "memes warfare" (Al-Rawi, 2020; Häkkinen & Leppänen, 2014). Again, this is in line with previous

research that examined social media and political polarization (Bay, 2018; Heikkilä, 2017) such as some polarized social media content posted during the 2016 US election (Denisova, 2019).

To answer the second research question on the main topics on Instagram in relation to fake news discourses, we find that the top five topics and their associated descriptors include four that are related to the Pro-Trump online community including Alllivesmatter, Rightwing, Draintheswamp, Republican, and only one relevant to the anti-Trump camp which is Notmypresident (Table 2). To provide more context, we will explore only a few ones here due to the limited space. The first topic is Alllivesmatter which is a term used by some conservatives to counter the claims of the Black Lives Matter movement. The use of this term suggests that all Americans including whites and police forces should be given equal weight and importance when it comes to issues of social justice and equality, ignoring the systemic injustice and historical circumstances that created economic and political inequality in the first place. Some of the associated words with this topic include a clear attack against Democrats with the word 'Liberallogic' to demean their arguments as well as other supportive terms like MAGA and Makeamericagreatagain. Criticism against liberals is also manifested in the third topic which is entitled 'Drain the Swamp,' a political term modified by President Trump to attack his Democratic opponents and accuse them of corruption and ineptness (Bierman, 2018). In relation to this topic, we find that some of the associated terms include Antifa, the anti-Fascist movement that is often the object of attacks by far-right groups, as well as CNN and its constant association with fake news. In fact, Trump encouraged his followers to troll CNN by sending funny or critical memes and videos (Gallagher, 2019). On the other hand, there are also words that show solidarity and unity among the Pro-Trump community members like MAGA, Mypresident, Deplorable, that was appropriated from

Figure 6. Anti-Trump Instagram users trolling Trump using #fakenews (https://www.instagram.com/ledsjam_trump/?hl=en).

Table 2. The major topics on Instagram in relation to #fakenews.

Topic	Keywords	Coherence
Alllivesmatter	Alllivesmatter; Constitution; Media; Democrats; Media; Liberallogic; Meme; La; Bluelivesmatter; Partners; Makeamericagreatagain; Para; Patriot; Donaldtrump; Memes; MAGA; Conservative; Republicans	0.514
Rightwing	Rightwing; Para; LOL; Time; Maga; Donaldtrump; Patriot; Deplorable; Republicans; Americafirst; America; President; Buildthewall; Conservative; Liberallogic; Fake; Makeamericagreatagain; CNN; La; Bluelivesmatter; Patriots; God; Republican; Politics; Follow; Media; Democrats; Meme; Partners	0.497
Draintheswamp	Draintheswamp; Merica; Qanon; Fakenews; Realdonaldtrump; Antifa; Follow; CNN; Deplorable; MAGA; De; Fake; Mypresident; America	0.429
Republican	Real; Republican; Resist; Buildthewall; Bluelivesmatter; Americafirst; American; Media; La; Makeamericagreatagain; Donaldtrump; Rightwing; Patriot; Maga; President; Conservative; Deplorable; LOL; Media; Republicans	0.409
Notmypresident	Government; Notmypresident; Foxnews; De; America; Media; GOP; President; Deplorable; God; Constitution; Meme; Rightwing; Americafirst; Real; Conservative; Resist; Time; Para; Make; Donaldtrump; LOL	0.341

Hillary Clinton's 2016 interview (Caffier, 2017), Merica instead of America which is often used in popular culture discourses to express patriotism, and Qanon.

To further investigate the main topics, we examined the most recurrent words to understand the salience of some terms that carry political or cultural meanings. Aside from the common words like 'fakenews' and 'fake,' Table 3 shows the top 50 most recurrent words and we can find that 'Trump' (n = 271,241) comes second due to his dominance in the discussion by the two main online communities. Yet, most of the discussion is supportive of the US President because the subsequent words that follow Trump are mostly related to his party and policies like 'MAGA'

Table 3. The most frequent words used in the Instagram posts on fake news.

No.	Word	Frequency	No.	Word	Frequency
1.	Fakenews	498,527	26.	People	35,601
2.	Trump	271,241	27.	Meme	34,143
3.	MAGA	128,505	28.	Funny	34,006
4.	Conservative	123,527	29.	Buildthewall	32,962
5.	News	117,173	30.	Truth	32,469
6.	Republican	89,493	31.	Liberallogic	32,037
7.	Donaldtrump	89,113	32.	Repost	31,366
8.	CNN	85,730	33.	Merica	29,365
9.	Makeamericagreatagain	84,974	34.	Obama	28,439
10.	America	81,886	35.	Trumpmemes	26,056
11.	USA	68,445	36.	Bluelivesmatter	26,003
12.	Fake	62,789	37.	Russia	25,456
13.	Follow	60,875	38.	Resist	25,374
14.	Politics	56,493	39.	Democrats	25,231
15.	Trumptrain	53,447	40.	Liberals	24,857
16.	Memes	51,352	41.	American	24,127
17.	Liberal	47,452	42.	Love	23,764
18.	Draintheswamp	46,622	43.	Ndamendment	23,613
19.	Presidenttrump	45,107	44.	Freedom	23,026
20.	Democrat	40,692	45.	Patriot	22,954
21.	Foxnews	40,152	46.	Libtards	22,735
22.	Potus	40,071	47.	Military	22,556
23.	President	39,895	48.	Notmypresident	21,774
24.	Media	38,803	49.	Partners	21,469
25.	Americafirst	38,330	50.	GOP	21,416

(Make America Great Again), 'conservative,' 'republican,' 'Donaldtrump,' 'Makeamericagreatagain,' 'Trumptrain,' 'Americafirst,' 'Buildthewall,' 'Blueivesmatter' etc. At the same time, there are many words used that indicate trolling activity or attacks against liberal democrats like 'Draintheswamp,' 'Liberallogic,' and 'Libtards.' On the other hand, the top 50 words only show two words that are often used by the anti-Trump online community including 'Resist' (n = 25,374) and 'Notmypresident' (n = 21,774), which indicates the minor presence of this online community.

In terms of media outlets, two channels are referenced in the list of top terms including CNN (n = 85,730) and Fox News (n = 40,152). To further understand how these outlets are used in the online discussion, we examined the most recurrent phrases or combination of words. In this regard, two of the most recurrent phrases are 'CNN fakenews' (n = 18382) and 'Fakenews CNN' (n = 14511) that are used by the Pro-Trump online community. We also examined the co-occurrence of the word 'CNN' and found that it is firstly connected with the term 'fakenews' (n = 72,765) with a very strong association (< 0.901)—1 is complete and 0 has no association. To a lesser extent, we find other phrases that indicate opposition to the Trump's administration including 'Fakenews Dumptrump Nevertrump' (n = 5924) and 'Fakenews Dumptrump Nevertrump Resist Fucktrump' (n = 5866).

To corroborate the above findings, we further examined the most mentioned users in the online discussion. Similar to the above findings, we found that President Donald Trump comes first as the most mentioned user (n = 10,743) immediately followed by CNN (n = 5,768). In the top 20 most referenced users, Trump's son, @Donaldjtrumpjr, is also among the most referenced as well as three other liberal mainstream

media outlets including the *New York Times*, *Washington Post*, and ABC, while Fox News comes near the end of the list. The only other politician included in the list is the Brazilian president, @Jairmessiasbolsonaro, and two of his sons Eduardo, @bolsonarosp, and Flavio, @flaviobolsonaro. In total and aside from five mainstream media outlets, there are nine Pro-Trump users, three Bolsonaro-related users, two deleted accounts, and one irrelevant one. Similar to the discussion found above, some active Pro-Trump users mention other ones in their community to get support such as @Teentrumpsupporter who lists 13 other Instagrammers as 'partners.' Other users with deleted accounts create similar profile names by sometimes adding a number, letter, or word to the original username. For example, after @But_Muh_Russia got deleted from Instagram, the user created @but_muh_clownshow, describing himself as follows: "(((They))) are scared of you. Nothing is more powerful than the public being awake and collectively free- thinking! Formerly @ but_muh_russia." In addition to the anti-Semitic use of the echoes, the image profile uses the far-right symbol of Pepe the Frog as well as a hidden lynching noose that is reminiscent of the slavery era though this time it looks like an implied threat against the LGBTQ community. This is evident in the head cap the Frog is wearing (Figure 7), for this clown pepe figure is often called Honk Honk or Honkler in popular culture and is intentionally used in an attempt to reclaim the rainbow symbol from the LGBTQ community (Jhaveri, 2019; Know Your Meme, 2018).

As for the visual analysis, we decided to combine more than one approach in the examination of visuals because it becomes a more effective method in understanding images (Rose, 2016, p. 99). As memes contain text embedded in the images, we decided to examine it. Aside from the word 'Trump' and a few other politicians
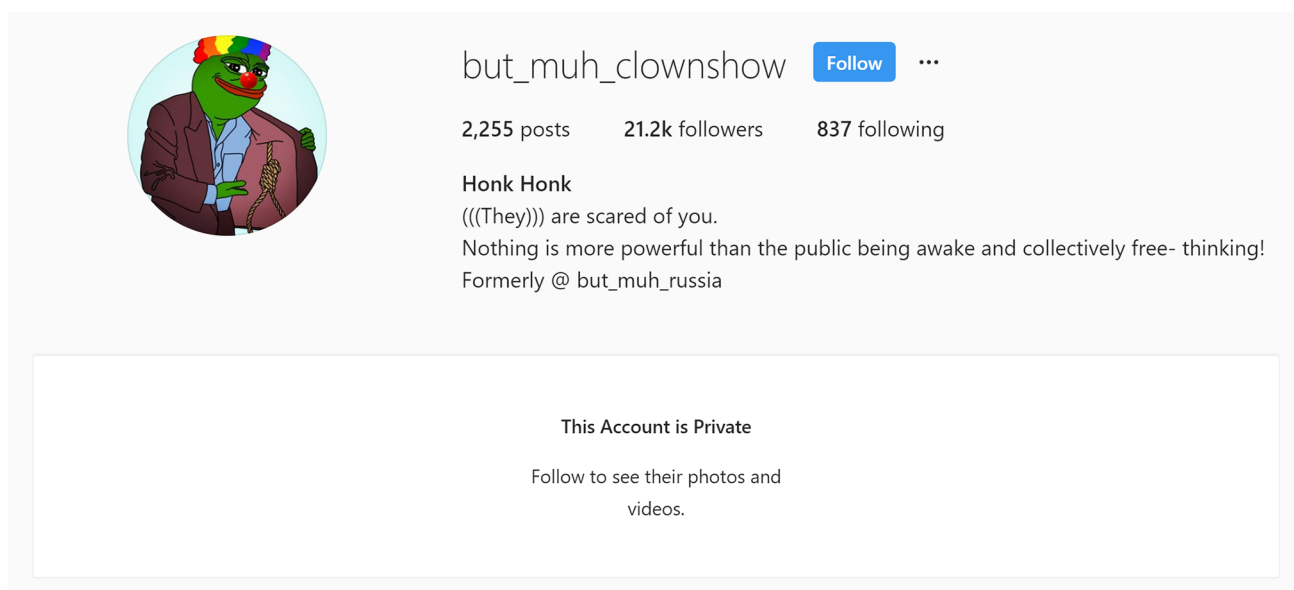

Figure 7. A Pro-Trump user on Instagram expressing implied threats.

mentioned above, we found that other frequent words and phrases in these memes include 'CNN' (n = 253) which is the top channel targeted followed by 'Fox News' (n = 105) and 'Media Is Fake News' (n = 34). This indicates the nature of visual attention to these two channels by the two online communities each accusing the other of being a fake news outlet. In this regard, we examined the co-occurrence of the words 'ABC,' 'CNN,' 'Fox,' and 'MSNBC' in the visual texts and found that 'CNN' is firstly and mostly linked to 'fakenews.' In fact, the terms 'fake' co-occurs 28 times with 'CNN' (< 0.015 association), 6 times with 'ABC,' 3 ones with 'NBC,' and 3 others with 'Fox.'

As for the other visual tags analysis, some of the identified objects can be relevant like 'stripes' which refer to the US flag, mostly denoting nationalistic expression through memes. This is, in fact, not unique to the US context, for using one's national flag is similar to the 'cyber-nationalism' sentiments expressed by some Chinese Instagram users (Fang, & Repnikova, 2018, p. 16). In our dataset, the US flag was used 790 times in these visuals in 9 different formats like '1 person, stripes' (n = 140), 'stripes and outdoor' (n = 66), 'standing and stripes' (n = 50), 'stripes and text' (n = 33). The qualitative assessment of a sample of these visual posts involving the US flag shows that the majority are Pro-Trump followers though there are different other users discussing it mostly to express nationalistic sentiment like the following user: "I'm a proud American. Are you?#MAGA#Trump #Bigdon#Thedon#Buildthewall#Makeamericagreatagain #Crookedhillary#Draintheswamp," while another one stated: "I love this country! Thank God we have a great President who fixing it! Look how beautiful our flag is! #MAGA" [sic].

Another aspect that requires analysis is the examination of the most recurrent posts that are used in the online discussion. Aside from the neutral, irrelevant, and ambiguous messages, we find that 15 out of the 20 top posts are part of the Pro-Trump online community that shares supportive hashtags on Donald Trump like #Trump2020 and #LoveTrump including some that are critical of mainstream media. For instance, the fifth most frequent post (n = 536) partly reads as follows: "#EnemyOfThePeople = @nytimes = #fakenews/fascists #EnemyOfThePeople = @cnn = #fakenews/fascists #EnemyOfThePeople = @washingtonpost = fakenews #fascists #JeffSessions = #EnemyOfThePeople @thejusticedept = #EnemyOfThePeople @abcnetwork = Nazis #WalkAway."

The same applies to the top 20 most commented on posts that include 9 out of 20 Pro-Trump messages which received 79,425 comments. On the other hand, there are only two anti-Trump posts that received 9,132 comments, while the other most commented on posts are neutral or did not show a clear political affiliation. Regarding the latter anti-Trump community, one of the top most commented on posts includes a call to follow other similar users in the community, stating: "Go fol-

low my partners: @_proud_liberal_, @texansocialdem, @the_norwegian_social_democrat…etc." This is, indeed, a tactic that is similar to what the Pro-Trump community follows.

To conclude, the visual and textual examination of fake news discourses on Instagram shows two highly polarized online communities separated by opposite political alliances. These two online communities exchange political memes that are mostly centered around individual prominent politicians especially the US President Donald Trump. However, the Pro-Trump community is much larger, more active, and organized than the anti-Trump community. At the same time, some active virtual 'warriors' of the Pro-Trump community show sympathy or affiliation with the far-right by using their hate and racist icons like Pepe the Frog, while expressing distrust in liberal mainstream media especially CNN. These findings indicate that Instagram has become weaponized by the two main online communities, and memes are used in an ongoing political online warfare to attack and demean the opponents. Meme War II is an ongoing daily reality on Instagram which requires more scholarly attention.

Finally, future research on fake news discourses need to include other under-researched visual platforms like Pinterest and Flickr in order to examine the nature of online communities and political memes that exist on these social media venues. In addition, in-depth interviews with members of the different Meme War communities need to be conducted to further understand their political strategies that are employed on Instagram and possibly elsewhere. Also, cross-national comparative research is vital, for this study identified an online community that is centered around the figure of the Brazilian President, Jair Bolsonaro, so it is important to understand how fake news discourses are used by the active members of this community.

Acknowledgments

Conflict of Interests

The author declares no conflict of interests.

References

Al-Rawi, A. (2018). Video games, terrorism, and ISIS's Jihad 3.0. *Terrorism and Political Violence*, *30*(4), 740–760.

Al-Rawi, A. (2019a). Gatekeeping fake news discourses on mainstream media versus social media. *Social Science Computer Review*, *37*(6), 687–704.

Al-Rawi, A. (2019b). The fentanyl crisis & the dark side of

social media. *Telematics and Informatics*, *45*, https://doi.org/10.1016/j.tele.2019.101280

Al-Rawi, A. (2020). Kekistanis and the meme war on social media. *The Journal of Intelligence, Conflict, and Warfare*, *3*(1), 13–13.

Al-Rawi, A., Groshek, J., & Zhang, L. (2019). What the fake? Assessing the extent of networked political spamming and bots in the propagation of #fakenews on Twitter. *Online Information Review*, *43*(1), 53–71.

Al-Rawi, A., & Shukla, V. (2020). Bots as active news promoters: A digital analysis of Covid-19 Tweets. *Information*, *11*(10), 461.

Aro, J. (2016). The cyberspace war: Propaganda and trolling as warfare tools. *European View*, *15*(1), 121–132.

Bay, M. (2018). Weaponizing the haters: The Last Jedi and the strategic politicization of pop culture through social media manipulation. *First Monday*, *23*(11).

Bayerl, P. S., & Stoynov, L. (2016). Revenge by photoshop: Memefying police acts in the public dialogue about injustice. *New Media & Society*, *18*(6), 1006–1026.

Belford, M., Mac Namee, B., & Greene, D. (2018). Stability of topic modeling via matrix factorization. *Expert Systems with Applications*, *91*, 159–169.

Bierman, N. (2018, February 9). Trump shifts meaning of 'Drain the Swamp' from ethics to anything he objects to. *Los Angeles Times*. Retrieved from https://www.latimes.com/politics/la-na-pol-swamp-20180209-story.html

Bogerts, L., & Fielitz, M. (2018). "Do you want meme war?" Understanding the visual memes of the German far right. In Maik Fielitz & Nick Thurston (Eds.), *Post-digital cultures of the far right: Online actions and offline consequences in Europe and the US* (pp. 137–153). Berlin: Majuskel Medienproduktion GmbH.

Burgess, J. (2008). 'All your chocolate rain are belong to us'? Viral video, YouTube and the dynamics of participatory culture. In G. Lovink & S. Niederer (Eds), *Video vortex reader: Responses to YouTube* (pp. 101–109). Amsterdam: Institute of Network Cultures.

Burroughs, B. (2013). FCJ-165 Obama trolling: Memes, salutes and an agonistic politics in the 2012 presidential election. *The Fibreculture Journal*, *22*. Retrieved from https://twentytwo.fibreculturejournal.org/wp-content/pdfs/FCJ-165Benjamin%20Burroughs.pdf

Caffier, J. (2017, February 9). Every insult the left uses to troll conservatives, explained. *Vice*. Retrieved from https://www.vice.com/en_ca/article/wn4vvq/every-insult-the-left-uses-to-troll-conservatives-explained

Chagas, V., Freire, F., Rios, D., & Magalhães, D. (2019). Political memes and the politics of memes: A methodological proposal for content analysis of online political memes. *First Monday*, *24*(2).

Coker, C. (2008). War, memes and memeplexes. *International Affairs*, *84*(5), 903–914.

Davey, J., Saltman, E. M., & Birdwell, J. (2018). The main-streaming of far-right extremism online and how to counter it. In L. Herman & J. Muldoon (Eds.). *Trumping the mainstream: The conquest of democratic politics by the populist radical right* (pp. 23–53). London: Routledge.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dematagoda, U. (2017). Revenge of the nerds: Recidivist masculinity, identity politics and the online 'culture wars.' *Journal of Extreme Anthropology*, *1*(3), 139–148.

Denisova, A. (2019). *Internet memes and society: Social, cultural, and political contexts*. London: Routledge.

Fang, K., & Repnikova, M. (2018). Demystifying "little pink": The creation and evolution of a gendered label for nationalistic activists in China. *New Media & Society*, *20*(6), 2162–2185.

Flisfeder, M. (2018). "Trump": What does the name signify? Or, Protofascism and the alt-right: Three contradictions of the present conjuncture. *Public Culture*, *14*(1), 1–19.

Flores-Saviaga, C. I., Keegan, B. C., & Savage, S. (2018). Mobilizing the trump train: Understanding collective action in a political trolling community. In K. Starbird & I. Weber (Eds.), Twelfth International AAAI Conference on Web and Social Media (pp. 82–91). Palo Alto, CA: AAAI Press.

Frier, S., & Dennis, S. (2018, December 17). Instagram was bigger Russian election tool than Facebook, senate report says. *Bloomberg*. Retrieved from https://www.bloomberg.com/news/articles/2018-12-17/instagram-was-bigger-russian-election-tool-than-facebook-report

Gal, N. (2018). Internet memes. In Barney Warf (Ed.). *The SAGE encyclopedia of the Internet* (pp. 529–530). Thousand Oaks, CA: SAGE.

Gallagher, T. (2019). Trump TV: The Trump campaign's real news update as competitor to cable news. *Visual Communication Quarterly*, *26*(1), 32–43.

Häkkinen, A., & Leppänen, S. (2014). YouTube meme warriors: Mashup videos as political critique. *Tilburg Papers Culture Studies*, *86*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.641.7774&rep=rep1&type=pdf

Hannan, J. (2018). Trolling ourselves to death? Social media and post-truth politics. *European Journal of Communication*, *33*(2), 214–226.

Harmer, E., & Lumsden, K. (2019). Online culture wars: The rise of the alt-right, Trumpism and white masculinities. In K. Lumsden & E. Harmer (Eds.), *Online othering: Exploring digital violence and discrimination on the web* (pp. 35–115). Berlin: Springer.

Heikkilä, N. (2017). Online antagonism of the alt-right in the 2016 election. *European journal of American studies*, *12*(12/2). https://doi.org/10.4000/ejas.12140

Highfield, T., & Leaver, T. (2015). A methodology for mapping Instagram hashtags. *First Monday*, *20*(1), 1–11.

Highfield, T., & Leaver, T. (2016). Instagrammatics and digital methods: Studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, *2*(1), 47–62.

Hine, G. E., Onaolapo, J., De Cristofaro, E., Kourtellis, N., Leontiadis, I., Samaras, R., . . . Blackburn, J. (2017). Kek, cucks, and god emperor trump: A measurement study of 4chan's politically incorrect forum and its effects on the web. In D. Ruths (Ed.), *Eleventh International AAAI Conference on Web and Social Media* (pp. 92–101). Palo Alto, CA: AAAI Press.

Howley, K. (2016). 'I have a drone': Internet memes and the politics of culture. *Interactions: Studies in Communication & Culture*, *7*(2), 155–175.

Jenkins, H. (2006). *Convergence culture*. New York, NY: New York University Press.

Jhaveri, I. (2019). When is a frog not a frog? Building a new digital tool to track political symbols. *Columbia Journalism Review*. Retrieved from https://www.cjr.org/tow_center/political-symbols-tow-center.php

Know Your Meme. (2018). Clown Pepe/Honk Honk/Clown World. *Know Your Meme*. Retrieved from https://knowyourmeme.com/memes/clown-pepe-honk-honk-clown-world

Lasn, K. (2012). *Meme wars: The creative destruction of neoclassical economics*. London: Penguin.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., . . . Schudson, M. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096.

Leskovec, J., Backstrom, L., & Kleinberg, J. (2009). *Meme-tracking and the dynamics of the news cycle*. Paper presented at ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France.

Lonkila, M. (2017). Social network sites and political governance in Russia. In V. Gel'man (Ed.), *Authoritarian modernization in Russia: Ideas, institutions, and policies* (pp. 113–127). London: Routledge.

Martínez-Rolán, X., & Piñeiro-Otero, T. (2016). The use of memes in the discourse of political parties on Twitter: Analysing the 2015 state of the nation debate. *Communication & Society*, *29*(1), 145–159.

Meme. (2020). In *Oxford English dictionary*. Retrieved from https://www.lexico.com/definition/meme

Merrin, W. (2019). President troll: Trump, 4Chan and memetic warfare. In C. Happer, A. Hoskins, & W. Merrin (Eds.), *Trump's media war* (pp. 201–226). London: Palgrave Macmillan.

Miltner, K. (2014). "There's no place for lulz on LOL-Cats": The role of genre, gender and group identity in the interpretation and enjoyment of an Internet meme. *First Monday*, *19*(8). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/5391/4103

Nissenbaum, A., & Shifman, L. (2017). Internet memes as contested cultural capital: The case of 4chan's/b/board. *New Media & Society*, *19*(4), 483–501.

Pauca, V. P., Shahnaz, F., Berry, M. W., & Plemmons, R. J. (2004). Text mining using non-negative matrix factorizations. In M. Berry, C. Kamath, U. Dayal, & D. Skillicorn (Eds.), *Proceedings of the 2004 SIAM International Conference on Data Mining* (pp. 452–456). Philadelphia, PA: Society for Industrial and Applied Mathematics.

Revesz, R. (2016, September 13). Hillary Clinton attacks Donald Trump for posting Pepe the Frog meme. *The Independent*. Retrieved from https://www.independent.co.uk/news/world/americas/donald-trump-hillary-clinton-pepe-frog-instagram-breitbart-white-supremacist-alex-jones-milo-a7240581.html

Roose, K. (2017, July 5). How a CNN investigation set off an Internet meme war. *The New York Times*. Retrieved from https://www.nytimes.com/2017/07/05/business/how-a-cnn-investigation-set-off-an-internet-meme-war.html

Rose, G. (2016). *Visual methodologies: An introduction to researching with visual materials*. California, CA: SAGE.

Schreckinger, B. (2017, March/April). World war meme. *Politico*. Retrieved from http://politi.co/2F5wv98

Seiffert-Brockmann, J., Diehl, T., & Dobusch, L. (2018). Memes as games: The evolution of a digital discourse online. *New Media & Society*, *20*(8), 2862–2879.

Shifman, L. (2012). An anatomy of a YouTube meme. *New Media & Society*, *14*(2), 187–203.

Shifman, L. (2014). *Memes in digital culture*. Cambridge, MA: MIT Press.

Tay, G. (2014). Binders full of LOLitics: Political humour, internet memes, and play in the 2012 US Presidential Election (and beyond). *European Journal of Humour Research*, *2*(4), 46–73. https://doi.org/10.7592/EJHR2014.2.4.tay

Townsend, L., & Wallace, C. (n.d.). *Social media research: A guide to ethics*. Aberdeen: The University of Aberdeen. Retrieved from https://www.gla.ac.uk/media/media_487729_en.pdf

Trump, D. [realDonaldTrump]. (2018, October 21). Facebook has just stated that they are setting up a system to "purge" themselves of Fake News. Does that mean CNN will finally be put out of business? [Tweet]. Retrieved from https://twitter.com/realDonaldTrump/status/1054142516881248258

Vaca, C. K., Mantrach, A., Jaimes, A., & Saerens, M. (2014, April). A time-based collective factorization for topic discovery and monitoring in news. In P. Ipeirotis (Eds.), *Proceedings of the 23rd International Conference on World Wide Web* (pp. 527–538). New York, NY: ACM.

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

Wiggins, B. E., & Bowers, G. B. (2015). Memes as genre: A structurational analysis of the memescape. *New Media & Society*, *17*(11), 1886–1906.

Williams, Z. (2016, June 12). (((Echoes))): Beating the far-right, two triple-brackets at a time. *The Guardian*. Retrieved from https://www.theguardian.com/technology/shortcuts/2016/jun/12/echoes-beating-the-far-right-two-triple-brackets-at-a-time

Wimmer, R., & Dominick, J. (2013). *Mass media research: An introduction*. Boston, MA: Wadsworth, Cengage Learning.

Woolley, S. C., & Guilbeault, D. R. (2017). *Computational propaganda in the United States of America: Manufacturing consensus online* (Computational Propaganda Research Project Working Paper No.2017.5). Oxford: University of Oxford.

Yang, G. (2014). Political contestation in Chinese digital spaces: Deepening the critical inquiry. *China Information*, *28*(2), 135–144.

About the Author

Ahmed Al-Rawi is an Assistant Professor of News, Social Media, and Public Communication at the School of Communication at Simon Fraser University, Canada. He is the Director of the Disinformation Project that empirically examines fake news discourses in Canada on social media and news media. Al-Rawi authored five books and over 90 peer reviewed book chapters and articles published in a variety of journals like *Information, Communication & Society*, *Online Information Review*, *Social Science Computer Review*, and *Telematics & Informatics*.

Article

Fighting Deepfakes: Media and Internet Giants' Converging and Diverging Strategies Against Hi-Tech Misinformation

Ángel Vizoso *, Martín Vaz-Álvarez and Xosé López-García

Faculty of Communication Sciences, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain;
E-Mails: angel.vizoso@usc.es (A.V.), martin.vaz.alvarez@usc.es (M.V.-A), xose.lopez.garcia@usc.es (X.L.-G.)

* Corresponding author

Abstract
Deepfakes, one of the most novel forms of misinformation, have become a real challenge in the communicative environment due to their spread through online news and social media spaces. Although fake news have existed for centuries, its circulation is now more harmful than ever before, thanks to the ease of its production and dissemination. At this juncture, technological development has led to the emergence of deepfakes, doctored videos, audios or photos that use artificial intelligence. Since its inception in 2017, the tools and algorithms that enable the modification of faces and sounds in audiovisual content have evolved to the point where there are mobile apps and web services that allow average users its manipulation. This research tries to show how three renowned media outlets—*The Wall Street Journal*, *The Washington Post*, and *Reuters*—and three of the biggest Internet-based companies—Google, Facebook, and Twitter—are dealing with the spread of this new form of fake news. Results show that identification of deepfakes is a common practice for both types of organizations. However, while the media is focused on training journalists for its detection, online platforms tended to fund research projects whose objective is to develop or improve media forensics tools.

Issue
This article is part of the issue "Disinformation and Democracy: Media Strategies and Audience Attitudes" edited by Pere Masip (University Ramon Llull, Spain), Bella Palomo (University of Málaga, Spain) and Guillermo López (University of Valencia, Spain).

1. Introduction

The implementation of artificial intelligence in technologically mediated communicative processes in the networked society poses new challenges for journalistic verification. Simultaneously, it has enhanced different stages of news production systems. The effects of the technology that houses artificial intelligence are present both in the communicative flows and in a large part of the socialization dynamics. Hence, the threat introduced by the emergence of deepfakes, doctored videos by using artificial intelligence, arises as one of the most recent hazards for journalistic quality and news credibility. Although

deepfakes are not only a concern for journalism, their existence has raised the uncertainty among users when trying to access news content. Likewise, the increasing sophistication of this form of fake news has put professionals on alert (Vaccari & Chadwick, 2020).

Misinformation has increased its relevance over the last few years, having now a major significance in the public agenda (Vargo, Guo, & Amazeen, 2018). In consequence, the number of projects and measures for counteracting this phenomenon has grown considerably. Example of this could be the *Action Plan Against Disinformation* developed by the European Commission (2018). Media and journalists are aware about how the

success of hoaxes undermines democracy and its reliability (Geham, 2017). Therefore, they try to react with actions that facilitate transparency and the fulfilment of their professional and ethical rules, like fact checking (Lowrey, 2017). This is an issue on which different lines of thinking have been opened. All of them try to counterbalance the result of misinformative political and social trends that became significative in 21st century societies (McNair, 2017) in a context where social media plays a central role as a space for the generation and dissemination of fake news and the consequences that this entails (Nelson & Taneja, 2018).

Techniques that guarantee the information verification's efficiency—one of the core elements of journalism since its consolidation as a communicative technique in the modern age (Kovach & Rosenstiel, 2014)—are looking inside technological innovation for tools with the ability to support professionals in their daily tasks. It is true that the norms followed for producing accurate informative pieces are in some cases unclear and nuanced (Shapiro, Brin, Bédard-Brûlé, & Mychajlowycz, 2013). Nonetheless, journalism should not retain antiquated verification techniques, but should rather update them to computational methods in order to evaluate dubious information (Ciampaglia et al., 2015). There are currently revamped verification systems with fact-checking techniques. Those may contribute to the elaboration of news pieces that, after the application of a complex group of cultural, structural, and technological relations would show the legitimation of news in the digital age (Carlson, 2017, p. 13). Although a high level of mistrust remains, some techniques used in these information verification services are able to build a bigger reliance by the users (Brandtzaeg & Følstad, 2017).

Furthermore, this scenario has seen the emergence of new proposals for renewed professional practices and profiles (Palomo & Masip, 2020). This could be the case of constructive journalism, whose objective is regaining the lost trust of the media (McIntyre & Gyldensted, 2017). This is a journalistic movement that explores new paths. However, it will take time to see if these new approaches fit in the frame of emerging journalism in the Third Millennium, with a clear commitment to social service, transparency and accuracy.

2. Literature Review

2.1. Misinformation Through the Ages

Falsehood, fantasy and fake news have walked along with the development of communication and journalism, initiating discussions about its practice and its role in society (McNair, 2017). Although there are evidence of misinformation since the Roman Empire (Burkhardt, 2017), its major development took place with the invention of the print in the 15th century. The possibility of disseminating written information in a faster and easier way made possible the circulation of falsehood too.

Hence, the advent of new means of communication, increased the presence of deliberated false content, not always with harmful purposes. In this regard, one of the greatest examples of misinformation of our times was the radio broadcast of *The War of the Worlds* directed by Orson Welles on October 30th, 1938. That radio show was followed by thousands of listeners, and some of them believed that the Earth was under an alien attack, thanks to the narration of Welles' cast (Gorbach, 2018).

This radio show wanted to entertain the audience using an alteration of reality. However, manipulation of the truth has been used as a weapon in military conflicts over the centuries in order to ascribe malicious acts or characteristics to the enemy (Bloch, 1999, p. 182). A good example of this use of misinformation was the sequence of news published after the explosion in the boilers of the United States Navy ship USS Maine on February 15th, 1898. In the middle of the fight for being the most read against Joseph Pulitzer's *The New York World*, William Randolph Hearst, editor of the *New York Journal*, sent a journalist to Cuba with the objective of telling the readers the details of a Spanish attack to this ship. Thus, when the correspondent arrived at the island reported that alleged attack did not exist. Nonetheless, the newspaper published a series of stories detailing the attack—even when they knew they were not accurate—causing a climate of hate against Spain and acceptance of the coming war. Finally, The United States declared the war against Spain (Amorós, 2018, p. 34). After this conflict, misinformation continued to be used against the enemy in war times. Thus, it is possible to identify strategies of its use in recent conflicts like the World War I and World War II, the Vietnam War or the Gulf War (Peters, 2018).

2.2. Fake News as a Threat to Journalism

Falsity has cast a shadow over the discipline of communication throughout history. One of the newest forms of misinformation is fake news, pieces that imitate the appearance of journalistic information, but deliberately altered (Rochlin, 2017). This form of deception has coexisted with true news. However, the current communicative scenario, marked by the utilisation of high speed and low contrast means of communication—and among all social media—provides a fertile soil for the dissemination of any form of misinformation (Lazer et al., 2018).

Platforms like Facebook or Twitter are now among the primary news sources for Internet users (Bergström & Jervelycke-Belfrage, 2018). Fake news producers are aware of this fact. As a result, they have made the web the main channel for false content distribution, taking advantage of the possibility of communicating anonymously provided by certain spaces (Vosoughi, Roy, & Aral, 2018). Furthermore, fake news producers have the chance of reaching as large audiences as consolidated journalistic brands (Fletcher, Cornia, Graves, & Nielsen, 2018), which makes the verification of this falsities more difficult.

During the last few years, there have been different proposals for classifying fake news. Among them, the one developed by Tandoc, Lim, and Ling (2018) is perhaps the most exhaustive: news satire, a very common form of fake news with a large presence in magazines, websites and radio or TV shows; news parody, which shares some of the characteristics of news satire, but it is not based on topical issues. These pieces are fictional elements specifically produced for certain purposes; news fabrication, unfounded stories that try to imitate the structure of news published by legacy media. The promoters of these pieces try to deceive by blending them among the truthful ones; photo manipulation—alteration of images—and more recently videos—for building a different reality; advertising and public relations—dissemination of advertising by masking it to look as journalistic reporting; and propaganda—stories from political organizations with the objective of influencing citizens' opinion on them. Like some of the previous ones, they imitate the formal structure of news pieces.

Regarding its formal structure, fake news try to imitate news items' formal appearance. Thus, visual codes and elements like headlines, images, videos hypertext and texts conceived like journalistic pieces are common features of this misinformation strategies (Amorós, 2018, p. 65). Nonetheless, its major particularity is that fake news tries to attack the readers' previous opinion, especially on controversial issues related to racism, xenophobia, homophobia and other forms of hate (Bennett & Livingston, 2018; Waisbord, 2018). This connection makes possible the rapid replication of such content thanks to the ease of sharing through spaces like social media platforms. Thus, episodes like electoral processes (Lowrey, 2017), or more recently the Covid-19 pandemic (Salaverría et al., 2020; Shimizu, 2020), resulted in a deep growth of fake news circulation, at times using simple methods but at times taking advance of the most advanced technology.

2.3. Deepfake: A Novel Form of Fake News

Deepfakes, a combination of 'deep learning' and 'fake' (Westerlund, 2019), are "highly realistic and difficult-to-detect digital manipulations of audio or video" (Chesney & Citron, 2019). It can be defined as "a technique used to manipulate videos using computer code" (Fernandes et al., 2019, p. 1721), generally replacing the voice or the face of a person with the face of the voice of another person. Although the photo and video manipulation have existed for a long time, the use of artificial intelligence methods for these purposes has augmented the number of fakes and its quality. Some of these videos are humorous, but the majority of them are damaging (Maras & Alexandrou, 2019). Hence, this is a recent movement whose beginnings date back to 2017, starting then a rapid popularisation until now (Deeptrace Labs, 2018, pp. 2–4).

This technique is the result of using Generative Adversarial Networks, algorithms designed to replace human faces or voices in thousands of images and videos in order to make them as realistic as possible (Li, Chang, & Lyu, 2018). The main advantage of these algorithms is that these systems are learning how to improve themselves by creating deepfakes. Therefore, future creations will be improved thanks to past experiences. This feature makes this misinformation procedure more dangerous, especially due to the emergence of mobile apps and computer programmes that allow users without computer programming training to produce deepfakes (Nirkin, Keller, & Hassner, 2019; Schwartz, 2018).

Farid et al. (2019, pp. 4–6) tried to label the different forms adopted by deepfakes in four categories: 1) face replacement or face swapping—this method involves changing one person's face, the source, for another one, the target; 2) face re-enactment—manipulation of the features of the features of one person's face like the movement of the mouth or the eyes, among others; 3) face generation—creation of a completely new face using all the potential provided by Generative Adversarial Networks; and 4) speech synthesis—alteration of someone's discourse in terms of cadence and intonation, or generation of a completely new one.

As with other technologies, the same algorithms used for creating deepfakes could have a beneficial application in the field of psychology, building digital synthetic identities for voiceless users; or in robot sketches through advanced facial recognition for law enforcement, for example (Akhtar & Dasgupta, 2019; Zhu, Fang, Sui, & Li, 2020). Notwithstanding, its use seems to be more harmful than beneficial nowadays with examples of the use of these technologies in acts of fraud and crime (Stupp, 2019).

Hence, one of the biggest challenges of deepfakes is to find out how to counteract them knowing that the debunking methods' development is always late regarding the production of misinformation (Galston, 2020). However, a great deal of effort has been made—and is still made—to develop technology-based tools for detecting and correcting it, both from public and private organizations (Deeptrace Labs, 2018, p. 2). These tools will be helpful in almost all areas of communication, especially for journalism.

2.4. Fact-Checking: Journalism's Response to the Misinformation Wave

In light of the above, verified information seems to be a necessity in our communicative context (Ekström, Lewis, & Westlund, 2020), especially because disruptive episodes like the coronavirus outbreak resulted in a clear increase of citizens' informative consumption (Masip et al., 2020). Furthermore, political communication has shifted to a model in which political leaders share their messages online instead of doing it through traditional media (López-García & Pavía, 2019).

At this juncture, the media has increased the importance of verification processes for correcting both internal and external errors (Geham, 2017). Consequently, a new professional profile—the fact-checker—has emerged with the mission of debunking misinformation and prevent audiences of its consumption. These professionals try to go to the origin of an information or a claim for gathering all the available data and contrasting it (Graves, 2016, p. 110). Fortunately, journalists have also benefited from the development of new technological tools designed for verifying images, videos or websites in an efficient manner (Brandtzaeg, Lüders, Spangenberg, Rath-Wiggins, & Følstad, 2016).

Although verification has always been part of any journalistic process, the rapid growth of the fake news phenomenon over the past few years made this activity more important than ever. Thus, the census created by the University of Duke Reporters' Lab counts now almost 300 fact-checkers in more than 60 countries by the by middle of 2020, a hundred more than on the same date in 2019 (Stencel & Luther, 2020). Regarding this, it is possible to talk about fact-checking as a transnational movement (Graves, 2018) were both legacy and independent media organizations try to restore the trust lost by the media (Bennett & Livingston, 2018).

3. Method

The starting point of this research will be the application of the Systematic Literature Review method (Kitchenham, 2004) as a method to set an approach on how deepfakes are being addressed and studied. Due to the novelty of this reality, this method will let us understand in an exhaustive way (Codina, 2017) what are researchers doing to assess this phenomenon and what efforts are being done to stop its spread.

Hence, our method consisted in the following phases: 1) topic identification—'deepfake' and 'deep fake'—and the period of analysis—all the available literature; 2) source selection—Web of Science's SCI-Expanded, SSCI, CPCI-S, CPI-SS, CPCI-SSH, and Scopus; 3) search in databases—the selection of Web of Science and Scopus is justified by the importance of these two databases, which contain the most relevant contributions for the Social Sciences field in general and deepfakes specifically; and 4) identification of the studied variables for each item—descriptive data (article title, date, journal or conference, number of authors, and keywords), type of study, research techniques (observation, survey, interview, content analysis, case study, experimental or

non-specified), principal contribution, DOI or URL, and institution and country.

This search resulted in 54 different research items: 28 presented at international conferences and 26 published in academic journals—all of that after deleting duplicities and texts that did not fit the criteria, such as editorial articles, call for papers, or interviews, among others. These 54 examples comprise our sample that will be addressed in the next section in order to understand the path followed by researchers on this subject.

Concerning the second stage of our study, it will analyse the approach taken by three renowned media outlets and news agencies—*The Washington Post*, *The Wall Street Journal*, and *Reuters*—and three of the most important Internet platforms—Google, Facebook and Twitter—in neutralizing the spread of deepfakes. Thus, case study of these six organizations will be applied in order to understand how they are managing to identify, label and notify deepfakes through different approaches—protocols, use of technology, collaboration with institutions, and funding of innovative projects. This will be done through the analysis of the available reports and statements of these six organizations.

Consequently, the main goal of our study will be to identify the coincidences and disparities in the strategy of three major media outlets and three of the most important online platforms when trying to stop the diffusion of deepfakes. This will be relevant in order to understand if six of the main representatives from these two communicative fields are joining efforts and strategies in limiting or not its spread, and how these procedures could be improved.

4. Findings

4.1. Results of the Systematic Literature Review

We will start by depicting the state of the research on deepfakes, especially the contributions indexed in the two main databases—Web of Science and Scopus. As shown in Table 1, research about this issue started in 2018 with four conference papers. However, it was quintupled in 2019, and during the first half of 2020 almost a half more of works on deepfakes than the previous year were published. Furthermore, the most salient element of this table is that this form of fake news used to have presence at conferences, but in 2020 they become a topic addressed in academic journals too. Nonetheless, it is necessary to note that the situation resulting from the Covid-19 pandemic

Table 1. Evolution of the studies on deepfake indexed in WoS and Scopus.

Year	Conference paper	Journal article	Total
2018	4	0	4
2019	16	5	21
2020 (1st half)	8	21	29

has provoked the cancellation or postponement of many conferences.

Regarding the authorship of this research, the most common approach is the participation of three authors. Thus, the arithmetic mean—3,13 authors—and the mode—14 articles have three authors—serve to confirm this. Also concerning the authorship, researchers from 24 countries were identified, most of them located in the United States and Asia—China, Japan, South Korea, India, or Taiwan.

Finally, this review shows a large degree of uniformity concerning the type of studies published on deepfakes. Almost all the reviewed articles and conference papers take a descriptive approach. This is because 32 of the items are the result of experimenting with new tools and algorithms to counteract it. Another important group of research is review articles on deepfake detection and prevention or even about legal framework and legal concerns of this form of misinformation, something that was found 21 times.

In sum, the novelty of deepfake implies a certain degree of youth for its research. At present, it is possible to see two trends: Studies that present new forms to stop its spread, or studies that try to create context on its emergence and development.

4.2. Counteracting Deepfakes at The Wall Street Journal, The Washington Post and Reuters

Recent advances in artificial intelligence and their democratisation have allowed average users to create deepfakes. This represents a major challenge for our society due to the potential harmful impact of these creations, especially before electoral processes. Looking to the United States 2020 general election, *The Wall Street Journal* has created a division of 21 journalists whose unique objective is detecting, labelling and debunking misinformation, particularly deepfakes (Southern, 2019). This team is a joint effort of Standards & Ethics and R&D departments, and this work is very linked to the use of technology with presence of journalist with video, photo, visuals, research and news experience that have been trained for deepfake detection (Marconi & Daldrup, 2018). Furthermore, *The Wall Street Journal* provides specialized training in fake news and deepfake identification in partnership with different researchers. This has led to the development of a protocol to find examples of this kind of misinformation with three stages: source examination (contact with the source, authorship identification, and metadata check, among others), search for older versions of the footage available online, and footage examination with video and photo editing programs.

Meanwhile, *The Washington Post* has applied to deepfake detection very similar criteria to other fake news detection. Thus, the *The Washington Post* has added video experts to the tasks developed by the team led by Glenn Kessler—also known as 'The Fact Checker' (Kessler, 2019). The most important contribution of this publication regarding this problem is the elaboration of a taxonomy to classify and label deepfakes. *The Washington Post* was also pioneering in the use of scales to highlight the degree of truth and lie of any content. Regarding doctored videos, the newspaper sets out three categories of manipulation (Ajaka, Samuels, & Kessler, 2019): missing context (presentation of the video without context or with a context intentionally altered), deceptive editing (rearrangement and edition of the video in certain parts or details), and malicious transformation (complete manipulation and transformation of the footage resulting in a completely new fabrication).

A third approach to this reality could be the one adopted by the news agency *Reuters*. The news services provider reports its awareness and concern on the deepfake spread (Crosse, n.d.). Hence, it has started a collaboration with Facebook for detecting as much doctored user-generated content as possible among all the videos and photos that run on the platform (Patadia, 2020).

In this regard, *Reuters* has started a blog whose objective is verifying doctored materials in English and Spanish. All of that with the objective of debunking as much information as possible ahead of the 2020 United States election. This is a clear example of the emerging collaboration among technological platforms and the media, a joint effort in trying to stop the rapid growth of fake news and deepfakes in such significant moments like a presidential run-up.

4.3. Internet Giants' Strategies Against Deepfakes

The spread of falsehood through social media platforms and other Internet spaces is now a challenge for providers like Google, Facebook or Twitter. As a result, over the last few months they have started different initiatives whose unique objective is finding efficient ways to detect and stop the misinformation and, more recently, deepfakes.

Regarding this, these three companies show different approaches against this matter. Google, for instance, has made available to the research community a large set of manipulated and non-manipulated videos (Dufour & Gully, 2019). With this initiative, they want to help in the development of identification techniques by taking advantage of the great amount of information saved in their files. In addition, they collaborate with the Defense Advanced Research Projects Agency to fund different researchers that are developing media forensic tools.

On the other hand, Facebook is financing different research projects within its 'Deepfake Detection Challenge.' This initiative, boosted by companies like Facebook, Microsoft and Amazon Web Services and research units from various universities across the United States, tries to assist researchers that are working on the development of artificial intelligence-based deepfake detection tools. Thus, a corpus of more than 100,000 videos was available to these researchers that

fight for presenting useful mechanism in order to win different awards.

Furthermore, Mark Zuckerberg's social network tries to counteract this form of misinformation by deleting doctored videos or photos, or labelling it as fake news with the help of fact-checking media outlets (Bickert, 2020). This is particularly important for those related to the 2020 United States run-up due to the influence that fake news could have in this process.

Finally, Twitter shows a simpler approach towards this problem. They summarize their strategy in the following four rules (Harvey, 2019): Identification through a notice of Tweets with manipulated content, warning of its manipulated condition before sharing it, inclusion of a link to news articles or other verified sources in which users can find out why and how the content has been doctored, and elimination of all that manipulated content potentially harmful or threatening to anyone's safety.

These diverging strategies on behalf of the major online platforms are in part the product of self-regulated methods for fighting deepfakes, as there is still incipient intervention on behalf of the states in regulating content on social media and other outlets. The question to be asked here is whether it is the online platforms' sole responsibility to tackle misinformation or if there are any social interests in this situation for which other public entities should allocate resources to.

The European Commission already pointed out in 2018 the need for governments to invest in research and detection of misinformation, while also prompting these to hold social media companies accountable (European Commission, 2018). So far, in the last two years the EU has launched a series of initiatives to tackle the issue: a code of practice against disinformation, the creation of the Social Observatory for Disinformation and Social Media Analysis and the set-up of the Rapid Alert System, among other R&D projects such as PROVENANCE, SocialTruth, EUNOMIA or WeVerify (European Commission, 2020). Despite a lot of efforts being made to avoid the spread of misinformation in the EU, deepfakes are still not as much on the agenda as other academics are asking for, while also describing their worry for seemingly understaffed programs (Bressnan, 2019). Measures taken by countries to prompt social media companies in acting against fake news contain different levels of intervention and are mainly dedicated to counteracting disinformation in political advertisement. France and Germany, for example, require online platforms to establish an easily accessible and visible way for users to flag false information, while Australia requires all paid electoral advertising, including advertisements on social media, to be authorized and to contain an authorization statement (Levush, 2019). In the United States, some states have already taken specific measures to counter deepfakes, although these are still merely reactive and not preventive, such as Texas passing a law that criminalizes publishing and distributing deepfake videos with the intention to harm

a candidate during the electoral process; or California, where a law was passed last October making it illegal for anyone to intentionally distribute deepfakes for deceiving voters or perjure a candidate (Castro, 2020).

The implications of these incipient interactions between governments and social media companies might have relevant governance questions in the forthcoming years, all the while these companies are also starting to take new approaches to their governance structures, such is the case of the Facebook, who set up the Independent Oversight Board, which "aims to make Facebook more accountable and improve the decision-making process," in the words of Nick Clegg, currently Facebook's VP of Global Affairs and Communications and former Deputy Prime Minister of the United Kingdom (Moltzau, 2020).

5. Discussion

As shown in the previous section, the media and Internet platforms have initiated different strategies to fight misinformation and, more particularly, the spread of deepfakes. In this regard, there are some similarities and differences among the strategies of these two communicative sectors.

First of all, it seems clear that the collaboration among platforms and media outlets increases over time. Example of this could be the agreements among *Reuters* and Facebook whose objective is to detect fake news and share its correction. Furthermore, other fact-checking organizations collaborate with this social network in labelling false content and warn users about this.

Another coincidence is the use of technology as a weapon in the battle against deepfakes. Both news media and digital platforms have understood that high technology and the use of algorithms as powerful as those used for creating fakes is the only chance to counteract them. Thus, media outlets are increasingly training journalists and interdisciplinary teams in the use of these mechanisms that allow them to identify this form of misinformation.

The third match could be the growing synergies between the academic and communicative sides. Thus, media outlets and platforms try to collaborate with researchers and institutions specialized in fake news detection, both in training and to apply their methods.

Regarding the divergences when dealing deepfakes, online platforms are able to fund research projects whose objective is developing artificial intelligence-tools for identifying this form of fake news. The media, however, does not have such possibilities due to the expenditure of these activities.

Another difference in dealing with this issue could be that the media use to correct misinformation instead deleting it. As shown before, some of the social media platforms have the elimination of doctored content among their strategies. This presents a clear challenge. Although deleting manipulated videos or photos ends

with the problem for future or potential users, does not for those users that have seen them. In contrast, labelling these materials as false or manipulated—the approach followed by verification media outlets—could be helpful for future users.

6. Conclusions

Deepfakes have become a reality in our communicative system. Media outlets and Internet services providers try to counteract it with different outlooks. However, the development of the techniques for producing misinformation seems to advance faster than those for debunking it. Regarding this, the available research is mainly focused on two aspects: On one hand detection tools, and on the other hand, the implications of this form of fake news for democracy and national security. The fact that so far only big technological giants are capable of introducing hi-tech expensive solutions for fighting deepfakes motivates that the available mass of research on the subject is fundamentally dedicated to address the questions raised by these corporations which are mainly technological. On the other hand, journalism focused media, which are not able to invest large amounts of money on deepfake detection are therefore unable to push their concerns into the research agenda. For this reason, producing research on the implications of deepfakes for journalism and under journalistic premises presents itself as elemental, as well as further investigation on how the media trains its professionals for detecting advanced misinformation.

The novelty of this deceiving technique provokes its understudied situation, but the constant growth of works on this matter show that it will be an important field for researchers on misinformation and media forensics in the following years. However, the study is able to show to some extent that the media and digital platforms' have notable similarities and differences when it comes to their strategies. This could be due to the different nature of their business models, but nevertheless sometimes it seems to be a matter of investment. Digital platforms have joined efforts with technological, academic or entrepreneurial partners, spending large amounts of money in this field, which is something that many media outlets cannot accomplish.

However, these collaborations go in some cases beyond the private sector. The growing pressure on behalf of states and governments, pushing for different levels of regulation for these social media companies is being noted around the world. Countries have started to put mechanisms in place that allow for a more scrutinized assessment of these corporations, since there is an increasingly wider knowledge on how their inner social interaction tools can affect basic democratic elements (elections, political advertisement, etc.). A representative example of this is Facebook's recent signing of Nick Clegg, former Deputy Prime Minister of the United Kingdom, in a move that seems to state its willingness to reform its governance structures towards more state-sensible policies. The tendency seems to describe a tension between the states and governments' purpose to regulate, and the effort of these companies to maintain their self-regulation.

Nonetheless, it is necessary to highlight the limitations of our study and our sample. As the reader may have noticed, our sample shows a western-oriented cultural bias, both for the media and digital platforms analysis. This could make it difficult to generalize our findings to similar organizations from other territories. However, this research presents a descriptive approach whose objective is to analyse how major communication companies are trying to counteract the spread of deepfakes, the most novel and hi-tech-based form of misinformation. Furthermore, the novelty and high level of sophistication of this new way of producing fake news requires a similar level of human and technological resources for the debunking process, something that only large companies—mainly United States-based—can afford at this particular time. In the future, it will be interesting to follow the progress of both deepfakes production and identification strategies as well as the paths adopted by researchers in this subject. As mentioned, this is a recent phenomenon, but its rapid growth will make necessary the setting up of protocols for containing its spread in both media outlets and online platforms.

Future lines of research regarding this topic might include a deep assessment of how journalists perceive deepfakes in their daily routines, and the challenge of verification for journalists, which includes questions of technological literacy and guides for good practices.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

References

Ajaka, N., Samuels, E., & Kessler, G. (2019). The Washington Post's guide to manipulated video. *The Washington Post.* Retrieved from https://www.

washingtonpost.com/graphics/2019/politics/fact-checker/manipulated-video-guide

Akhtar, Z., & Dasgupta, D. (2019). A comparative evaluation of local feature descriptors for deepfakes detection. In *2019 IEEE international symposium on technologies for homeland security, HST 2019*. Woburn, MA: IEEE. https://doi.org/10.1109/hst47167.2019.9033005

Amorós, M. (2018). *Fake news: La verdad de las noticias falsas* [Fake news: The truth of false news]. Barcelona: Plataforma Editorial.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Bergström, A., & Jervelycke-Belfrage, M. (2018). News in social media. *Digital Journalism*, *6*(5), 583–598. https://doi.org/10.1080/21670811.2018.1423625

Bickert, M. (2020, January 6). Enforcing against manipulated media. *Facebook*. Retrieved from https://about.fb.com/news/2020/01/enforcing-against-manipulated-media

Bloch, M. (1999). *Historia e historiadores* [History and historians]. Madrid: Akal.

Brandtzaeg, P. B., & Følstad, A. (2017). Trust and distrust in online fact-checking services. *Communications of the ACM*, *60*(9), 65–71. https://doi.org/https://doi.org/10.1145/3122803

Brandtzaeg, P. B., Lüders, M., Spangenberg, J., Rath-Wiggins, L., & Følstad, A. (2016). Emerging journalistic verification practices concerning social media. *Journalism Practice*, *10*(3), 323–342. https://doi.org/10.1080/17512786.2015.1020331

Bressnan, S. (2019, September 19). Can the EU prevent deepfakes from threatening peace? *Carnegie Europe*. Retrieved from https://carnegieeurope.eu/strategiceurope/79877

Burkhardt, J. M. (2017). Chapter 1: History of fake news. *Library Technology Reports*, *53*(8), 5–9.

Carlson, M. (2017). *Journalistic authority: Legitimating news in the digital era*. New York, NY: Columbia University Press.

Castro, D. (2020). Deepfakes are on the rise: How should government respond? *Government Technology*. Retrieved from https://www.govtech.com/policy/Deepfakes-Are-on-the-Rise-How-Should-Government-Respond.html

Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war. *Foreign Affairs*. Retrieved from https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakes-and-new-disinformation-war

Ciampaglia, G. L., Shiralkar, P., Rocha, L. M., Bollen, J., Menczer, F., & Flammini, A. (2015). Computational fact checking from knowledge networks. *PLOS ONE*, *10*(6). https://doi.org/10.1371/journal.pone.0128193

Codina, L. (2017, April 20). Revisiones sistematizadas y cómo llevarlas a cabo con garantías: Systematic reviews y SALSA Framework [Systematized reviews and how to conduct it with guarantees: Systematic reviews and SALSA Framework]. *Lluiscodina*. Retrieved from https://www.lluiscodina.com/revision-sistematica-salsa-framework

Crosse, G. (n.d.). Truth and authentication: The frightening world of video fakes. *Reuters*. Retrieved from https://agency.reuters.com/en/insights/articles/articles-archive/truth-and-authentication-the-frightening-world-of-video-fakes.html

Deeptrace Labs. (2018). *The state of deepfakes: Reality under attack*. Amsterdam: Deeptrace Labs. Retrieved from https://s3.eu-west-2.amazonaws.com/rep2018/2018-the-state-of-deepfakes.pdf

Dufour, N., & Gully, A. (2019, September 24). Contributing data to deepfake detection research. *Google Blog*. Retrieved from https://ai.googleblog.com/2019/09/contributing-data-to-deepfake-detection.html

Ekström, M., Lewis, S. C., & Westlund, O. (2020). Epistemologies of digital journalism and the study of misinformation. *New Media & Society*, *22*(2), 205–212. https://doi.org/10.1177/1461444819856914

European Commission. (2018). *Action plan against disinformation*. Brussels: European Commission. Retrieved from https://eeas.europa.eu/sites/eeas/files/action_plan_against_disinformation.pdf

European Commission. (2020). *Tackling online disinformation*. Brussels: European Commission. Retrieved from https://ec.europa.eu/digital-single-market/en/tackling-online-disinformation

Farid, H., Davies, A., Webb, L., Wolf, C., Hwang, T., Zucconi, A., & Lyu, S. (2019). *Deepfakes and audio-visual disinformation*. London: Centre for Data Ethics and Innovation. Retrieved from https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai/snapshot-paper-deepfakes-and-audiovisual-disinformation

Fernandes, S., Raj, S., Ortiz, E., Vintila, I., Salter, M., Urosevic, G., & Jha, S. (2019). Predicting heart rate variations of deepfake videos using neural ODE. In *Proceedings: 2019 international conference on computer vision workshop, ICCVW 2019* (pp. 1721–1729). Seoul: IEEE. https://doi.org/https://doi.org/10.1109/iccvw.2019.00213

Fletcher, R., Cornia, A., Graves, L., & Nielsen, R. K. (2018). *Measuring the reach of fake news and online disinformation in Europe*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from https://reutersinstitute.politics.ox.ac.uk/our-research/measuring-reach-fake-news-and-online-disinformation-europe

Galston, W. A. (2020). *Is seeing still believing? The deepfake challenge to truth in politics*. Washington, DC: Brookings. Retrieved from https://www.brookings.edu/research/is-seeing-still-believing-the-deepfake-

challenge-to-truth-in-politics

Geham, F. (2017, July 12). *Le fact-checking: Une réponse à la crise de l'information et de la démocratie* [Fact-checking: The answer to the information and democracy crisis]. Paris: Fondapol. Retrieved from http://www.fondapol.org/etude/farid-gueham-le-fact-checking-une-reponse-a-la-crise-de-linformation-et-de-la-democratie

Gorbach, J. (2018). Not your grandpa's hoax: A comparative history of fake news. *American Journalism*, *35*(2), 236–249. https://doi.org/10.1080/08821127.2018.1457915

Graves, L. (2016). *Deciding what's true*. New York, NY: Columbia University Press.

Graves, L. (2018). Boundaries not drawn. *Journalism Studies*, *19*(5), 613–631. https://doi.org/10.1080/1461670X.2016.1196602

Harvey, D. (2019, November 11). Help us shape our approach to synthetic and manipulated media. *Twitter.* Retrieved from https://blog.twitter.com/en_us/topics/company/2019/synthetic_manipulated_media_policy_feedback.html

Kessler, G. (2019, June 25). Introducing The Fact Checker's guide to manipulated video. *The Washington Post.* Retrieved from https://www.washingtonpost.com/politics/2019/06/25/introducing-fact-checkers-guide-manipulated-video

Kitchenham, B. (2004). *Procedures for performing systematic reviews*. Keele: Keele University. Retrieved from http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf

Kovach, B., & Rosenstiel, T. (2014). *The elements of journalism, revised and updated: What newspeople should know and the public should expect* (3rd ed.). New York, NY: Three Rivers Press.

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., . . . Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Levush, R. (2019). Government responses to disinformation on social media platforms. *Library of Congress*. Retrieved from https://www.loc.gov/law/help/social-media-disinformation/compsum.php

Li, Y., Chang, M.-C., & Lyu, S. (2018). In ictu oculi: Exposing AI created fake videos by detecting eye blinking. In *2018 10TH IEEE international workshop on information forensics and security (WIFS)* (pp. 1–5). Hong Kong: IEEE. https://doi.org/10.1109/wifs.2018.8630787

López-García, G., & Pavía, J. M. (2019). Political communication in election processes: An overview. *Contemporary Social Science*, *14*(1), 1–13. https://doi.org/10.1080/21582041.2018.1479040

Lowrey, W. (2017). The emergence and development of news fact-checking sites. *Journalism Studies*, *18*(3), 376–394. https://doi.org/10.1080/1461670X.2015.1052537

Maras, M.-H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. *The International Journal of Evidence & Proof*, *23*(3), 255–262. https://doi.org/10.1177/1365712718807226

Marconi, F., & Daldrup, T. (2018, November 15). How The Wall Street Journal is preparing its journalists to detect deepfakes. *NiemanLab.* Retrieved from https://www.niemanlab.org/2018/11/how-the-wall-street-journal-is-preparing-its-journalists-to-detect-deepfakes

Masip, P., Aran-Ramspott, S., Ruiz-Caballero, C., Suau, J., Almenar, E., & Puertas-Graell, D. (2020). News consumption and media coverage during the confinement by Covid-19: Information overload, ideological bias and sensationalism. *El Profesional de la Información*, *29*(3). https://doi.org/10.3145/epi.2020.may.12

McIntyre, K., & Gyldensted, C. (2017). Constructive journalism: An introduction and practical guide for applying positive psychology techniques to news production. *The Journal of Media Innovations*, *4*(2), 20–34. https://doi.org/https://doi.org/10.5617/jomi.v4i2.2403

McNair, B. (2017). *Fake news: Falsehood, Fabrication and Fantasy in Journalism*. London: Routledge.

Moltzau, A. (2020, January 9). What strategy does Europe have to tackle deepfakes? *Medium*. Retrieved from https://medium.com/dataseries/what-strategy-does-europe-have-to-tackle-deepfakes-fb159040f0c

Nelson, J. L., & Taneja, H. (2018). The small, disloyal fake news audience: The role of audience availability in fake news consumption. *New Media & Society*, *20*(10), 3720–3737. https://doi.org/10.1177/1461444818758715

Nirkin, Y., Keller, Y., & Hassner, T. (2019). FSGAN: Subject agnostic face swapping and reenactment. *Nirkin*. Retrieved from https://nirkin.com/fsgan

Palomo, B., & Masip, P. (2020). Journalistic reinvention for an automated and polarized scenario. In C. Toural Bran, Á. Vizoso, S. Pérez-Seijo, M. Rodríguez-Castro, & M.-C. Negreira-Rey (Eds.), *Information visualization in the era of innovative journalism* (pp. 161–170). New York, NY: Routledge.

Patadia, D. (2020, February 12). Reuters launches fact-checking initiative to identify misinformation, in partnership with Facebook. *Reuters*. Retrieved from https://www.reuters.com/article/rpb-fbfactchecking/reuters-launches-fact-checking-initiative-to-identify-misinformation-in-partnership-with-facebook-idUSKBN2061TG

Peters, M. A. (2018). The information wars, fake news and the end of globalisation. *Educational Philosophy and Theory*, *50*(13), 1161–1164. https://doi.org/10.1080/00131857.2017.1417200

Rochlin, N. (2017). Fake news: Belief in post-truth. *Library*

Hi Tech, *35*(3), 386–392. https://doi.org/10.1108/LHT-03-2017-0062

Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I., & Erviti, M.-C. (2020). Disinformation in times of pandemic: Typology of hoaxes on Covid-19. *El Profesional de la Información*, *29*(3). https://doi.org/https://doi.org/10.3145/epi.2020.may.15

Schwartz, O. (2018, November 12). You thought fake news was bad? Deep fakes are where truth goes to die. *The Guardian.* Retrieved from https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth

Shapiro, I., Brin, C., Bédard-Brûlé, I., & Mychajlowycz, K. (2013). Verification as a strategic ritual how journalists retrospectively describe processes for ensuring accuracy. *Journalism Practice*, *7*(6), 657–673. https://doi.org/10.1080/17512786.2013.765638

Shimizu, K. (2020, February 29). 2019-nCoV, fake news, and racism. *The Lancet*, *395*(10225), 685–686. https://doi.org/10.1016/S0140-6736(20)30357-3

Southern, L. (2019, July 1). 'A perfect storm': The Wall Street Journal has 21 people detecting 'deepfakes.' *Digiday*. Retrieved from https://digiday.com/media/the-wall-street-journal-has-21-people-detecting-deepfakes

Stencel, M., & Luther, J. (2020, June 22). Annual census finds nearly 300 fact-checking projects around the world. *Duke Reporters' Lab*. Retrieved from https://reporterslab.org/annual-census-finds-nearly-300-fact-checking-projects-around-the-world

Stupp, C. (2019, August 30). Fraudsters used AI to mimic CEO's voice in unusual cybercrime case. *The Wall Street Journal*. Retrieved from https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news." *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, *6*(1). https://doi.org/10.1177/2056305120903408

Vargo, C. J., Guo, L., & Amazeen, M. A. (2018). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New Media & Society*, *20*(5), 2028–2049. https://doi.org/10.1177/1461444817712086

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Waisbord, S. (2018). Truth is what happens to news. *Journalism Studies*, *19*(13), 1866–1878. https://doi.org/10.1080/1461670X.2018.1492881

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, *9*, 40–53. https://doi.org/http://doi.org/10.22215/timreview/1282

Zhu, B., Fang, H., Sui, Y., & Li, L. (2020). Deepfakes for medical video de-identification: Privacy protection and diagnostic information preservation. In *AIES 2020: Proceedings of the AAAI/ACM conference on AI, ethics, and society* (pp. 414–420). New York, NY: ACM. https://doi.org/https://doi.org/10.1145/3375627.3375849

About the Authors

Ángel Vizoso is a Researcher at Novos Medios Research Group (University of Santiago de Compostela) and Beneficiary of the Training University Lecturers´ Program (FPU) funded by the Spanish Ministry of Science, Innovation and Universities. (Government of Spain). His research lines are mainly information visualization, fact-checking and journalistic production for online media and he was Visiting Scholar at Universidade Nova de Lisboa (Portugal).

Martín Vaz-Álvarez is a Researcher at Novos Medios Research Group (University of Santiago de Compostela) and Beneficiary of the Training University Lecturers' Program (FPU) funded by the Spanish Ministry of Science, Innovation and Universities. (Government of Spain). His research lines are focused on co-creation in European public broadcasters, innovation and new Technologies.

Xosé López-García is a Professor of Journalism at University of de Compostela, PhD in History and Journalist. Coordinator of Novos Medios Research Group since 1994, whose research lines are focused on the study of digital and printed media, the analysis of the implications of technology for mediated communications, as well as the performance and funding of cultural industries or the history of communication, among others.

Article

Deepfakes on Twitter: Which Actors Control Their Spread?

Jesús Pérez Dasilva *, Koldobika Meso Ayerdi and Terese Mendiguren Galdospin

Department of Journalism II, University of the Basque Country, 48940 Leioa, Spain;
E-Mails: jesusangel.perez@ehu.eus (J.P.D.), koldo.meso@ehu.eus (K.M.A.), terese.mendiguren@ehu.eus (T.M.G.)

* Corresponding author

Abstract

The term deepfake was first used in a Reddit post in 2017 to refer to videos manipulated using artificial intelligence techniques and since then it is becoming easier to create such fake videos. A recent investigation by the cybersecurity company Deeptrace in September 2019 indicated that the number of what is known as fake videos had doubled in the last nine months and that most were pornographic videos used as revenge to harm many women. The report also highlighted the potential of this technology to be used in political campaigns such as in Gabon and Malaysia. In this sense, the phenomenon of deepfake has become a concern for governments because it poses a short-term threat not only to politics, but also for fraud or cyberbullying. The starting point of this research was Twitter's announcement of a change in its protocols to fight fake news and deepfakes. We have used the Social Network Analysis technique, with visualization as a key component, to analyze the conversation on Twitter about the deepfake phenomenon. NodeXL was used to identify main actors and the network of connections between all these accounts. In addition, the semantic networks of the tweets were analyzed to discover hidden patterns of meaning. The results show that half of the actors who function as bridges in the interactions that shape the network are journalists and media, which is a sign of the concern that this sophisticated form of manipulation generates in this collective.

Issue

This article is part of the issue "Disinformation and Democracy: Media Strategies and Audience Attitudes" edited by Pere Masip (University Ramon Llull, Spain), Bella Palomo (University of Málaga, Spain) and Guillermo López (University of Valencia, Spain).

1. Introduction

The recent upsurge in artificial intelligence (AI), along with image processing and machine learning, have made deepfake production possible. A video scarcely a minute long that featured Barack Obama spouting harsh criticism against current US President, Donald Trump, went viral in early April 2018 (Fagan, 2018). In fact, the previous US leader had said nothing, although it was his image that appeared in the video. The person who made it was actor Jordan Peele. He sought to sound the alarm on how dangerously easy it was to use new technologies to manipulate and falsify someone's identity. Deepfake

videos entail risks, and can potentially undermine truth, confuse citizens and falsify reality. With the arrival of social media, the spread of this sort of content seems to be unstoppable. Potentially, it may worsen issues related to disinformation and conspiracy theories (Hasan & Salah, 2019). They could even be weaponized to unleash national or international crises (Stover, 2018).

The firstly widely known examples of deepfakes appeared in November 2017, when a Reddit user called Deepfakes uploaded a series of videos with the faces of famous actresses, including Gal Gadot and Scarlett Johansson, over the faces of pornographic actresses (Rense, 2018). Since then, the media and the general

public have begun using the term deepfakes to refer to this sort of video made with AI, where one person's face can be confused with another's.

When the computer code used to make the fakes was shared, it sparked great interest in the Reddit community and the amount of fake content spread and increased. The fakes' initial targets were celebrities, including actors, singers and politicians. There are two possible main reasons that they were successful: accessibility and credibility (Kietzmann, Lee, McCarthy, & Kietzmann, 2020), since we tend to trust more in voices we know and in videos we see (Brucato, 2015).

2. State of Play

Manipulating photographs and videos, altering the reality of the recorded moment, came before the Internet. Different countries have carried out propaganda campaigns since World War II (Rutenberg, 2017). However, deepfakes account for a fundamental paradigm shift in how the world will operate online (Chesney & Citron, 2019).

Driven by the latest technological progress in AI and machine learning, there is a growing number of tools that make it possible for any unqualified individual to relatively easily create fake content that is increasingly more difficult to detect. In 2018, the popular face-swap program FakeApp required huge amounts of input data to create fakes (Robertson, 2018). One year later, similar applications like Zao, Doublicat and DiffSnap were more accessible and less demanding (Mehta, 2020).

This technical resource is being widely used in action films to replace actors with digital avatars in certain dangerous scenes, or even to digitally resurrect actors who have passed away (Atkin, 2019). However, when we observe their use in information systems, there are a great number of dangers and ethical challenges (Sora, 2018).

While there are those who mention the entertaining and even positive side of fakes (Kietzmann et al., 2020), some works address the use of deepfakes in online disinformation campaigns to manipulate public opinion (Riechmann, 2018). Many authors warn of the important repercussions that failure to curb their spread may pose, both to the population (Newman et al., 2015) and to democracy (Bennett & Livingston, 2018; Chadwick, Vaccari, & O'Loughlin, 2018; Rojecki & Meraz, 2016; Waisbord, 2018). There are even those who state that their fast and widespread dissemination can lead to great economic loss or national security risks (Yadlin-Segal & Oppenheim, 2020). In parallel fashion, if deepfakes contribute to increased uncertainty regarding the information they contain, another one of the risks of their use is reduced trust in the news media (Fletcher, 2018; Vaccari & Chadwick, 2020). Credibility in the news is falling around the world (Hanitzsch, Van Dalen, & Steindl, 2018) and trust in social media news is now less than news accessed through other channels (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018).

Given that there is a great deal at stake, automatic detection of deepfakes is an important problem, although difficult to undertake. Some argue that they can be fought through legislation and regulations, company policies, education and training (Westerlund, 2019). There are others who advocate for developing technology to detect deepfakes and to authenticate content and for prevention. In fact, many tools have been created to automatically detect deepfakes. To date, methods to detect these digital manipulations were based on intrinsic contradictions in the algorithm synthesis. For example, a lack of actual eye blinking (Li, Chang, & Lyu, 2018), or mismatching lip movement with speech (Korshunov & Marcel, 2018). There are systems that use a Convolutional Neural Network that extracts frame-level features that are then used to train a Recurrent Neural Network that learns to determine whether or not a video has been manipulated (Güera & Delp, 2018; Li & Lyu, 2018). There are even those who suggest tracking and monitoring the source and history of content to the origin, based on the principle that if it is reliable or prestigious, then the content can be real and authentic (Hasan & Salah, 2019).

Deepfakes promise to be one of the greatest challenges for social media platforms in 2020. Some, like Facebook and Adobe, raised policies to detect and fight deepfakes. The latest was Twitter, which announced a new policy in February to fight the impact of manipulated content (Robertson, 2020). Moreover, Google has also decided to take action to limit their reach by creating an algorithm to detect and automatically delete deepfakes uploaded to YouTube and other Google services. It also created a tool called Assemble that helps journalists to identify manipulated images (Alba, 2020).

Although deepfakes have become a topic of debate, academic research has only just recently begun addressing digital disinformation on social media (Anderson, 2018), which can be dangerous to the public sphere given the potential to create states of false opinion (Pennycook, Cannon, & Rand, 2018). In this regard, this study contributes to this debate by analyzing the conversation on Twitter about the deepfake phenomenon and which type of actors are most referenced and made viral by users, all after the news that Twitter was going to double down in its efforts to fight fake news and videos.

3. Objectives

This study's generic objective is to analyze the conversation and the structure of relationships on the net arising around the term deepfake on Twitter by means of the social network analysis technique. Deepfakes is still a relatively new and 'fluid' phenomenon in the making. This article may help people understand how different actors try to shape and 'crystalize' our understanding of the emerging issue, as well as mapping the most important actors in this debate. It contains the following specific objectives: 1) Identify the main actors and research

which ones hold a greatest advantage when controlling the spread of messages—all actors who posted messages containing the term deepfake or who were replied to or mentioned in those messages have been examined; and 2) analyze the semantic network arising around the search term deepfake and discover which content predominates in messages.

4. Hypothesis

The following hypothesis are formulated:

> H1: Politicians and the media are amongst the actors who are most referenced and made viral by third parties when speaking of deepfakes on Twitter (related to the first objective).

Politicians, because they often become the involuntary protagonists of videos which, with a humorous tone, form part of disinformation campaigns that affect their image and credibility. The news media, because they are worried about the consequences that improper use of this face-swapping technology may have for governments, companies and the general population.

> H2: The most relevant topics when users discuss deepfakes on Twitter (related to the second objective) are related to politics and concern over the growing difficulty in distinguishing between reality and fiction in the near future.

It is important to examine whether the content about deepfakes also relates to politics because "political deepfakes are an important product of the Internet's visual turn. They are at the leading edge of online, video-based disinformation and, if left unchallenged, could have profound implications for journalism, citizen competence, and the quality of democracy" (Vaccari & Chadwick, 2020, p. 2). In this sense, according to Maddocks (2020), most of the deepfakes that are spread on the Internet today are pornographic in nature, but public attention is typically focused on political deepfakes. Often simulating the image of high-profile politicians, these videos spread hoaxes and lead to political instability.

5. Methodology

Using the Social Network Analysis technique (Borgatti, Mehra, Brass, & Labianca, 2009; Freeman, 2004; Otte & Rousseau, 2002; Wasserman & Faust, 1994), this article studies the structure of networked relationships woven around the term deepfake on the social media platform Twitter. This platform was selected because it is open and creates a huge amount of interpersonal interactions that can be collected by academic researchers to study processes of how information is spread on social networks (Benson, 2016; Boyd, 2014; Brubaker & Wilson, 2018; Evans, 2016; Tolson, 2010).

To explore the properties of the net (relevance of actors and information flows), open-source software NodeXL Pro was used, one of the programs to study digital networks that is most used by the scientific community (Hansen, Shneiderman, & Smith, 2010; Ricaurte & Ramos-Vidal, 2015; Smith et al., 2010). This tool was used in different works of research, such as the one analyzing connections between politicians and journalists in Holland (Verweij, 2012), the use of hashtags and trending topics (Dossis, Amanatidis, & Mylona, 2015; Wukich & Steinberg, 2013), news-spreading processes (Ahmed & Lugovic, 2019), the spread of hoaxes regarding the coronavirus (Pérez-Dasilva, Meso-Ayerdi, & Mendiguren-Galdospín, 2020), and more.

The software captured a network of 15,885 actors who posted messages containing the term deepfake or who were replied to or mentioned in those messages. The sample was taken from a data set limited to a maximum of 18,000 tweets (formal limits of the NodeXL software sample universe). The database was obtained through Twitter's streaming API February 28th, 2020, at 09:41 UTC. The reason for this choice is that on February 5th, the platform created by Jack Dorsey announced a new policy to fight content manipulation like fake news and fake videos (Robertson, 2020). The collected tweets were posted between February 7, 2018, at 11:17 UTC and February 28, 2020, at 09:28 UTC. Users were grouped by hierarchical conglomerates (or cluster analysis; Kaleel & Abhari, 2015; Paolillo, 2008), using the algorithm by Clauset, Newman, and Moore (2004). To visualize the network, Harel and Koren's (2000) multiscale design algorithm was used, which facilitates identification of actors and their links. Analysis was based on directed and weighted edges. The weight reflected the number of times that actors posted messages containing the term deepfake or who were replied to or mentioned in those messages.

To analyze the role held by actors and the relationships that occur between them on the network revolving around the term deepfake, two of the most common centrality indicators in the SNA were used: in-degree and the degree of betweenness. The in-degree is the number of interactions an actor has received from other users forming the structure (Aguilar-Gallegos et al., 2016; Fernández, 2019). Actors with the highest numbers were the most-referenced and made most-viral, so their content is the most influential. On the other hand, the degree of betweenness is the capacity to control spreading of a message (Gibbs & McKendrick, 2015; Hansen et al., 2010). Users with the highest numbers acted as bridges over which relevant information flowed, and they contributed to spreading or blocking it for other parts of the network. The color and size of the nodes show the most relevant accounts, and the strength of the bonds between them was shown with the intensity of the lines joining them.

To study the semantic network created around the search term deepfake, words such as conjunctions and

prepositions, which are not relevant, were eliminated. Next, a data-mining strategy based on word-matching was applied (on nouns, verbs, adverbs and adjectives) to identify the strongest connections (Seo, Kim, Kim, Kim, & Kim, 2019) and its presence in each message was studied from a relational perspective. These data were interpreted as non-directed graphs.

6. Results

Structure indicators or cohesion measurements, such as density or reciprocity, that analyze the complete network's properties were 0.006393862 for the ratio of reciprocal vertex pairs and 5.981029 for the average geodesic distance. The first data indicates that 6 of every 100 users held mutual communication during the period of study, and the second that the actor was located at almost six steps on average from any other in the analyzed structure. Moreover, density or cohesion was 6.86199%. These data indicate that this is a dense network, where nodes are not very far from one another and with a high speed of information transmission.

Centrality indicators, which show the position a node or actor holds on the network, also bore interesting results. Figure 1 shows the existence of different commu-

nities. Of note are eight large-sized clusters (light-blue, orange, red, green, dark green, yellow, light green and maroon), followed by ten moderately-sized groupings.

According to the nodes' in-degree, of the first 20, a group of 10 accounts from India stands out, followed by eight profiles from the US or headquartered in the US (such as Mashable or Elon Musk), plus another one from Brazil and another from France (Figures 2 and 3). The one with the highest value is @techreview, MIT's magazine. This actor's contents are amongst the most-referenced and most virally spread by third parties. One of the messages reports on the purchase by Fintech Square, headed by Twitter CEO Jack Dorsey, of the AI research company Dessa, a company known for its deepfake-detection software. Its technology became known thanks to the deepfake on Joe Rogan, a mixed martial arts commentator and one of the most popular podcasters in the world on May 17, 2019 (Vincent, 2019).

Moreover, another post under the same profile is of note, which went widely viral and reported on the use of a deepfake to win voters, used by Indian politician Manoj Tiwari, president of the Bharatiya Janata party (BJP). According to MIT's magazine, this was the first time in the world that a political party used a deepfake for an electoral campaign. The controversy arose when
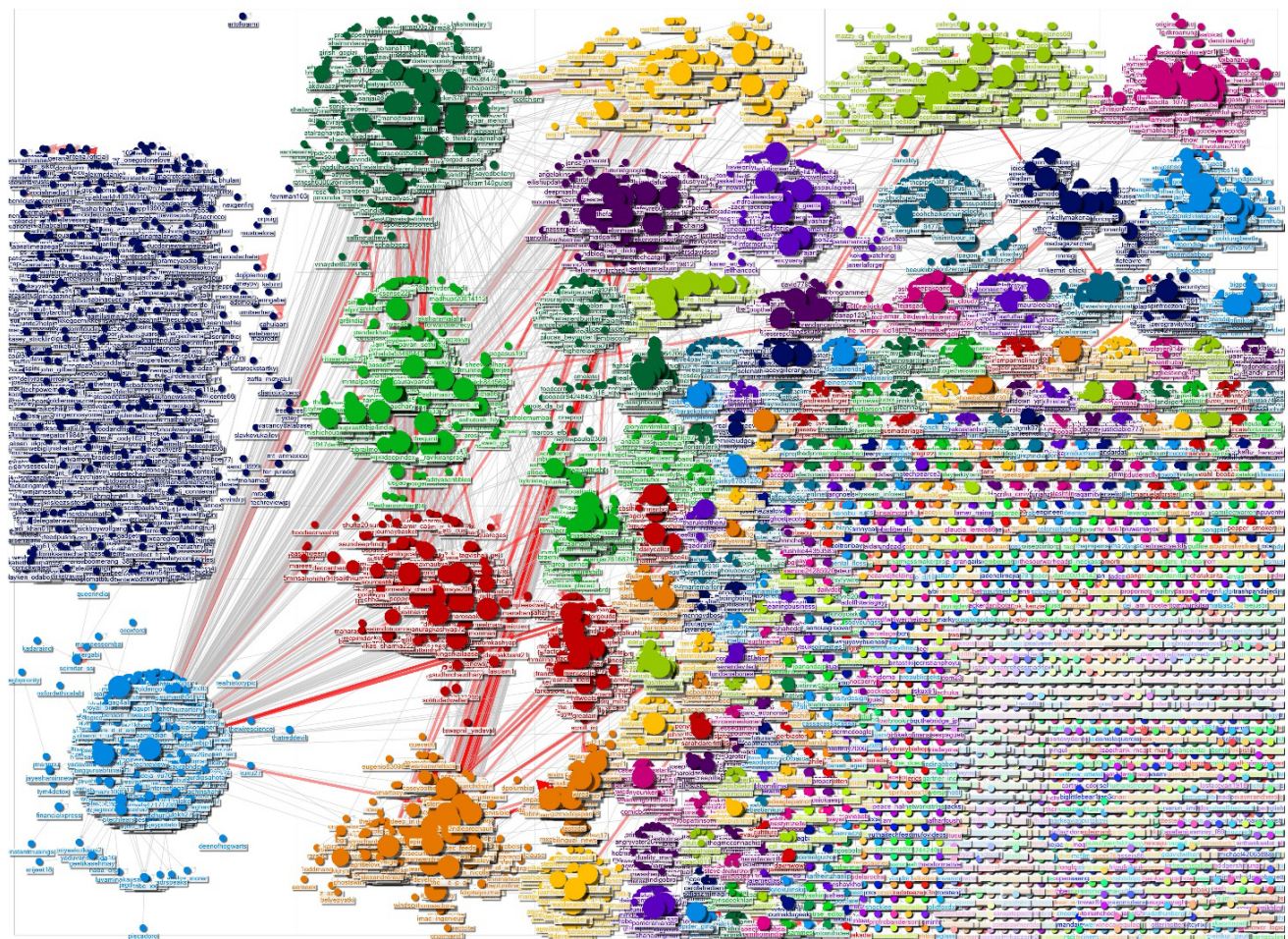

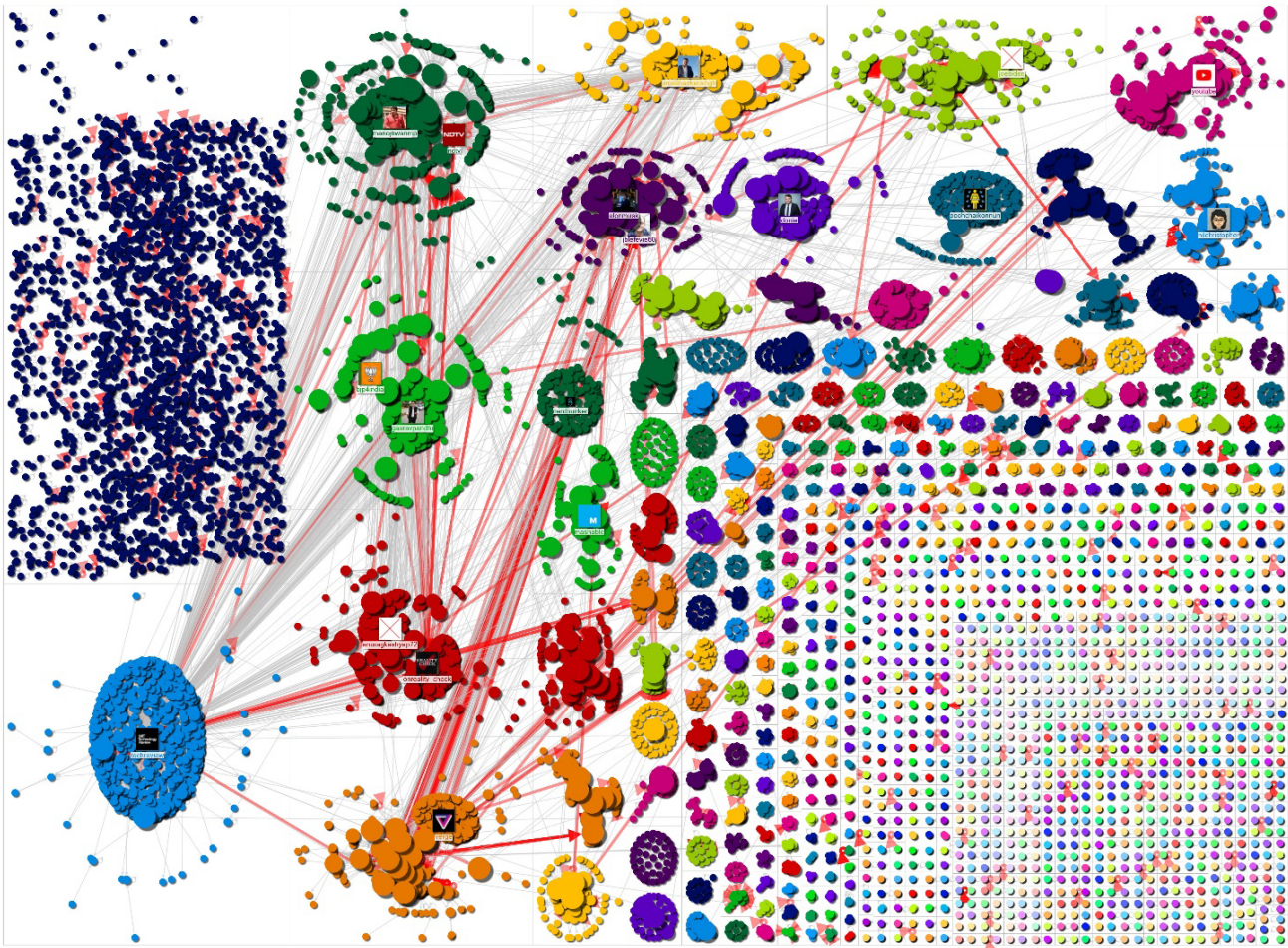**Figure 1.** Illustration of the network around the term deepfake.

Figure 2. Illustration of the 20 actors most-referenced by users.

Tiwari manipulated one of his electoral videos with deep-fake technology to simulate that he was speaking a Hindi dialect, and thus reach millions of voters that would have been unreachable otherwise, since they only speak this dialect. According to the party itself, they hired the company Ideaz Factory to create deepfakes to reach voters in the 22 different languages and 1,600 dialects used in India.

An examination of the content around the 20 most-referenced actors detects three aspects. On the one hand, the presence of politics when speaking of deepfakes must be mentioned. Throughout the period of study, ten of the actors most-referenced by third parties are related to Manoj Tiwari (the magazine @techreview, activist @GuaravPandhi, @ManojTiwariMP, journalist @UmashankarSingh, TV program @OnReality_Check, politician @amitmalviya, film director @anuragkashyap72, the party @BJP4India, journalist @NilChristopher and television channel @ndtv). In addition to this group is US politician Joe Biden, another node to which a huge number of users go in an attempt to generate a direct link with him. Former vice president Biden became a protagonist based on a video related to the Democratic Primary debate of 2020. The original recording of the debate in Nevada

was edited by Mike Bloomberg, one of the participants, to improve his image since he did not appear in a flattering light. The billionaire modified the audio and order of scenes in the video and included grasshopper sounds when his adversaries responded. The video obtained 4,2 million views. Shortly thereafter, Twitter announced that it would sanction the video for violating its new media manipulation policy.

Another topic revolves around film. Six of the most viralized actors are related to two manipulations of popular films using AI to swap the faces of movie stars in one or several iconic scenes. One of the deepfakes, with almost half a million views, fakes a moment from the well-known science fiction series *Star Trek*. The video was made by The Fakening, a famous YouTube channel owned by programmer Paul Shales, devoted to creating fake videos with AI. This face-swapping technology places Elon Musk and Jeff Bezos in the role of the series' actors, and is associated with the media profiles @verge, owner of Tesla @elonmusk, owner of Amazon @JeffBezos and French influencer @jblefevre60. The other manipulated film, the work of YouTuber EZRyderX47 with almost nine million views, is *Back to the Future*. Thanks to its quality, it is referenced by commentator @PoohChaikonNun, entertain-
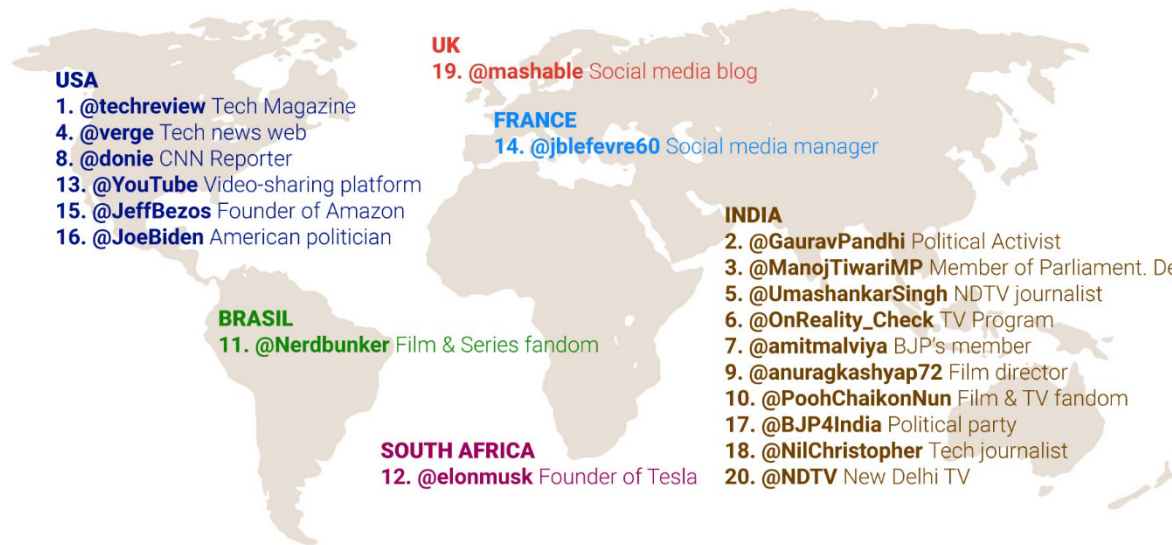
Figure 3. In-degree: Actors most-referenced and viralized by other users.

ment website @Nerdbunker and another media outlet @mashable.

Lastly, the third topic area has to do with the very technology used to produce deepfakes. In this area, five profiles are of note whose contents are amongst the most-referenced and made most-viral by third parties. As mentioned previously, one of them is @techreview, MIT's magazine, reporting on Jack Dorsey's purchase of the deepfake software company Dessa. Another is *CNN* journalist Donie O'Sullivan (@donie), author of the highly viral news piece on the dangers associated with improper use of this technology. This information is also related to @NilChristopher, another one of the actors with a high in-degree. Within this scope, we also see French influencer @jblefevre60 holding one of the top positions. He is mentioned in a very widespread tweet by global influencer Spiros Margaris, explaining how video faking technology works. The fifth actor is @Youtube, since there are plenty of videos about the dangers of deepfakes that end with the phrase "via @Youtube" to indicate the platform from whence the content was obtained.

Regarding the degree of betweenness (Figures 4 and 5), we observe that 13 of the actors mentioned in the section above appear again in the 20 top positions. They have the highest values, which means that these nodes are intermediaries through which relevant information related to deepfakes is spread. These users are the ones who contribute the most to spreading or blocking messages to other people that give shape to the structure. In this regard, it is interesting to highlight that, of the 20 actors with the most favorable positions, there are six media outlets (@techreview, @verge, @mashable, @OnReality_Check, @CNN, @YouTube) and three journalists (@donie, @UmashankarSingh, @NilChristopher) who act as bridges in interactions giving shape to the

network. In this regard, also in analyzing the role played by certain actors in configuring the structure, we must make special mention of @thefakening (15th position), because this YouTube channel creates a good portion of the most popular deepfakes spread amongst social media. Barring exceptions, many of his fake videos garner no more than 25,000 views, but the fake with Elon Musk and Jeff Bezos as actors on *Star Trek* reached almost half a million views and became his most viral video.

Regarding semantic analysis of the network, the most relevant conversation threads revolve around the video of Indian politician Manoj Tiwari, highlighting that this is the first time in the world that a political party used deepfake technology to conduct an electoral campaign (Figure 6). Moreover, it is rated as "dangerous and illegal" (Pandhi, 2020). The second most-significant association has to do with the deepfake of *Star Trek*, with the owners of Amazon and Tesla. In third position, we find references to the *Back to the Future* video.

7. Conclusions

This study shows the network woven around the term deepfake after Twitter's announcement that it was tightening its protocols to fight fake news and videos. The data indicate that this is a dense network with high connectivity where information on deepfakes quickly spreads. Although reports state that 96% of these fakes are non-consensual pornography (Patrini, 2019), this piece of research observes that in the microblogging network, the most important topics are not related to pornographic content. The nodes with the most favorable positions in the structure converse on fake videos related to politicians (H1). This coincides with studies such as those by Maddocks (2020) which explain that, although most of the deepfakes that spread over the

Figure 4. Illustration of the main actors control the flow of information on the network.

CANADA
15. @thefakening Deep fake videos

UK
5. @mashable Social media blog

USA
1. @techreview Tech Magazine
2. @verge Tech news web
3. @donie CNN Reporter
16. @THollandNews Tom Holland fandom
19. @YouTube Video-sharing platform
20. @CNN TV Channel

THAILAND
18. @NKengkun T. Holland fandom

INDIA
4. @GauravPandhi Political Activist
6. @ManojTiwariMP Member of Parliament. Delhi
8. @UmashankarSingh NDTV journalist
10. @OnReality_Check TV Program
12. @NilChristopher Tech journalist
13. @BJP4India Political party
17. @PoohChaikonNun Film & TV fandom

SOUTH AFRICA
14. @elonmusk Founder of Tesla

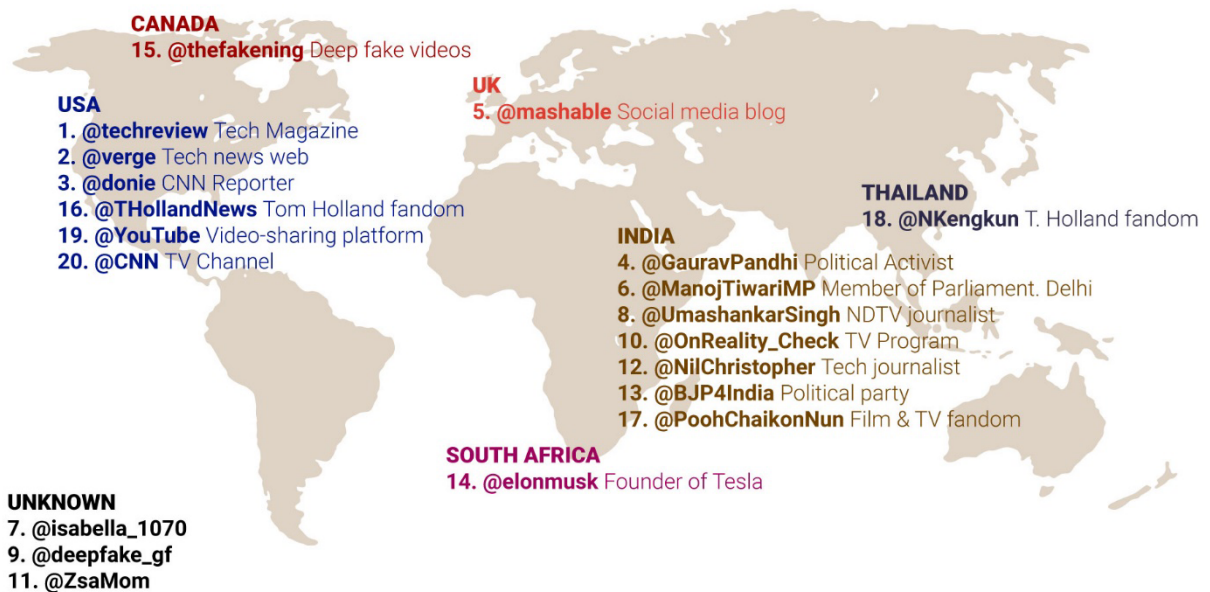**UNKNOWN**
7. @isabella_1070
9. @deepfake_gf
11. @ZsaMom

Figure 5. Betweenness degree: Actors who act as bridges in the interactions that shape the network.
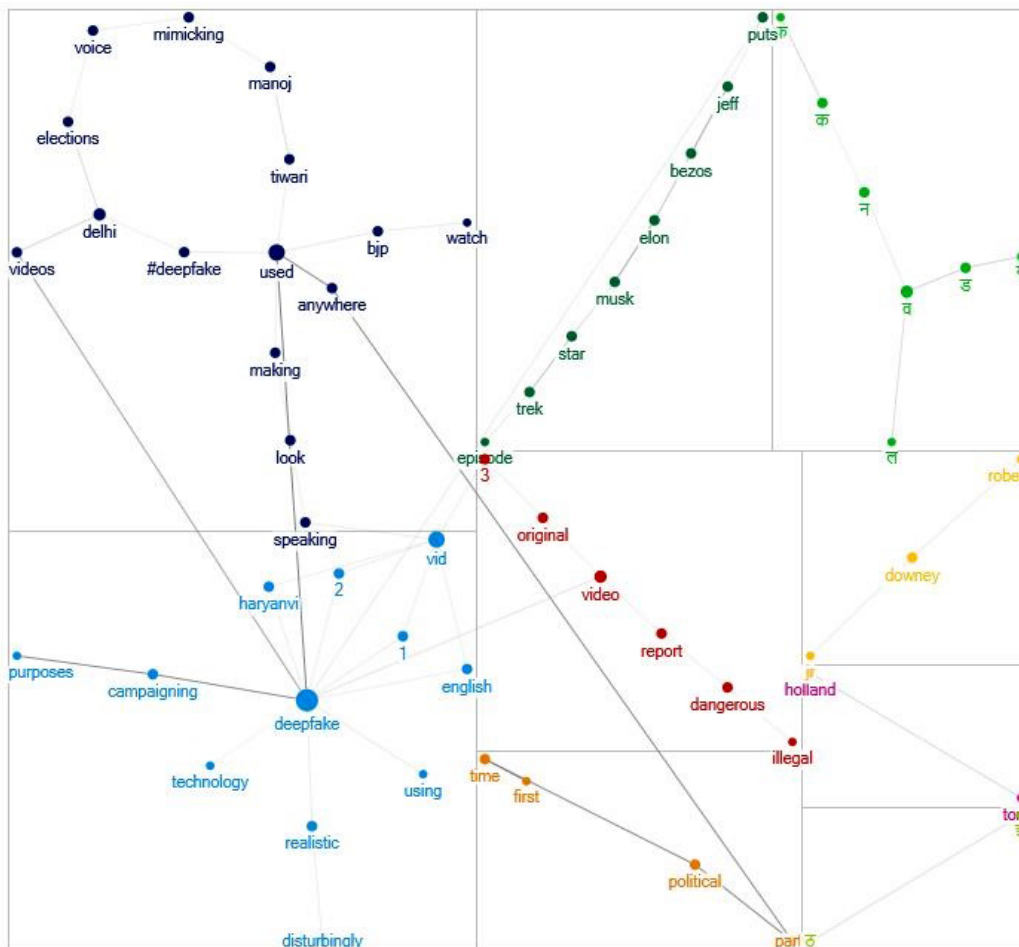
Figure 6. Illustration of the most relevant conversation topics regarding deepfakes.

Internet are pornographic in nature, public attention is focused above all on political deepfakes because of their ability to generate political instability. In contrast, other authors such as Westerlund (2019) conclude that the reason many counterfeits focus on celebrities and politicians is basically because they are public figures who have a large number of free videos and photos on the Internet and it is an easy way to train an AI deepfake system. According to these two recent studies, this is only the beginning. There are going to be more and more deepfakes that use AI to spread fake political videos tailored to the preferences of social media users.

Secondly, the nodes that have the greatest structural advantage in the network also refer to satirical videos of famous films where face-swapping technology is used on the actors in one or several iconic scenes from said films.

The result of this is that if we consider the network according to the in-degree, the most-referenced and viralized users are celebrities (politicians, businessmen or businesswomen, singers, athletes and more) that see how they become the target of manipulations. Thirdly, and in relation to the above two points, we also observe concern, especially with news media, for the consequences that improper use of this AI technology may have for citizens, companies and governments (H2).

In all their tweets, analyzed in this study, news media talk about the potential danger of this technology. This research coincides with recent studies such as those by Yadlin-Segal and Oppenheim (2020, p. 1) because it "shows how journalists frame deepfakes as a destabilizing platform that undermines a shared sense of social and political reality." On the other hand, if we consider the network in terms of degree of betweenness, we observe that half the actors with the greatest capacity to control the spread of messages on deepfakes are also journalists or news media (H1). In this study, although most of the videos are entertaining and easy to spot, these professionals are clearly concerned and have the responsibility of discrediting these fake videos and avoiding the manipulation of public opinion: "Authentication of video is especially important to news media companies who have to determine authenticity of a video spreading in a trustless environment, in which details of the video's creator, origin, and distribution may be hard to trace" (Westerlund, 2019, p. 46)

In this work, attention has been focused on the conversation about deepfakes, about the way users talk about the subject, and it has been shown who are the most referenced actors and whose contents are the most viralized by users. As mentioned above in the arti-

COGITATIO

cle, deepfakes is still a relatively new phenomenon and the purpose of this manuscript has been to help understand how different actors try to shape and crystalize our understanding of the emerging issue, as well as mapping the most important actors in this debate. Current research could be extended to include the study of the spread of deepfakes. Work is continuing in this area in order to overcome limitations connected to this study.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

References

Aguilar-Gallegos, N., Martínez-González, E. G., Aguilar-Ávila, J., Santoyo-Cortés, H., Muñoz-Rodríguez, M., & García-Sánchez, E. I. (2016). Análisis de redes sociales para catalizar la innovación agrícola: De los vínculos directos a la integración y radialidad [Social network analysis for catalysing agricultural innovation: From direct ties to integration and radiality]. *Estudios Gerenciales*, *32*(140), 197–207. https://doi.org/10.1016/j.estger.2016.06.006

Ahmed, W., & Lugovic, S. (2019). Social media analytics: Analysis and visualisation of news diffusion using NodeXL. *Online Information Review*, *43*(1), 149–160. https://doi.org/10.1108/OIR-03-2018-0093

Alba, D. (2020, February 4). Tool to help journalists spot doctored images is unveiled by jigsaw. *The New York Times*. Retrieved from https://www.nytimes.com/2020/02/04/technology/jigsaw-doctored-images-disinformation.html

Anderson, K. E. (2018). Getting acquainted with social networks and apps: Combating fake news on social media. *Library Hi Tech News*, *35*(3), 1–6. https://doi.org/10.1108/LHTN-02-2018-0010

Atkin, M. (2019, March 27). Human assets. *Medium*. Retrieved from https://immerse.news/human-assets-624f3066c2ce

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Benson, P. (2016). *The discourse of YouTube: Multimodal text in a global context*. London: Routledge. https://doi.org/10.4324/9781315646473

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892–895. https://doi.org/DOI:10.1126/science.1165821

Boyd, M. S. (2014). (New) participatory framework on YouTube? Commenter interaction in US political speeches. *Journal of Pragmatics*, *72*, 46–58. https://doi.org/10.1016/j.pragma.2014.03.002

Brubaker, P. J., & Wilson, C. (2018). Let's give them something to talk about: Global brands' use of visual content to drive engagement and build relationships. *Public Relations Review*, *44*(3), 342–352. https://doi.org/10.1016/j.pubrev.2018.04.010

Brucato, B. (2015). Policing made visible: Mobile technologies and the importance of point of view. *Surveillance & Society*, *13*(3/4), 455–473. https://doi.org/10.24908/ss.v13i3/4.5421

Chadwick, A., Vaccari, C., & O'Loughlin, B. (2018). Do tabloids poison the well of social media? Explaining democratically dysfunctional news sharing. *New Media & Society*, *20*(11), 4255–4274. https://doi.org/10.1177/1461444818769689

Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs*. Retrieved from https://www.questia.com/magazine/1P4-2161593888/deepfakes-and-the-new-disinformation-war-the-coming

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6). https://doi.org/10.1103/PhysRevE.70.066111

Dossis, M., Amanatidis, D., & Mylona, I. (2015, November 9). Mining Twitter data: Case studies with trending hashtags. In M. Mokryš & Š. Badura (Eds.), *Proceedings in ARSA-advanced research in scientific areas* (pp. 242–246). Zilina: EDIS Publishing Institution of the University of Žilina.

Evans, M. (2016). Information dissemination in new media: YouTube and the Israeli-Palestinian conflict. *Media, War & Conflict*, *9*(3), 325–343. https://doi.org/10.1177/1750635216643113

Fagan, K. (2018, April 17). A viral video that appeared to show Obama calling Trump a "dips" shows a disturbing new trend called "deepfakes." *Business Insider*. Retrieved from https://www.businessinsider.com/obama-deepfake-video-insulting-trumpp.2018--4

Fernández, D. (2019, August 19). Análisis de grafos en redes sociales: Medidas de centralidad [Analysis of networks in social networks: Measures of centrality]. *Datahack*. Retrieved from https://www.datahack.es/grafos-redes-sociales-centralidad

Fletcher, J. (2018). Deepfakes, artificial intelligence, and some kind of dystopia: The new faces of online post-fact performance. *Theatre Journal*, *70*(4), 455–471.

https://doi.org/10.1353/tj.2018.0097

Freeman, L. (2004). *The development of social network analysis: A study in the sociology of science*. Vancouver: Empirical Press.

Gibbs, W. J., & McKendrick, J. (Eds.). (2015). *Contemporary research methods and data analytics in the news industry*. Pennsylvania, PA: IGI Global.

Güera, D., & Delp, E. J. (2018). Deepfake video detection using recurrent neural networks. In *2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS)* (pp. 1–6). Los Alamitos, CA: IEEE Computer Society.

Hanitzsch, T., Van Dalen, A., & Steindl, N. (2018). Caught in the nexus: A comparative and longitudinal analysis of public trust in the press. *The International Journal of Press/Politics*, 23(1), 3–23. https://doi.org/10.1177/1940161217740695

Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Burlington, MA: Morgan Kaufmann.

Harel, D., & Koren, Y. (2000). A fast multi-scale method for drawing large graphs. In S. Levialdi, V. Di Gesu, & L. Tarantino (Eds.), *Proceedings of the working conference on advanced visual interfaces: AVI '00* (pp. 282–285). New York, NY: ACM Press. https://doi.org/10.1145/345513.345353

Hasan, H. R., & Salah, K. (2019). Combating deepfake videos using blockchain and smart contracts. *IEEE Access*, 7, 41596–41606. https://doi.org/10.1109/ACCESS.2019.2905689

Kaleel, S. B., & Abhari, A. (2015). Cluster-discovery of Twitter messages for event detection and trending. *Journal of Computational Science*, 6, 47–57. https://doi.org/10.1016/j.jocs.2014.11.004

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006

Korshunov, P., & Marcel, S. (2018). Speaker inconsistency detection in tampered video. In *2018 26th European signal processing conference (EUSIPCO)* (pp. 2389–2393). Los Alamitos, CA: IEEE Computer Society.

Li, Y., Chang, M.-C., & Lyu, S. (2018). *In ictu oculi: Exposing AI generated fake face videos by detecting eye blinking*. ArXiv. http://arxiv.org/abs/1806.02877

Li, Y., & Lyu, S. (2018). *Exposing deepfake videos by detecting face warping artifacts*. ArXiv. https://arxiv.org/abs/1811.00656

Maddocks, S. (2020). A deepfake porn plot intended to silence me: Exploring continuities between pornographic and 'political' deep fakes. *Porn Studies*. https://doi.org/10.1080/23268743.2020.1757499

Mehta, I. (2020, January 13). New deepfake app pastes your face onto GIFs in seconds. *The Next Web*. Retrieved from https://thenextweb.com/artificial-intelligence/2020/01/13/new-deepfake-app-pastes-your-face-onto-gifs-in-seconds

Newman, E. J., Garry, M., Unkelbach, C., Bernstein, D. M., Lindsay, D. S., & Nash, R. A. (2015). Truthiness and falsiness of trivia claims depend on judgmental contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(5), 1337–1348. https://doi.org/10.1037/xlm0000099

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., & Nielsen, R. K. (2018). *Reuters Institute digital news report 2018*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf

Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science*, 28(6), 441–453.

Pandhi, G. [GauravPandhi]. (2020, February 19). Watch how BJP used #deepfake in Delhi elections where a voice mimicking Manoj Tiwari is used, making it look like he is speaking. Report: https://t.co/8zcKwnSxbL This is dangerous, should be illegal! https://t.co/WEXb0zaXdl [Tweet]. Retrieved from https://twitter.com/gauravpandhi/status/1229966733370257411

Paolillo, J. C. (2008). Structure and network in the YouTube core. *Proceedings of the 41st annual Hawaii international conference on system sciences (HICSS 2008)* (pp. 156–156). Los Alamitos, CA: IEEE Computer Society. https://doi.org/10.1109/HICSS.2008.415

Patrini, G. (2019, October 7). Mapping the deepfake landscape. *Deeptrace*. Retrieved from https://deeptracelabs.com/mapping-the-deepfake-landscape

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880. https://doi.org/10.1037/xge0000465

Pérez-Dasilva, J.-A., Meso-Ayerdi, K., & Mendiguren-Galdospín, T. (2020). Fake news y coronavirus: Detección de los principales actores y tendencias a través del análisis de las conversaciones en Twitter [Fake news and coronavirus: Detecting key players and trends through analysis of Twitter conversations]. *El Profesional de la Información*, 29(3). https://doi.org/10.3145/epi.2020.may.08

Rense, S. (2018, February 12). What are 'deepfakes,' and why are Pornhub and Reddit banning them? *Esquire*. Retrieved from https://www.esquire.com/lifestyle/sex/a17043863/what-are-deepfakes-celebrity-porn

Ricaurte, P., & Ramos-Vidal, I. (2015). Investigación en redes sociales digitales: Consideraciones metodológicas desde el paradigma estructural [Research on digital social networks: methodological considerations from the structural paradigm]. *Virtualis: Revista de Cultura Digital*, 6(11), 165–194.

Riechmann, D. (2018). I never said that! High-tech deception of 'deepfake' videos. *WCJB*. Retrieved from https://www.wcjb.com/content/news/I-never-

said-that-High-tech-deception-of-deepfake-videos-487147011.html

Robertson, A. (2018, February 11). I'm using AI to face-swap Elon Musk and Jeff Bezos, and I'm really bad at it. *The Verge*. Retrieved from https://www.theverge.com/2018/2/11/16992986/fakeapp-deepfakes-ai-face-swapping

Robertson, A. (2020, February 4). Twitter will ban 'deceptive' faked media that could cause 'serious harm.' *The Verge*. Retrieved from https://www.theverge.com/2020/2/4/21122661/twitter-deepfake-manipulated-media-policy-rollout-date

Rojecki, A., & Meraz, S. (2016). Rumors and factitious informational blends: The role of the web in speculative politics. *New Media & Society*, *18*(1), 25–43. https://doi.org/10.1177/1461444814535724

Rutenberg, J. (2017, September 13). RT, Sputnik and Russia's new theory of war. *The New York Times*. Retrieved from https://www.nytimes.com/2017/09/13/magazine/rt-sputnik-and-russias-new-theory-of-war.html

Seo, S., Kim, J.-K., Kim, S.-I., Kim, J., & Kim, J. (2019). Semantic hashtag relation classification using co-occurrence word information. *Wireless Personal Communications*, *107*(3), 1355–1365. https://doi.org/10.1007/s11277-018-5745-y

Smith, M., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., & Dunne, C. (2010). *NodeXL: A free and open network overview, discovery and exploration add-in for Excel 2007/2010*. Redwood City, CA: Social Media Research Foundation.

Sora, C. (2018). El futuro digital de los hechos [The digital future of facts]. *Hipertext.net: Revista Académica sobre Documentación Digital y Comunicación Interactiva*, *0*(17), 1–10. https://doi.org/10.31009/hipertext.net.2018.i17.01

Stover, D. (2018). Garlin Gilchrist: Fighting fake news and the information apocalypse. *Bulletin of the Atomic Scientists*, *74*(4), 283–288. https://doi.org/10.1080/00963402.2018.1486618

Tolson, A. (2010). A new authenticity? Communicative practices on YouTube. *Critical Discourse Studies*, *7*(4), 277–289. https://doi.org/10.1080/17405904.2010.511834

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, *6*(1). https://doi.org/10.1177/2056305120903408

Verweij, P. (2012). Twitter links between politicians and journalists. *Journalism Practice*, *6*(5/6), 680–691. https://doi.org/10.1080/17512786.2012.667272

Vincent, J. (2019, May 17). This AI-generated Joe Rogan fake has to be heard to be believed. *The Verge*. Retrieved from https://www.theverge.com/2019/5/17/18629024/joe-rogan-ai-fake-voice-clone-deepfake-dessa

Waisbord, S. (2018). Truth is what happens to news: On journalism, fake news, and post-truth. *Journalism Studies*, *19*(13), 1866–1878. https://doi.org/10.1080/1461670X.2018.1492881

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge: Cambridge University Press.

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, *9*(11), 39–52. https://doi.org/10.22215/timreview/1282

Wukich, C., & Steinberg, A. (2013). Nonprofit and public sector participation in self-organizing information networks: Twitter hashtag and trending topic use during disasters: Self-organizing information networks. *Risk, Hazards & Crisis in Public Policy*, *4*(2), 83–109. https://doi.org/10.1002/rhc3.12036

Yadlin-Segal, A., & Oppenheim, Y. (2020). Whose dystopia is it anyway? Deepfakes and social media regulation. *Convergence: The International Journal of Research into New Media Technologies*. https://doi.org/10.1177/1354856520923963

About the Authors

Jesús Pérez Dasilva is Professor of Online Journalism at the University of the Basque Country. His research focuses on digital journalism and social networks. He has published widely on these topics in peer-reviewed international journals included in listings as JCR or Scopus. Nowadays is member of the research project "News, Networks and Users in the Hybrid Media System: News Creation and Sharing in Online Media," funded by Ministry of Science, Innovation and Universities of Spain.

Koldobika Meso Ayerdi is a Senior Lecturer at the Department of Journalism II of the University of the Basque Country. He has researched online journalism and blogging and organizes the annual International Conference on Online Journalism and Web 2.0 at the University of the Basque Country in Bilbao, Spain.

Terese Mendiguren Galdospin is a Lecturer at the Department of Journalism II at the University of the Basque Country. Author of works on cyberjournalism, she investigates trends in contemporary journalism such as citizen journalism or the use of social networks in the dissemination of information. She is Member of the consolidated investigation group Gureiker. In her professional role, she has worked as an Editor and Program Coordinator for the Basque television channel (EITB) and Bilbovision Channel.

Article

Post-Truth as a Mutation of Epistemology in Journalism

Pablo Capilla

Blanquerna Institute of Research in Communication and International Relations, Ramon Llull University, 08001 Barcelona, Spain; E-Mail: pablocg@blanquerna.url.edu

Abstract
In recent years, many authors have observed that something is happening to the truth, pointing out that, particularly in politics and social communication, there are signs that the idea of truth is losing consideration in media discourse. This is no minor issue: Truth, understood as the criterion for the justification of knowledge, is the essential foundation of enlightened rationality. The aim of this article, based on prior research on social communication (especially as regards journalism), is to elucidate an explanation of this phenomenon, known as 'post-truth.' Because it is an epistemological question, the three main variables of the problem (reality, subject and truth) have been analysed by taking into account the manner in which digital social communication is transforming our perception of reality. By way of a conclusion, we propose that (a) the ontological complexity of reality as explained by the news media has accentuated the loss of confidence in journalism as a truth-teller, and that (b) truth is being replaced by sincerity, as an epistemological value, in people's understanding of the news. The result, using Foucault's concept of Regime of Truth, suggests a deep change in the global framework of political, economic, social and cultural relations, of which post-truth is a symptom.

Issue
This article is part of the issue "Disinformation and Democracy: Media Strategies and Audience Attitudes" edited by Pere Masip (Ramon Llull University, Spain), Bella Palomo (University of Málaga, Spain) and Guillermo López (University of Valencia, Spain).

1. Introduction

The idea of 'post-truth' may evolve into a 'zombie concept' if its meaning is not determined and if it fails to show some kind of capacity to explain our reality. Since its popularisation in 2016, following the victory of Donald Trump in the US and Brexit in the UK, post-truth has come to form part of public and academic discourse. However, in social communication, it has only been used to designate a vague series of phenomena: fake news, disinformation, loss of trust in the media, the 'emotional turn' caused by the influence of social media (Wahl-Jorgensen, 2016) and, above all, the rise of populism, aided by political communication practices. The epistemological implications of post-truth have hardly been analysed, despite the fact that it is clearly an epistemological concept, since it deals with truth, that is, with the validation of statements about reality.

Philosophy, for its part, has shown little interest in post-truth. Susan Haack (2019) admits that perhaps concern for truth is on the decline, but that does not imply that the idea of truth is in crisis. For Lorna Finlayson, post-truth is nothing more than an 'act of saying':

> What, finally, is being done with the word 'post-truth' when it is used? As with the sort of political speech it is used to talk about, talk about post-truth appears to make little sense when taken at face value: It is either totally banal…or it is both wildly audacious and philosophically confused. (Finlayson, 2019, p. 78)

These authors argue that the idea of post-truth adds nothing new, only more confusion, because it is an imprecise and politically charged term usually reserved for discrediting opponents.

However, research in various academic fields indicates that in recent decades there has been a deep transformation not only of the material means (economic and technological) of our societies, but also of the social structures and the forms of subjectivity in which these material means act (Boler & Davis, 2018), and that this could be affecting the idea of truth. Part of this approach revolves around 'neoliberalism,' understood not only as an economic theory and practice, but as an alternative to enlightened rationality that affects the way in which people perceive themselves, others and reality. Wendy Brown defines neoliberalism as "an order of normative reason that, when it becomes ascendant, takes shape as a governing rationality extending a specific formulation of economic values, practices and metrics to every dimension of human life" (Brown, 2015, p. 30).

Researchers in the fields of general sociology (Gane, 2014), cognitive sociology (Leyva, 2020), social psychology (Gjorgjioska & Tomici, 2019) and education (Goldstein, Macrine, & Chesky, 2011) have also pointed in the same direction.

This is how the central question of this article arises: Do post-truth and neoliberal subjectivity refer to similar things? Is this means of naming phenomena associated with neoliberalism? Does post-truth specifically refer to an epistemological mutation caused by these phenomena? Obviously, these are not new questions: authors such as Calin Cotoi (2011), Barbara Biesecker (2018) and Sergei Prozorov (2019) have linked post-truth to neoliberalism under the Foucaultian concept of 'Regimes of Truth' (RoT), which defines the general framework in which the relationship between truth, power and subjectivity is established. It is not the aim of this article to engage in a debate on the notion of RoT, nor whether Foucault is at the philosophical origin of post-truth (McIntyre, 2018), but to propose an explanation for the change that may be occurring in the perception of the truth within the framework of this neoliberal rationality. For this, we turn to previous studies on the issue conducted within the field of journalism, since journalism is a gnoseological activity.

Thus, this article seeks (a) to problematise journalistic ontology, in line with other authors, as the first step to addressing the issue of the validation of statements about reality: We start with ontology, because, without knowing the facts, it makes no sense to consider the possibility of knowledge, and journalism has tended to uncritically accept that facts simply exist. The hypothesis is that, if the news media were to spread different types of reality, it would be impossible to establish a single epistemological justification, and doubts might even be cast about the very idea of verifying facts. Next (b), we will analyse how these problematic journalistic facts are being validated in a digital media context: We place this question in the conceptual framework in which post-truth and neoliberal rationality converge, using the Foucaultian concept of RoT because it enables us to integrate the subjective dimension into the gnose-

ological process, and because it explains how epistemology is determined by the neoliberal hegemony.

2. Theoretical Framework

Discussion surrounding the concept of truth falls within the ambit of knowledge, i.e., what do we know, how and to what certainty, and revolves around three factors: (a) reality (ontology)—what we want to know, which entails the implicit acceptance that something outside the subject exists; (b) the subject—the individual who makes statements based on their perception of reality—these statements emerge as a conviction that what they are saying is what they have perceived (sincerity) and are shared with other subjects with a view to instilling in them the same conviction; and (c) truth as a shared criterion for justifying statements about reality. This is what gives others a reason to accept the statement and subsequently hold it as true.

According to Bernard Williams, knowledge is based on the values of 'sincerity' (people believe what they say) and 'accuracy' (what people say is caused by contact with reality and may be checked against reality). Sincerity pertains to the subject, to their beliefs, and entails a willingness to ensure that our statements about reality express what we really believe. Sincerity therefore also has a social dimension, since it is assumed that whoever communicates something wants others to share their idea:

> The connections between belief and truth explain why, in the case of sincere assertion, a speaker's intention to inform the hearer about the truth, and to inform him about the speaker's beliefs, fit naturally together—they are two sides of the same intention. (Williams, 2002, p. 75)

Accuracy refers to the methods used to justify the statements, drawing a distinction between methods which prove more reliable than others when it comes to reflecting reality. But accuracy requires sincerity, because a person can lie using exact data: What prevails in the lie is the issuer's willingness to hide what they really think in order to manipulate the reality they present to others, effectively trapping them by their will. According to this idea of knowledge, accuracy corrects the false security that sincerity can provide, by establishing the need to contrast the inner sense of security we have that we are telling the truth with some external element, allowing us to share and reinforce this sense of security with others. Thus, a gnostic statement would be a 'justified belief': "One that is arrived at by a method, or supported by considerations, that favour it, not simply by making it more appealing or whatever, but in the specific sense of giving reason to think it is true" (Williams, 2002, p. 129). The problem here lies in adequately justifying that what is said is true. This is the key question in epistemology. To refrain from providing a detailed description of this

endless philosophical discussion, we will focus on the issues most relevant to the epistemology of journalism.

The most persistent justification of knowledge is the suspicion that there is some correspondence between statements and reality. Thus, what we say about reality is reality. The simplest expression of this concept is the 'Tarski sentence': 'Snow is white' (statement) is true if and only if snow is white (reality). This is the approach inherent to the correspondence theory (Haack, 2019), which associates reality and truth, yet omits the subject, who is perceived as a contaminant, because the subject introduces their biases into their statements about reality. This is the basis of the idea of objectivity in journalism, characterised by the strict separation of information (pure facts, reality) and opinion (values, the subject; Maras, 2013).

The problem with the 'Tarski sentences' is that they only work with very simple logical-formal statements, but not with news: Readers are rarely able to verify a statement against their perception of the events. Correspondence also raises a circular problem: To justify a statement generated based on a perception, we need another perception of the reality, which is what we want to justify. The only way to escape this vicious circle is through a metaphysical justification, as posed by Aristotle: Between reality and statements there is a shared essence, logos.

Pragmatist philosophers avoided this metaphysical dimension by invoking utility: The truth depends on its practical results, on what we can do with it, i.e., it may be verified in reality. Pragmatists advocate a form of truth based on 'common sense,' which people apply to their lives without asking themselves big questions (Frankfurt, 2007), because, in pragmatism, the debate about what is the truth is of little importance: According to Charles S. Peirce, truth is the result of an inquiry that is carried on indefinitely, an idea applied to verification in journalism (Kovach & Rosenstiel, 2007).

Despite their differences, the correspondence and pragmatic theories of truth share an imperative view of the truth: It is the judge that resolves whether or not what we say is correct, and once established, the truth is necessary, at least until there appears another truth that explains reality better. The imperative nature of truth, the result of the universal sense of reason that emerged during the Enlightenment, has been the subject of constant criticism over the past century (Falomi, 2019).

Constructivist epistemology tempered this imperative nature by focusing on the social processes that construct reality, turning it into a 'social reality,' which, according to Niklas Luhmann (2000), is the result of communicative processes. While constructivism does not deny the existence of an external reality, it considers it of secondary importance, reduced to the mission of providing materials to build the social reality. And by dissociating the social reality from the physical reality, the question of truth becomes blurred. Studies in the field of journalism began integrating constructivism with

Gaye Tuchman (1978): Journalism does not reflect reality, it constructs it (Poerksen, 2011), with the risk of anti-realism and relativism that comes with bracketing its ontological basis (Hearns-Branaman, 2016).

Foucault also rejects correspondence and imperative truth, yet approaches the issue differently: He goes beyond epistemology and frames it within a broader concept, the RoT, which is chiefly concerned not with how reality is constructed (as in constructionism), but with how truth is produced (giving truth an historical and contingent character):

> By 'truth' it is meant a system of ordered procedures or the production, regulation, distribution and circulation of statements….'Truth' is linked by a circular relation to systems of power, which produce it and sustain it, and to effects of power, which it induces and which redirect it. (Foucault, 1977, p. 14)

The RoT is structured around power and subjectivity: Power imposes its interpretation of reality (what is true) on the individual, and the individual constructs their subjectivity by integrating this schema and accepting it as conviction, basing their knowledge on it:

> Each society has its regime of truth, its 'general politics' of truth: That is, the types of discourse it harbours and causes to function as true; the mechanisms and instances which enable one to distinguish true from false statements, the way in which each is sanctioned; the techniques and procedures which are valorised for obtaining truth; the status of those who are charged with saying what counts as true. (Foucault, 1977, p. 13)

Thus, truth ceases to justify knowledge and becomes an instrument for the hegemonic forms of society, which imposes its view of the world not by violent imposition, but through the different means of socialisation, including media and journalism: "News discourse can be seen as a particular instance of the more general 'will to truth' which motivates and constrains institutional forms of knowing in modern society" (Matheson, 2004, p. 445).

For Foucault, the RoT in place since the modern era is the scientific or epistemological conception of truth, 'truth-demonstration,' targeted on reality and a 'technology of demonstration,' characterised by the omnipresence of truth ("the question of truth can be posed about anything and everything") and universal access to truth, in the sense that the subject, to grasp the truth, relies on "the instruments required to discover it, the categories necessary to think it and an adequate language for formulating it in propositions, and not on the 'mode of being' of the subject himself or herself" (Foucault, as cited in Lorenzini, 2016, p. 64).

Thus, truth, far from being transcendental, is moulded to fit the political, social, economic and cultural environment in which the subject operates, who adopts it

COGITATIO

as their own. Subsequently modifying that environment could lead to a reconsideration of what is true, which would vary according to the new forms of hegemony. This is what may be occurring with neoliberalism, under which a form of individualism that affects the way individuals construct their subjectivity has become more intense. According to David Harvey (2005, p. 23), individualism is one of the foundations of neoliberalism:

> All forms of social solidarity were to be dissolved in favour of individualism, private property, personal responsibility and family values. The ideological assault along these lines that flowed from Thatcher's rhetoric was relentless. 'Economic are the method,' she said, 'but the object is to change the soul.'

In the same line, the sociologist Ulrich Beck asserts that individualism is the hallmark of current modernity:

> The basic figure of fully developed modernity is the single person….The form of existence of the single person is not a deviant case along the path of modernity. It is the archetype of the fully developed labour market society. The negation of social ties that takes effect in the logic of the market begins in its most advanced stage to dissolve the prerequisites for lasting companionship. (Beck, 1992, pp. 122–123)

This growing individualism has been studied in the field of social communication in relation to the development of new technologies, particularly social media, which are regarded as promoting new models of human relationships such as 'individual networking' or 'networked individualism' (Rainie & Wellman, 2012). Jayson Harsin (2018) draws a link between these forms of digital communication and an underlying logic focused on recent forms of consumer capitalism, such as the 'attention economy,' in which the problem no longer lies in accessing information, but in how news media capture the audience's attention in increasingly personal and individual ways, segmenting audiences in the same manner as marketing, and in the increasing use of cognitive-oriented commercial tools in political communication and marketing. Maddalena and Gili (2020) agree that the interest of sociology and psychology in understanding human behaviour, in personalising the messages broadcast by commercial, political and media sources, is one of the keys elements in the current individualisation process; one that changes the way individuals, increasingly dependent on their emotions and personal beliefs, think, feel and act. Journalism studies have confirmed this increase in the emotional content of media (Papacharissi, 2014) and the change this is having on information:

> As journalism and society change, emotion is becoming a much more important dynamic in how news is produced and consumed. Emphasising emotion as the key redefines the classic idea of journalistic

objectivity—indeed, it is reshaping the idea of news itself. (Beckett & Deuze, 2016, p. 2)

The result of this individualisation process, based on the marketization of information and the primacy of emotional content, would be, according to Harsin, the proliferation of 'truth games' within communication markets devoid of an authority that imposes a truth.

3. The Reality of News Media: Beyond the Facts

Maurizio Ferraris warned that the crisis of epistemology in the 20th century has called into question the idea of reality, and that this epistemological confusion may lead to the belief that "the real world ended up being a tale" (Ferraris, 2014, p. 2). Journalism has, until now, been grounded in its faith in facts ('facts are sacred') and has justified this faith through the notion of objectivity. In journalism, epistemological doubts have been formulated almost exclusively in the academic sphere (Muñoz-Torres, 2012), as the journalistic profession has remained a staunch defender of objectivity (Maras, 2013), either as a reflection of reality, as a process in which the truth gradually takes shape or as a ritual through which journalists justify their profession (Tuchman, 1978). Few authors, such as Hearns-Branaman (2016) and his adaptation to Baudrillard's hyperreality, have questioned reality as an a priori. Concern for the ontology of journalism has focused on the emergence of new actors that spread news through social media (Ryfe, 2019).

However, concern about fake news, considered the most visible manifestation of post-truth (McIntyre, 2018), has indirectly revealed the problem surrounding the ontology of journalism, by reinforcing the objectivist approach to news. Discussion has revolved around defining what constitutes fake news and what it brings to the age-old practice of disinformation (Tandoc, Zheng, & Ling, 2018), and around offsetting its effects through fact-checkers and media literacy (Wardle & Derakhshan, 2017). Yet the struggle against fake news remains ineffective (Chan, Jones, Jamieson, & Albarracín, 2017; Clayton et al., 2020; Thorson, 2016). Pennycook and Rand (2017) cautioned against (a) the limited effect that labelling information as false has on readers and (b) the 'implied truth effect,' i.e., branding certain news as false leads people to believe that the rest has been verified. Other studies have revealed just how little use audiences make of fact-checkers (Guess, Nyhan, & Reifler, 2018) and the fact that people are unable to perceive a clear difference between fake and real news (Nielsen & Graves, 2017). These difficulties suggest that perhaps the problem does not reside in finding the correct verification method, but rather in what is verified, in the sense that not all methods allow for the same degree of verification because not all journalistic events allude to the same type of reality. To justify this assertion, we will analyse the typologies of events in journalism based on the hypothesis that, as

opposed to Ferrari's idea, the heterogeneous nature of these events is causing epistemological confusion. This heterogeneity is not new: Harvey Molotch and Marilyn Lester (1974, p. 106) described it in their research on how "public events" are presented to journalists based on "the circumstances of the promotion work which makes them available to publics." In our analysis of journalistic ontology, we will take the Molotch and Lester typology as a basis and update it with the changes brought about by digital communication.

First, Molotch and Lester refer to 'accidents,' in which (a) the event from which they derive is not intentional and (b) whoever reports it as news has not been affected by the 'accident' or hopes to benefit from it. An 'accident' is a surprise to everyone: to the witnesses of the event, to the journalist who, as a truth-teller, collects the witnesses' accounts, and to the institutions, public or private, affected by the 'accident.' Nonetheless, social media has emerged as a new collector of 'accidents,' interacting with professional truth-tellers (journalists, official sources) as part of a 'hybrid' media system (Chadwick, 2013). People not only provide accounts to the journalist, but are also able to relay them directly to the media sphere: "One thing that crowds do better than journalists is collect data" (Anderson, Bell, & Shirky, 2012, p. 24). As a result, the role of the journalist has been thrust into a crisis (Broersma, 2013), and the factual truth of the news media is forced to compete with the truths of non-journalists who are not held to journalistic standards (Deuze & Witschge, 2017).

The second type of events are 'routine events,' which:

> Are distinguishable by the fact that the underlying happenings on which they are presumably based are purposive accomplishments and by the fact that the people who undertake the happening (whom we call 'effectors') are identical with those who promote them into events. (Molotch & Lester, 1974, p. 106)

There is no surprising occurrence at their origin, but rather a source who creates events, reports them to a journalist (as a witness) and directs their meaning for their benefit. Such is the case of the institutional statements that have colonised the news media (Berkowitz, 2009). These 'verbal manifestations' are a reality whose sole purpose is to be communicated and induce effects through public disclosure. Daniel J. Boorstin (1987) labels them 'pseudo-events,' which he defines as (a) 'not spontaneous,' but rather planned by the party concerned, (b) planted for the purpose of gaining media coverage, (c) discussion surrounding the nature of the event is limited to determining whether it has happened and why, and (d) such events are a self-fulfilling prophecy: Interviewing a 'distinguished' person makes that person distinguished.

The hegemony of institutions as truth-tellers turns 'routine events' into a key tool for political commu-

nication and promotional culture, something that has become standardised in journalism through the use of quotes, which should be accurate and balanced. Thanks to this journalistic routine, institutions generate stories about themselves (statements are signs of power) and about the reality they communicate (what they say is real), and they do so by pushing the possibilities of language beyond the strict confines of reality (Hearns-Branaman, 2016). By using Austin's 'speech acts' (1962), we can further explore the linguistic performance of these institutions, which create and spread 'routine events' through: (a) Information locutions, when the institution reports events that cannot be known through other channels (official data, internal events)—these events are difficult to verify, unless some manner of internal betrayal occurs, which, as we will see later, would be considered a 'scandal'; (b) illocutionary statements, when the institution mentions something known by the journalist (and audience) to explain, clarify and interpret with a view to fixing the meaning of the event—these statements do not usually provide new facts (except for the statement itself), yet feed public opinion through discussion in the media; (c) perlocutionary statements, when the institution announces it is going to do something—what is newsworthy is the institution's commitment, yet the only verifiable event is the statement, created to be reported by the journalist and for the institution to gauge the audience's reaction, which may generate a verifiable event.

Nonetheless, the power of institutions as truth-tellers is being called into question due to two factors: (a) the loss of credibility currently being experienced by all institutions (Edelman, 2020), and (b) social media's capacity to provide individuals direct access to public debates and even constitute an official source when able to capture a collective sentiment and become its spokesperson (Masip, Ruiz, Suau, & Puertas, 2020), and, in extreme circumstances, even stand as a threat to these institutions, as occurred during the Arab Spring (Wolfsfeld, Segev, & Sheafer, 2013).

The third type of event are 'scandals,' an anomaly among 'routine events,' in which a source (anonymous) intentionally breaks with the meaning fixed by the institution. This is the category that encompasses 'investigative journalism,' based on revealing what's hidden. A 'scandal' is unexpected for the institution that tries to conceal it, but may be planned by the person who reveals it. Social media greatly increases the chances of there being a 'scandal,' given its unprecedented dissemination capacity, without the intermediary of journalists, as demonstrated by the WikiLeaks case (Marmura, 2018).

The last event type is 'serendipity': The news story originates in an unanticipated event ('accident'), which is handled by an 'effector' (an institution) as if it were a 'routine event.' This occurs when news stories about natural disasters are handled by institutions, as has been the case with the Covid-19 pandemic. With serendipity, different public discourses compete to impose meaning

on an event, as analysed by agenda setting and framing (D'Angelo, 2019; McCombs, Shaw, & Weaver, 2014).

Yet, to round off the ontological framework of journalism, a fifth element should be added to the typology proposed by Molotch and Lester, data, given their importance in journalism and their epistemological prestige. Even though behind each datum there is usually an institution (the only entity with the capacity to create it), it is presented as an 'objective witness' of the reality: A seemingly aseptic and neutral 'information package' which lends the news story a factual basis.

The datum operates as a concept: It reduces the complexity of the reality by selecting the quantifiable aspects of an event and discarding all others. Once obtained, the datum may be incorporated into homogeneous datasets to compare, infer or anticipate results. Can a datum be verified? Reality is no help in this regard: it is an abstract of reality, not a raw sample of reality (Rosenberg, 2013). The verification of a datum is methodological (how was it obtained?), although the selection of data used to explain an aspect of reality may always be subject to dispute.

Digitisation has enhanced the ability to use data and has been harnessed by journalism to apply Big Data techniques (Lewis & Westlund, 2015), intensifying the tendency to datify virtually all aspects of our existence: "Datafication is a contemporary phenomenon which refers to the quantification of human life through digital information, very often for economic value" (Mejias & Couldry, 2019, p. 1). The datum, when incorporated into a news story, is situated on an ontological level similar to that of facts and statements, despite belonging to substantially different realities (Uscinski & Butler, 2013).

From this capacity of the news media to compress various realities emerges Baudrillard's notion of 'hyperreality' (Baudrillard, 1994), whereby any element of reality, or fiction, matched by the 'common code' (technological and symbolic) used for dissemination by the media becomes something other than reality (Hearns-Branaman, 2016). Hyperreality is a simulacrum, an illusion of reality through which journalism justifies itself as a profession. Hyperreality replaces the physical reality with the media reality, yet the reality that fact-checkers try and verify reduces the media reality to only physical events, ignoring all of the other realities depicted by the news media.

Of these 'public events,' only 'accidents' can be considered verifiable from an objectivist point of view, as they come close to the idea of 'pure fact' without human intervention. Yet reducing all the realities the news media spread to 'accidents' implies leaving out other events, despite the fact that they form the basis of a large amount of news. This is not only an ontological problem: Forcing the objectivist validation of facts shaped by human intervention requires the epistemology to do the impossible.

4. Epistemology: Just Me

As we have seen, news media spread news based on facts that people are virtually unable to verify by themselves (Read & Uçan, 2019), statements by institutional actors who may or may not be trusted, with no details as to how the data were created or for what purpose:

> The ability of mere individuals to understand the social world has decreased because they do not have the tools to comprehend what is happening around them or the meaning of events and their consequences, let alone the possibility of directing or influencing those events. (Maddalena & Gili, 2020, p. 6)

The subject is not only confronted with the media's heterogeneous reality, but, in a digital context, they do so increasingly by themselves, in the sense that the crisis of institutional authority dilutes the global processes of knowledge construction within societies (Berger & Luhman, 1967) and that digital media and social media tend to isolate the individual in bubbles resistant to any input that jeopardises constructed subjectivity (Fletcher & Nielsen, 2018; Sunstein, 2018). Katharine Viner (2016) and Jihii Jolly (2014) examined how algorithms heighten this individualisation by personalising searches, prioritising past results and restricting access to new information. While several other studies (Trilling & Schoenbach, 2015) have dissected the processes of 'selective exposure' in traditional media; a phenomenon intensified by the Internet (Laybats & Tredinnick, 2016). This results in fragmented audiences and transforms the public sphere, taken to mean a 'general conversation' (the Habermasian public opinion) on the truth of the matter (Sunstein, 2009), dividing it into a series of 'partial conversations' tailored to reflect each group's expectations. Segmentation, typical of marketing, is present in journalistic practices: Tandoc and Vos (2016) dubbed it the 'marketization' of information; for Maria Karidi (2018) it constitutes the application of 'commercial logic' to new media; while, according to Harsin (2015), truth-tellers have become 'Truth markets,' groups competing to impose their truth with no ultimate authority.

A fragmented audience does not imply setting aside the idea of truth, but rather questioning the authority that establishes this truth. In Steve Fuller's (2018) opinion, the battles waged over the truth are not battles of the first order (what is true and what is false), but rather battles of the second order (the criteria of truth and who determines these criteria). For Yael Brahms (2020, p. 16), "In the post-truth era, the power to decide between the narratives is no longer held by the customary sources of authority, but rather, is held by anyone who positions himself opposite these sources of authority."

The lack of authority affects not only truth, but also journalistic facts. An objective reality, such as 'accidents,' implies a single reference, in which sense it would be as authoritative as the truth of correspondence theories

and pragmatism: There would be only one reality, which remains out there, beyond our control. But facts in which humans intervene, in one way or another, can give rise to what the comedian Stephen Colbert described, in a satirical manner, as 'truthiness': "Everyone was entitled to their own opinion, but not their own facts. But that's not the case anymore. Facts matter not at all" (Rabin, 2006). For Jeffrey Jones, this 'truthiness' represents an emblematic change in the journalistic RoT based on 'truth in fact,' hegemonic until now, which has transformed into an RoT in which a group of actors (citizens, politicians, journalists) create 'believable fictions,' defined as "constructions of reality where truth in fact is less important than truth in essence. Indeed, the word 'truthiness' is designed to highlight this sleight of hand in the contested terrain of politically motivated constructions of truth" (Jones, 2009, p. 135). Numerous studies have shown how deniers, of everything from vaccines to climate change, reject any scientific fact that does not fit their narratives, and instead accept facts provided by their own supposed experts who reaffirm those narratives (Diethelm & McKee, 2009).

With fragmented and atomised audiences and a slew of facts with no authority capable of justifying them beyond doubt, what is the criterion of truth in the post-truth era? In the framework of 'truth-demonstration,' a gnostic statement was a 'justified belief,' which forced the subject to leave themselves to validate their statement against reality. But, if emotions and beliefs are at the centre of the new RoT, statements no longer need external validation for the subject to accept them as authentic, making them fundamental constituents of their own subjectivity. In this way, post-truth removes the need for empirical justification. As a result, the statement is reduced to a belief, reinforced by the experience of truth of sincerity, which makes the individual feel good about themselves: That is what Jordi Ibáñez (2017) calls 'collective hedonism.'

This sincerity, if it aspires to stand as justification of knowledge, must be capable of being shared, yet not on the basis of universal reason, as occurs with the RoT of 'truth-demonstration,' but rather through a series of experiences of truth on the same subjects. And a perfect vehicle for this purpose are communities of believers, fostered by the fragmentation of audiences, which are perceived by news media not only as markets for their advertisers, but as reader markets:

> Marketing implies attention to market demand. In a period when journalists are faced with shrinking audience sizes, decreasing revenues and an overflow of different forms of audience feedback, pandering to audience choice—that is, giving in to market demand—becomes an easy alternative to privileging editorial autonomy. (Tandoc & Vos, 2016, p. 13)

Megan Boler and Elizabeth Davis (2018, p. 82) explored how the "affective feedback loop" promoted by social media is a central element in "shaping the networked subjectivity fundamental to computational propaganda and algorithmic governance."

One of the common arguments is that post-truth is just another name for the common lie. Yet participants in one of these communities of believers do not believe they are lying, in the sense that they do not formulate beliefs in which they do not believe, but rather use their belief to justify their knowledge, which is therefore perceived as being certain. And given the adjustment problems that may arise between their beliefs and other beliefs or reality, the subject fills in their knowledge gaps "using a set of beliefs and personal opinions, their sentiments towards this or that politician, their confidence or lack thereof in the various sources of information, and their personal interpretation of the information made public" (Brahms, 2020).

Knowledge gained in this manner finds in news an ideal vehicle for formulating and transmitting itself, since facts, statements and data presented by the news media become meaningful to the individual. News is a story based in reality (Schudson, 2005), albeit the objective reality of correspondence, the social reality of constructivism or Baudrillard's notion of hyperreality. Thanks to this narrative form, the truth of the news does not stem solely from its semantic content, but is part of a Wittgensteinian 'language game' that involves symbolic elements and the rules inherent to any story. The narrative searches for the truth of life, and does not seek to reflect reality as it is (Lule, 2001). This centrality of narration is what Lynn Smith labelled 'narrative turn':

> Since the postmodern literary movement of the 1960s swept out of academia and into the wider culture, narrative thinking has seeped into other fields. Historians, lawyers, physicians, economists and psychologists have all rediscovered the power of stories to frame reality, and storytelling has come to rival logic as a way to understand legal cases, geography, illness or war. Stories have become so pervasive, critics fear they have become a dangerous replacement for facts and reasoned argument. (Smith, 2001)

Having removed literal (semantic) meaning as the only means of understanding a text, the statement is relieved of all need for references (D'Ancona, 2017): The fact becomes a free signifier in search of meaning, and the meaning ends up configuring the signifier, not the other way around. The subject, armed with the confidence that their sincerity confers on them, having built their subjectivity in an RoT characterised by individualism, finds no obstacles to prioritising the meaning they want to reaffirm a priori, modelling the signifier at their convenience.

5. Conclusions

There is sufficient evidence to support the idea that a change is occurring in the way people perceive reality

through the news, and that this shift is affecting the perception of what is true or false in the news. While this change has been dubbed 'post-truth,' without specifying exactly what it is, it could really be called anything, because what matters is not the name, but the phenomenon itself. In this article, we have looked to substantiate the term post-truth through what we have called 'epistemological mutation,' which eliminates the subject's need to validate their statements against reality (which was the foundation of the 'truth-demonstration' of the RoT that emerged during the Enlightenment) and replaces it with the sense of security that stems from the sincerity with which the subject formulates their statements, in an context in which individualism has weakened social ties and the construction of knowledge has ceased to be a global endeavour. To substantiate the existence of this mutation, we have framed it within the Foucaultian concept of the RoT, which views epistemology as a product of the hegemonic forms existing in society at a specific moment in history. In this way, this mutation likely corresponds to a change in the RoT resulting from the triumph of a neoliberal form of rationality that has permeated all aspects of life, enhancing individuality and shaping social communication. Digital technologies have accelerated and intensified this change, spreading a neoliberal form of economic logic that tends to reduce human experience to marketing and datafication. We believe that this theoretical interpretative framework could help to pinpoint the origin of some of the current problems (such as the struggle against fake news) and conceptually frame some of the incessant transformations that are taking place in social communication and journalism.

Acknowledgments

Conflict of Interests

The author declares no conflict of interests.

References

Anderson, C. W., Bell, E., & Shirky, C. (2012). *Post-industrial journalism: Adapting to the present*. New York: NY: Tow Center for Digital Journalism.

Austin, J. L. (1962). *How to do things with words*. London: Oxford University Press.

Baudrillard, J. (1994). *Simulacra and simulation*. Ann Arbor, MI: University of Michigan Press.

Beck, U. (1992). *Risk society: Towards a new modernity*. London: Sage.

Beckett, C., & Deuze, M. (2016). On the role of emotion in the future of journalism. *Social Media + Society*, *2*(3), 1–6.

Berger, P., & Luhman, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge*. London: Penguin Books.

Berkowitz, D. (2009). Reporters and their sources. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The handbook of journalism studies* (pp. 102–115). New York, NY: Routledge.

Biesecker, B. A. (2018). Introduction: Toward an archaeogenealogy of post-truth. *Philosophy & Rhetoric*, *51*(4), 329–341.

Boler, M., & Davis, E. (2018). The affective politics of 'post-truth' era: Feeling and networked subjectivity. *Emotion, Space and Society*, *27*, 75–85.

Boorstin, D. J. (1987). *The image: A guide to pseudo-events in America* (25th ed.). New York, NY: Atheneum.

Brahms, Y. (2020). Philosophy of post-truth. *Institute for Nacional Security Studies*. Retrieved from https://www.jstor.org/stable/resrep23537?seq=1#metadata_info_tab_contents

Broersma, M. (2013). A refractuted paradigm: Journalism, hoaxes and the challenge of trust. In C. Peters & M. Broersma (Eds.), *Rethinking journalism: Trust and participation in a transformed news landscape* (pp. 28–44). Oxon: Routledge.

Brown, W. (2015). *Undoing the demos: Neoliberalism's stealth revolution*. New York, NY: Zone Books.

Chadwick, A. (2013). *The hybrid media system: Politics and power*. Oxford: Oxford University Press.

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

Clayton, K., Blair, S., Busam, J. A., Forstner, S., Glance, J., Green, G., . . . Nyhan, B. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behaviour*, *42*, 1073–1095.

Cotoi, C. (2011). Neoliberalism: A Foucauldian perspective. *International Review of Social Research*, *1*(2), 109–124.

D'Ancona, M. (2017). *Post truth: The new war on truth and how to fight back*. London: Ebury Press.

D'Angelo, P. (2019). Framing theory and journalism. In P. T. Vos, F. Hanusch, A. Sehl, D. Dimitrakopoulou, & M. Geertsema-Sligh (Eds.), *The international encyclopedia of journalism studies* (pp. 1–10). New York, NY: Wiley.

Deuze, M., & Witschge, T. (2017). Beyond journalism: Theorizing the transformation of journalism. *Journalism*, *19*(2), 165–181.

Diethelm, P., & McKee, M. (2009). Denialism: What is it and how should scientists respond? *European Journal of Public Health*, *19*(1), 2–4.

Edelman. (2020). 2020 Edelman trust barometer. *Edelman*. Retrieved from https://www.edelman.com/trustbarometer

Falomi, M. (2019). Two forms of domination by reason. *Nordic Wittgenstein Review*, *8*, 43–61.

Ferraris, M. (2014). *Manifesto of new realism*. Albany, NY: The State University of New York Press.

Finlayson, L. (2019). What to do with post-truth. *Nordic Wittgenstein Review*. 63–79. https://doi.org/10.15845/nwr.v8i0.3502

Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society*, *20*(7), 2450–2468.

Foucault, M. (1977). The political function of the intellectual. *Radical Philosophy*, *17*, 12–14.

Frankfurt, H. G. (2007). *On truth*. London: Pimlico.

Fuller, S. (2018). *Post truth: Knowledge as a power game*. London, Anthem.

Gane, N. (2014). Sociology and neoliberalism: A missing history. *Sociology*, *48*(6), 1092–1106.

Gjorgjioska, M. A., & Tomici, A. (2019). The crisis in social psychology under neoliberalism: Reflections from social representations theory. *Journal of Social Issues*, *75*(1), 169–188.

Goldstein, R. A., Macrine, S., & Chesky, N. (2011). Welcome to the "new normal": The news media and neoliberal reforming education. *Journal of Inquiry & Action in Education*, *4*(1), 112–131.

Guess, A., Nyhan, B., & Reifler, J. (2018). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign*. Brussels: European Research Council. Retrieved from http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf

Haack, S. (2019). Post 'post-truth': Are we there yet? *Teoria*, *85*, 25–275.

Harsin, J. (2015). Regimes of posttruth, postpolitics, and attention economies. *Communication, Culture & Critique*, *8*, 327–333.

Harsin, J. (2018). Post-truth and critical communication. *Oxford Research Encyclopedia of Communication*. Retrieved from https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-757

Harvey, D. (2005). *A brief history of neoliberalism*. Oxford: Oxford University Press.

Hearns-Branaman, J. O. (2016). *Journalism and the philosophy of truth: Beyond objectivity and balance*. New York, NY: Routledge.

Ibáñez, J. (2017). Una introducción [An introduction]. In J. Ibáñez (Ed.), *En la era de la posverdad: 14 ensayos* [In the era of post-truth: 14 essays] (pp. 11–36). Barcelona: Calambur Editorial.

Jolly, J. (2014, May 20). How algorithms decide the news you see. *Columbia Journalism Review*. Retrieved from https://archives.cjr.org/news_literacy/algorithms_filter_bubble.php

Jones, J. P. (2009). Believable fictions: Redactional culture and the will of truthiness. In B. Zelizer (Ed.), *The changing faces of journalism: Tabloidization, technology and truthiness* (pp. 127–143). New York, NY: Routledge.

Karidi, M. (2018). News media logic on the move? In search of commercial media logic in German news. *Journalism Studies*, *19*(9), 1237–1256.

Kovach, B., & Rosenstiel, T. (2007). *The elements of journalism: What news people should know and the public should expect*. New York, NY: Three Rivers Press.

Laybats, C., & Tredinnick, C. (2016). Post truth, information, and emotion. *Business Information Review*, *33*(4), 204–206.

Lewis, S., & Westlund, O. (2015). Big data and journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism*, *3*(3), 447–466.

Leyva, R. (2020). *Brains, media, and politics: Generating neoliberal subjects*. Abingdon and New York, NY: Routledge.

Lorenzini, D. (2016). Foucault, regimes of truth and the making of the subject. In L. Cremonesi, O. Irrera, & D. Lorenzini (Eds.), *Foucault and the making of subjects* (pp. 63–76). London and Lanham, MD: Rowman & Littlefield International.

Luhmann, N. (2000). *The reality of mass media*. Stanford, CA: Stanford University Press.

Lule, J. (2001). *Daily news, eternal stories: The mythological role of journalism*. New York, NY: The Guilford Press.

Maddalena, G., & Gili, G. (2020). *The history and theory of pos-truth*. London: Palgrave Macmillan.

Maras, S. (2013). *Objectivity in journalism*. Cambridge: Polity Press.

Marmura, S. (2018). *The WikiLeaks paradigm: Paradoxes and revelations*. Basingstoke: Palgrave Macmillan.

Masip, P., Ruiz, C., Suau, J., & Puertas, D. (2020). Media and Twitter agendas for social mobilizations: The case of the protests in defense of the public healthcare system in Spain. *International Journal of Communication*, *14*(2020), 3355–3376.

Matheson, D. (2004). Weblogs and the epistemology of the news: some trends in online journalism. *New Media & Society*, *6*(4), 443–468.

McCombs, M., Shaw, D. L., & Weaver, D. H. (2014). New directions in agenda-setting theory and research. *Mass Communication and Society*, *17*, 781–802.

McIntyre, L. (2018). *Post-truth*. Cambridge, MA and London: The MIT Press.

Mejias, U. A., & Couldry, N. (2019). Datafication. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1428

Molotch, H., & Lester, M. (1974). News as purposive behavior: On the strategic use of routine events, accidents and scandals. *American Sociological Review*, *39*(1), 101–112.

Muñoz-Torres, J. R. (2012). Truth and objectivity in journalism: Anatomy of an endless misunderstanding. *Journalism Studies*, *13*(4), 566–582.

Nielsen, R. K., & Graves, L. (2017). *"News you don't*

believe": Audience perspectives on fake news. Oxford: Reuters Institute for the Study of Journalism.

Papacharissi, Z. (2014). *Affective publics: Sentiment, technology, and politics*. Oxford: Oxford University Press.

Pennycook, G., & Rand, D. G. (2017). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analitic thinking. *Journal of Personality*, *88*(2), 185–200.

Poerksen, B. (2011). *The creation of reality: A constructivist epistemology of journalism and journalism education*. Charlottesville, VA: Imprint Academic Philosophy Documentation Center.

Prozorov, S. (2019). Why is there truth? Foucault in the age of post-truth politics. *Constellations*: *An International Journal of Critical and Democratic Theory*, *26*, 18–30.

Rabin, N. (2006, January 25). Interview: Stephen Colbert. *A. V. Club*. Retrieved from http://tv.avclub.com/stephen-colbert-1798208958

Rainie, L., & Wellman, B. (2012). *Networked: The new social operating system*. Cambridge, MA: MIT Press.

Read, R., & Uçan, T. (2019). Introduction: 'post-truth'? *Nordic Wittgenstein Review*. https://doi.org/10.15845/nwr.v8i1.3508

Rosenberg, D. (2013). Data before the fact. In L. Gitelman (Ed.), *"Raw data" is an oxymoron* (pp. 15–40). Cambridge, MA: The MIT Press.

Ryfe, D. (2019). The ontology of journalism. *Journalism*, *20*(1), 206–209.

Schudson, M. (2005). News as stories. In E. Rothenbuhler & M. Coman (Eds.), *Media anthropology* (pp. 121–128). Thousand Oaks, CA: Sage.

Smith, L. (2001, November 11). Not the same old story. *Los Angeles Times*. Retrieved from http://articles.latimes.com/2001/nov/11/news/cl-2758

Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton, NJ:
Princeton University Press.

Sunstein, C. R. (2018). *#Republic: Divided democracy in the age of social media*. Princeton, NJ: Princeton University Press.

Tandoc, E., & Vos, T. P. (2016). The journalist is marketing the news. *Journalism Practice*, *10*(8), 950–966.

Tandoc, E., Zheng, W. L., & Ling, R. (2018). Defining "fake news." *Digital Journalism*, *6*(2), 137–153.

Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, *33*, 460–480.

Trilling, D., & Schoenbach, K. (2015). Challenging selective exposure. *Digital Journalism*, *3*(2), 140–157.

Tuchman, G. (1978). *Making news: A study in the construction of reality*. New York, NY: Free Press.

Uscinski, J. E., & Butler, R. W. (2013). The epistemology of fact checking. *Critical Review: A Journal of Politics and Society*, *25*(2), 162–180.

Viner, K. (2016, July 12). How technology disrupted the truth. *The Guardian*. Retrieved from https://www.theguardian.com/media/2016/jul/12/how-technology-disrupted-the-truth

Wahl-Jorgensen, K. (2016). Emotion and journalism. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida (Eds.), *The SAGE handbook of digital journalism* (pp. 128–143). London: SAGE.

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making* (Report DGI(2017)09). Strasbourg: Council of Europe.

Williams, B. (2002). *Truth and truthfulness*. Princeton, NJ: Princeton University Press.

Wolfsfeld, G., Segev, E., & Sheafer, T. (2013). Social media and the Arab Spring: Politics comes first. *The International Journal of Press/Politics*, *18*(2), 115–137.

About the Author

Pablo Capilla is a Journalist and Professor at the Blanquerna School of Communication and International Relations (Ramon Llull University), where he teaches subjects related to journalism, linguistics and international relations. He is also a Professor in the Master's in Political Communication and in the Master of International Journalism at the same university. He is part of the research group "Digilab Media, Strategy and Regulation," which studies innovation in the media and the participation of citizens in the digital public sphere.

Article

Digital Disinformation and Preventive Actions: Perceptions of Users from Argentina, Chile, and Spain

Jordi Rodríguez-Virgili *, Javier Serrano-Puche and Carmen Beatriz Fernández

University of Navarra, 31009 Pamplona, Spain; E-Mails: jrvirgili@unav.es (J.R.-V.), jserrano@unav.es (J.S.-P.), cbfer@datastrategia.com (C.B.F.)

* Corresponding author

Abstract

This article explores audience perceptions of different types of disinformation, and the actions that users take to combat them, in three Spanish-speaking countries: Argentina, Chile, and Spain. Quantitative data from the Digital News Report (2018 and 2019), based on a survey of more than 2000 digital users from each country was used for the analysis. Results show remarkable similarities among the three countries, and how digital users identically ranked the types of problematic information that concerned them most. Survey participants were most concerned by stories where facts are spun or twisted to push a particular agenda, followed by, those that are completely made up for political or commercial reasons, and finally, they were least concerned by poor journalism (factual mistakes, dumbed-down stories, misleading headlines/clickbait). A general index of "Concern about disinformation" was constructed using several sociodemographic variables that might influence the perception. It showed that the phenomenon is higher among women, older users, those particularly interested in political news, and among left-wingers. Several measures are employed by users to avoid disinformation, such as checking a number of different sources to see whether a news story is reported in the same way, relying on the reputation of the news company, and/or deciding not to share a news story due to doubts regarding its accuracy. This article concludes that the perceived relevance of different types of problematic information, and preventive actions, are not homogeneous among different population segments.

Issue

This article is part of the issue "Disinformation and Democracy: Media Strategies and Audience Attitudes" edited by Pere Masip (University Ramon Llull, Spain), Bella Palomo (University of Málaga, Spain) and Guillermo López (University of Valencia, Spain).

1. Introduction

In recent years, academics and the media have been paying attention to the phenomenon of disinformation (Freelon & Wells, 2020; McKay & Tenove, 2020). The fact that disinformation poses a threat to democratic institutions has contributed to mounting concern regarding this problem (Miller & Vaccari, 2020). Disinformation is a modality within a broader field called 'problematic information' (Jack, 2017), which includes various

types of information that are considered inaccurate, misleading, or found to be improperly or totally manufactured. Although it is true that this type of media is not novel, today's information ecosystem is without precedent, due to its sheer scale and scope. This fosters new ways in which problematic information can be created, circulated, and received by users, increasing its potential effects (Lewandowsky, Ecker, & Cook, 2017). Said information ecosystem is technological in nature, but it also has political, social, and economic implications, giving

rise to what Wardle and Derakhshan (2017) have called an 'information disorder.'

2. Theoretical Framework

The academic literature on disinformation has grown remarkably in the last decade (Ha, Perez, & Ray, 2019). Following Freelon and Wells (2020), two main areas of enquiry are particularly worth noting: On the one hand, research that focuses on content, and on the other, reception studies. In the following pages we analyze these areas in detail.

2.1. Types of 'Problematic Information'

Regarding content, various types of information are included under the paradigm of 'problematic information.' Among them, the term 'fake news' has become the most popular. However, Egelhofer and Lecheler (2019) caution that the term fake news should not be used to group instances of falsehood indiscriminately. As the concept should differentiate between fake news as a genre and the fake news label as applied to specific news by political figures to discredit journalism (Khaldarova & Pantti, 2016). For this reason, it is preferable to broaden the range of concepts included in this phenomenon. Allcott and Gentzkow (2017) pointed out six different types of problematic information:

> 1) Unintentional reporting mistakes; 2) rumours that do not originate from a particular news article; 3) conspiracy theories (these are, by definition, difficult to verify as true or false, and they typically originate from those who believe them to be true); 4) satire that is unlikely to be misconstrued as factual; 5) false statements by politicians; and 6) reports that are slanted or misleading but not outright false. (p. 214)

In addition, Tandoc, Lim, and Ling (2018), through a bibliographic review of the academic literature, identify six types, partially coinciding with the above— news satire, news parody, fabrication, manipulation, propaganda, and advertising—which they organize along two dimensions according to their level of facticity and deception.

According to the High Level Expert Group on Fake News and Disinformation, designated by the European Commission, disinformation includes "all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit" (High Level Expert Group on Fake News and Disinformation, 2018, p. 5). This definition introduces a parameter that helps catalogue the cast of disinformation genres, namely the content creator's motivation. Indeed, the intention to cause harm or to seek profit is a vector that allows malicious disinformation and deliberately polarized content to be distinguished from other situations where online content may be untruthful, such

as satirical news or journalistic pieces that turn out to be imprecise due to lack of professionalism. Although such pieces do not intend to cause deception or confusion, this may in fact, be the result. The content itself, its factuality, the lexical and syntactic features within it, and any evidence presented might inadvertently cause confusion. The sources that are consulted for these journalistic pieces, the possible intentions of the issuing agent, and other structural elements (such as URL, or website transparency), allow for the establishment of a taxonomy of modalities of false content that contrast with the 'real news.' These range from fabricated news to disguised advertising, including also parodic news or information from hyperpartisan sources (Molina, Sundar, Le, & Lee, 2019). Therefore, we ask:

> RQ1: What are the types of disinformation that raise more concern among digital users?

2.2. Factors that Influence Disinformation Reception

Firstly, and from a macro point of view, researchers have centred on which structural conditions of countries can lead to a greater or lesser resilience towards disinformation. Following Humprecht, Esser, and Van Aelst (2020), carried out cross-national comparative research in 18 countries, these factors include the country's level of social polarization, political populism, trust in the news, or the strength of the public media, among others.

Descending to the individual level, a review of the literature suggests that exposure and vulnerability to disinformation are mediated by several different factors: political ideology, age, gender, and level of interest in public affairs. Regarding ideological self-positioning, some studies agreed that the American right-wing is substantially more vulnerable to disinformation attacks, and more likely to accept them, than the left-wing. It is an 'ideological asymmetry' (Freelon et al., 2020), showed in two different dimensions: First, the content dimension, wherein disinformation providers produce vastly more conservative-oriented messages than liberal ones (Allcott & Gentzkow, 2017; Howard, Ganesh, Liotsiou, Kelly, & Francois, 2018; Mckay & Tenove, 2020); second, the reception dimension, showing that conservatives are also more likely to engage with the disinformation messages that target them than liberals (Grinberg, Joseph, Friedland, Swire-Thompson, & Lazer, 2019; Hjorth & Adler-Nissen, 2019).

Regarding age as a significant variable in relation to disinformation reception, Guess, Nagler, and Tucker (2019) identify a strong age effect, through their research on individual-level characteristics associated with sharing false articles during the 2016 US presidential campaign. Respondents in an older age category were more likely to share fake news than respondents in the next-youngest group. Overall, users over 65 years old, shared nearly seven times as many articles from fake news domains as the youngest age group. Other

studies, such as that of Serrano-Puche, Fernández, and Rodríguez-Virgili (2021), using the Venezuelan political landscape as a case study, have analysed vulnerability to disinformation as dependent, besides age, on educational level and the main source of information (analogue or digital).

The relationship between gender and attitudes towards disinformation has been understudied and lacks conclusive results. Studies such as Reuter, Hartwig, Kirchner, and Schlegel (2019) have found significant differences in Germany among men and women regarding perceptions of disinformation's effects on society and the agents that cause it. However, gendered differences were not found among ways to react to disinformation. In a similar vein, in a survey with Portuguese college students, Morais and Cruz (2020) found gender differences in the use of media and information consumption, but not in the skills to distinguish between credible and false sources of information.

Regarding interest in political matters, research indicates that using social media for news can lead to the spread of misinformation, albeit indirectly, due to its association with individuals' political participation (Valenzuela, Halpern, Katz, & Miranda, 2019). Although being politically engaged does not make people more or less likely to be misinformed, active users are more likely to share inaccurate contents than those who are less politically engaged (Bail et al., 2020). Other studies indicate that political participation can promote the formation of closed groups in which disinformation is more likely to be sent and received. In this way, with the consolidation of such polarized groups, the opportunity to access divergent information diminishes, favouring selective exposure to partisan sources (Iyengar & Hahn, 2009; Lazer et al., 2018). As with selective exposure, motivated reasoning can contribute to an individual becoming misinformed (Scheufele & Krause, 2019). When users engage in goal-directed processing of new information to protect preexisting values, beliefs, and ideologies (Kunda, 1990), and when such directional goals influence reasoning processes, individuals become prone to 'biased assimilation' (Lord & Taylor, 2009). This is the tendency to privilege information that is consistent with one's predispositions, discrediting information that seems contradictory. Giglietto, Iannelli, Valeriani, and Rossi (2019) affirm that those who receive disinformation can in turn become its propagators—whether intentionally or not. This might generate propagation cascades where the intention of the 'injector' does not determine the future evolution of the false information cycle, meaning that what was born as a parody or a journalistic error can end up being amplified for manipulative purposes.

Taking all this into consideration, we also ask:

RQ2: Do sociodemographic factors, such as political leaning, age, gender, and interest in political news, influence the perception of disinformation?

2.3. Coping with Disinformation: The Role of Media Literacy

The phenomenon of disinformation requires a reconceptualization affecting news media practices, media policies, and how media literacy initiatives are designed and mapped out. In this sense, it is necessary to situate media literacy within an ample framework of actions, centring in the defence of the public sphere and common interests. According to Lazer et al. (2018), there are two categories of interventions that might be effective to stop the flow and influence of disinformation: "(i) those aimed at empowering individuals to evaluate the fake news they encounter, and (ii) structural changes aimed at preventing exposure of individuals to fake news in the first instance" (p. 1095).

Regarding the first category, there is an academic consensus that citizens are largely uninformed due to their inability to critically examine and evaluate information (Scheufele & Krause, 2019). Although it may also be due to the aforementioned reasons for cognitive biases, there is no doubt that a part of the disinformation problem is citizens' low level of media literacy. Starting from the basic definition that media literacy is "the ability of a citizen to access, analyze, and produce information for specific outcomes" (Aufderheide, 1993, p. 6), it can be concluded that it is the evaluation skill that poses the most relevant challenge for disinformation, as those with limited ability to evaluate "cannot distinguish dated, biased or exploitative sources" (Livingstone, 2004, p. 6). Therefore, as Klurfeld and Schneider (2014) point out, "the ultimate check against the spread of rumour, pernicious falsehood, disinformation, and unverified reports masquerading as fact" is a "generation of astutely educated news consumers" who can "identify for themselves fact-and-evidence-based news and information'' (p. 19).

In regards to the actions taken by citizens to combat disinformation, a review of the literature allows us to identify some recurring measures. In a qualitative study with 71 American users from different cities, Wagner and Boczkowski (2019) identify as main measures: drawing upon the experience and knowledge to assess news quality, triangulating sources, fact-checking, seeking for repetition of information across outlets, consumption of cross-ideological sources, and relying on certain personal contacts on social media who are perceived as good assessors of news quality. For their part, Tandoc, Ling, Westlund, Duffy, Goh, and Zheng (2018) utilize a survey of 2,501 Singaporeans to propose a conceptual framework to understand how individuals authenticate the information they encounter on social media. The results suggest that users rely on their own judgment of both the source and the message, and when this does not adequately provide a definitive answer, they turn to external resources to authenticate news items (from their social sphere or other institutional sources). On the other hand, fact-checking can be incidental and can simply

arise from the process of interacting with friends or consuming media.

The option of consulting a fact-checking website is not frequent among Portuguese university students, according to Figueira and Santos (2019). When in doubt about the veracity of a news item, the young people surveyed were more likely to check various sources, in addition to consulting trustworthy organizations. Finally, in a survey among Germans on 'fake news' perceptions, Reuter et al. (2019) indicate that even though about half of the respondents (48%) had noticed fake news, most participants report never having liked, shared, or commented on it.

Since user response is important in the spread and control of fake news on social media, finally we ask the following:

RQ3: What are the most common actions taken by Internet users to avoid being deceived by disinformation?

3. Method

3.1. Context

This work is part of a line of research by the authors that seeks to broaden the understanding of digital information consumption. Most empirical studies to date have a narrow geographical perspective, which limits the universality of inquiry in those topics, and suffer from a lack of contextualization (Rojas & Valenzuela, 2019). Thus, the geographical focus of our inquiry is centred in Iberoamerica, which encompasses Latin America and Spain (Fernández & Rodríguez-Virgili, 2019; Serrano-Puche et al., 2021; Serrano-Puche, Fernández, & Rodríguez-Virgili, 2018). A region where the academic literature on disinformation has grown in recent years (Guallar, Codina, Freixa, & Pérez-Montoro, 2020), but which still lacks empirical studies. Our analysis focuses on three Spanish speaking democracies with high Internet penetration rates, ranging from 77.5% of the population in Chile to 92.5% and 93.1% in Spain and Argentina, according to data from Internet World Stats (2019; Table 1).

Considering our objective to develop the understanding of disinformation in Iberoamerica, one thing to note is that countries such as Argentina, Chile, and Spain are experiencing symptoms of information disorder as do the other countries of the global North

(Valenzuela et. al., 2019; Wardle & Derakhshan, 2017). The Reuters Institute Digital News Report (Newman, Fletcher, Kalogeropoulos, & Nielsen, 2019) shows a generalized and steady decline in the average level of trust in the news and in traditional media worldwide. In addition, these three countries have experienced recent electoral or post-electoral contexts: Chile held presidential elections in 2017, Argentina in 2019, and Spain held two general elections in 2019. Objectives will be addressed in this comparative study using a most-similar systems approach (Meckstroth, 1975).

3.2. Sample

This work was developed from surveys carried out annually by the Reuters Institute Digital News Report (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018; Newman et. al., 2019), an international study on the consumption of digital information coordinated since 2012 by the University of Oxford, which currently includes 40 countries. For this specific study, surveys from the years 2018 and 2019 were consulted. Specifically, this study analysed user samples from Argentina (2018: n = 2,012; 2019: n = 2,006), Chile (2018: n = 2,008; 2019: n = 2,004), and Spain (2018: n = 2,023; 2019: n = 2,005).

The participants of the Digital News Report are adult Internet users who have consumed news in the last month, which serve as representatives of the connected population according to sociodemographic and geographic criteria. The data has been weighted according to official censuses and data accepted by the industry, and classified by age, sex, region, newspaper reading, and educational level, in order to better reflect the population of the three countries analysed. However, one should bear in mind that although these samples represent the digital population well, they do not necessarily represent the general population, and this can be considered a methodological limitation of the research.

3.3. Questionnaire and Variables

The online questionnaires cover a wide range of questions about news consumption, from which this research selected those that relate to disinformation. Questionnaires were not identical for both years. First, we used a question related to the different types of problematic information, from the 2018 survey:

Table 1. Penetration of the Internet in analysed countries.

Country	Population	Internet Users	Penetration Rate
Argentina	44,688,864	41,586,960	93.1%
Chile	18,197,209	14,108,392	77.5%
Spain	46,441,049	42,961,230	92.5%

Source: Internet World Stats (2019).

COGITATIO

To what extent, if at all, are you concerned about the following:

- Stories where facts are spun or twisted to push a particular agenda
- Stories that are completely made up for political or commercial reasons
- Poor journalism (factual mistakes, dumbed-down stories, misleading headlines/clickbait)
- The use of the term fake news (e.g., by politicians, others) to discredit news media they don't like
- Headlines that look like news stories but turn out to be advertisements
- Stories that are completely made up to make people laugh (satire)

From the 2019 survey we used a question about the different actions that users can take to protect against disinformation:

Have you done any of the following in the last year? Please select all that apply.

- I checked a number of different sources to see whether a news story was reported in the same way
- I decided not to share a news story because I was unsure about its accuracy
- I discussed a news story with a person I trust because I was unsure about its accuracy
- I stopped paying attention to news shared by someone because I am unsure whether I trust that person
- I stopped using certain news sources because I was unsure about the accuracy of their reporting
- I started relying more on sources of news that are considered more reputable

Considering our review of the literature and following the aforementioned research questions, the analysis of the answers takes into account sociodemographic variables such as age, gender, political leaning, access to, and interest in news.

3.4. Procedure

The fieldwork was carried out between the end of January and the beginning of February of both years, 2018 and 2019. It was conducted by the firm YouGov, which sent an invitation via email for users to complete an online poll.

3.5. Analyses

First, within each country, we identified what the participants' most prevalent concerns were regarding disinformation. Differences were tested with McNemar's test, after which Chi-squared tests were performed to identify differences according to country, gender, and political self-positioning. The association between the concerns

and age was tested with Spearman correlations. Finally, we created a composite measure of 'Concern about disinformation' based on the answers to the 6 items. A multivariate linear regression was then performed with this composite measure as a dependent variable, and different independent variables such as demographics and others related to news and politics. We performed the same analyses for the variables regarding actions against disinformation.

4. Results

In response to the aforementioned research objectives and questions, results are presented below under two main headings. First, we analyse which types of problematic information are of most concern for digital users in Argentina, Chile, and Spain (RQ1). The most common measures adopted by these users to combat disinformation (RQ3) are presented below. In both cases, and from a comparative perspective, the incidence of sociodemographic factors (RQ2) on the phenomenon of disinformation is examined.

4.1. Concerns about Disinformation

According to the Digital News Report of 2018, in general, the phenomenon of disinformation is a major concern among digital users in Argentina (60%), Chile (66%) and, to a greater extent, in Spain (69%), but the different manifestations of disinformation do not worry users to the same degree (Newman et. al., 2018). The questionnaire probes a range of six types of problematic information to which the interviewee assigns a greater or lesser level of concern. It should be noted that for the three countries, the final user ranking of problematic information is very similar (Table 2). The most worrying type of problematic information within the three countries analysed is "news that are created or altered in favor of a particular agenda" (77% of Spaniards, 74% of Argentinians, and 69% of Chileans). In the second spot, citizens placed "news that are completely invented, with political or commercial motivations" (in both Argentina and Spain, 73% of surveyed were concerned, with slightly lower numbers in Chile). In third place, participants selected errors that could be explained by mediocre journalism, such as inaccuracies, wrong or striking headlines that only seek to generate clicks (with 68% of the participants from these three countries showing concern). Likewise, for the three countries surveyed, the existence of parody news, or those just invented to make people laugh, occupy the last place in the ranking of concerns among those survey participants (with only 29–32% of participants showing concern for this type of media). News stories that can be considered as covert advertising, straying far from the news, are not reported to be of much concern either.

There does not seem to be a distinctive pattern related to concerns about problematic information when considering the gender of the interviewees (see

Table 2. Concern about disinformation by country.

	Argentina (N = 2012)	Chile (N = 2008)	Spain (N = 2023)	p *
Stories where facts are spun or twisted to push a particular agenda	74% [a]	69% [a]	77% [a]	< .001
Stories that are completely made up for political or commercial reasons	73% [a]	68% [a]	73% [b]	< .001
Poor journalism (factual mistakes, dumbed-down stories, misleading headlines/clickbait)	68% [b]	68% [a]	68% [c]	.745
The use of the term fake news (e.g., by politicians, others) to discredit news media they don't like	58% [b]	66% [d]	60% [c]	< .001
Headlines that look like news stories but turn out to be advertisements	43% [d]	43% [c]	53% [e]	< .001
Stories that are completely made up to make people laugh (satire)	29% [e]	33% [d]	34% [f]	.001

Notes: Question "To what extent, if at all, are you concerned about the following" (answers: Very + Extremely concerned). Within each country, different superscripts indicate a statistically significant (p < .05) difference in percentages, according to the McNemar's test ([a] > [b] > [c] > [d] > [e] > [f]). * p-value of the Chi-squared test for the inter-country comparison. Source: Adapted from the Reuters Institute Digital News Report survey 2018, conducted by YouGov.

Supplementary File, Table A). Instead, when analysing responses to concern about types of disinformation by age group, as shown in Table 3, findings suggest a pattern: in all cases, concern over disinformation increased with age.

When we analyse the concern regarding disinformation according to the reported ideological position (Table 4), as measured by the political self-positioning of the interviewees, we find an interesting pattern in the three countries analysed: those who see themselves as 'left-leaning' tend to worry more about disinformation. This pattern is evident in the three countries analysed and when analysing the forms of disinformation that were considered most worrying.

4.2. Actions against Disinformation

There is a similar pattern in the three countries analysed in relation to the actions they take to protect themselves against disinformation (Table 5). The checking of different sources and how they contrast with each other is the main measure that Argentinian (60%), Chilean (63%), and Spanish (56%) respondents take against disinformation. In second place, participants attempt to avoid viralizing news that is not entirely reliable (52%, 50%, and 40%, respectively).

However, although the relative importance attached to the different measures of disinformation is the same, the intensity of the reaction is not equal. There is an important gender gap in the actions taken when faced with problematic information (Supplementary File, Table B).

Although a clear pattern was detected in the three countries (Table 3) regarding the direct relationship

between age and concern about 'problematic information,' the same does not occur with reported actions to combat disinformation (Table 6). Only in a few cases do we find a significant linear trend. A clear trend was found in Spain, where younger participants reported more frequently discussing news with a person they trust. Another clear trend was found in Chile, where older participants reported a higher incidence of checking different sources and of stopping paying attention to news shared by someone because of being unsure of that person's reliability.

Regarding the relationship between ideological self-positioning and the different measures to protect against disinformation, divergences are observed (Table 7). In Argentina and Chile, those who declare themselves as being on the left of the political spectrum stop consulting some media when they feel the medium has failed them (49% and 47%, respectively) more often than centrists (39% and 38%) and right-wingers (37% and 38%). Moreover, in Argentina, left-wingers discussed news with a person they trust (44%) more often than centrists (41%) and right-wingers (32%).

4.3. Predicting Concern and Actions

The composite measure of 'Concern about disinformation' showed a good internal consistency (Cronbach alpha = .84). In the multiple regression to predict this measure (see Table 8), concern about disinformation was higher among women, older participants, those with high interest in political news, left-wingers, and Spaniards.

Regarding actions against disinformation, the internal consistency of the composite measure was lower (Cronbach alpha = .56), probably due to the fact that these variables were dichotomous. According to the

Table 3. Concern about disinformation by age.

	Argentina						Chile						Spain					
	18–24 (N = 300)	25–34 (N = 418)	35–44 (N = 398)	45–54 (N = 302)	55+ (N = 594)	p *	18–24 (N = 281)	25–34 (N = 416)	35–44 (N = 357)	45–54 (N = 361)	55+ (N = 592)	p *	18–24 (N = 166)	25–34 (N = 293)	35–44 (N = 413)	45–54 (N = 399)	55+ (N = 753)	p *
Stories where facts are spun or twisted to push a particular agenda	66%	68%	71%	76%	82%	< .001	61%	58%	65%	70%	82%	< .001	74%	68%	72%	79%	83%	< .001
Stories that are completely made up for political or commercial reasons	66%	68%	71%	76%	82%	< .001	58%	61%	64%	69%	78%	< .001	68%	61%	68%	75%	80%	< .001
Poor journalism (factual mistakes, dumbed-down stories, misleading head-lines/clickbait)	60%	64%	66%	68%	76%	< .001	63%	62%	65%	68%	78%	< .001	62%	60%	62%	72%	75%	< .001
The use of the term fake news (e.g., by politicians, others) to discredit news media they don't like	54%	56%	54%	59%	68%	< .001	51%	53%	53%	58%	69%	< .001	64%	58%	59%	71%	71%	< .001
Headlines that look like news stories but turn out to be advertisements	36%	41%	39%	39%	54%	< .001	36%	39%	37%	43%	53%	< .001	54%	47%	46%	52%	59%	< .001
Stories that are completely made up to make people laugh (satire)	31%	25%	26%	28%	34%	.046	29%	29%	30%	30%	41%	< .001	31%	30%	33%	34%	36%	.033

Notes: Question "To what extent, if at all, are you concerned about the following" (answers: Very + Extremely concerned). * p-value of the Spearman correlation between each type of concern and age (within each country). Source: Adapted from the Reuters Institute Digital News Report survey 2018, conducted by YouGov.

Table 4. Concern about disinformation due to political position.

	Argentina				Chile				Spain			
	Left (N = 204)	Centre (N = 1138)	Right (N = 201)	p *	Left (N = 313)	Centre (N = 934)	Right (N = 285)	p *	Left (N = 585)	Centre (N = 1108)	Right (N = 143)	p *
Stories where facts are spun or twisted to push a particular agenda	80%	78%	72%	.095	76%	75%	68%	.039	80%	78%	75%	.381
Stories that are completely made up for political or commercial reasons	82%	77%	66%	< .001	76%	73%	65%	.006	75%	73%	74%	.440
Poor journalism (factual mistakes, dumbed-down stories, misleading headlines/clickbait)	73%	72%	70%	.634	75%	74%	70%	.277	70%	68%	71%	.569
The use of the term fake news (e.g., by politicians, others) to discredit news media they don't like	63%	62%	60%	.687	67%	64%	53%	< .001	72%	65%	64%	.009
Headlines that look like news stories but turn out to be advertisements	52%	45%	42%	0.69	46%	46%	45%	.935	56%	52%	50%	.290
Stories that are completely made up to make people laugh (satire)	30%	28%	30%	.506	31%	34%	35%	.711	32%	35%	42%	.075

Notes: Question "To what extent, if at all, are you concerned about the following" (answers: Very + Extremely concerned). * p-value of the Chi-squared test. Source: Adapted from the Reuters Institute Digital News Report survey 2018, conducted by YouGov.

Table 5. Actions against disinformation by country.

	Argentina (N = 2006)	Chile (N = 2004)	Spain (N = 2005)	p *
I checked a number of different sources to see whether a news story was reported in the same way	60% [a]	63% [a]	56% [a]	< .001
I decided not to share a news story because I was unsure about its accuracy	52% [b]	50% [b]	40% [b]	< .001
I discussed a news story with a person I trust because I was unsure about its accuracy	37% [d]	40% [c]	37% [b]	.090
I stopped paying attention to news shared by someone because I am unsure whether I trust that person	39% [c]	40% [c]	28% [c]	< .001
I stopped using certain news sources because I was unsure about the accuracy of their reporting	35% [d]	36% [d]	29% [c]	< .001
I started relying more on sources of news that are considered more reputable	27% [e]	27% [e]	20% [d]	< .001

Notes: Question "Have you done any of the following in the last year? Please select all that apply." Within each country, different super-scripts indicate a statistically significant (p < .05) difference in percentages, according to the McNemar's test ([a] > [b] > [c] > [d] > [e]). * p-value of the Chi-squared test. for the inter-country comparison. Source: Adapted from the Reuters Institute Digital News Report survey 2019, conducted by YouGov.

regression, a greater number of actions were carried out by participants with post-secondary education, by those who access news more frequently, by those who have greater interest in political news, and by Argentinians and Chileans.

5. Discussion and Conclusion

In the previous sections, we comparatively analysed the perceptions of disinformation by digital users in three Spanish-speaking countries.

After analysing the data, and in response to the first research question (RQ1), Internet users from Argentina, Chile, and Spain agree both on the types of problematic information that they consider to be of most concern and on their ranking, considering the use of news to push particular agendas as the most worrying. These are closely followed by stories that are completely made up for political or commercial reasons, and thirdly, by instances of poor journalism such as factual mistakes and misleading headlines. Regardless, people see the difference between fake news and news as one of degree rather than a clear distinction, following research by Nielsen and Graves (2017) who conducted focus groups on different countries combined with Reuters Digital News Reports. Thus, from an audience perspective, these modalities of problematic information seem to be part of a broader discontent with the news industry, as well as platform companies and politicians.

Given these remarkable similarities in general perceptions of problematic information, the differences found among the analysed sociodemographic segments are even more relevant (RQ2). A general index of 'Concern about disinformation' showed that it is higher among women, older people, those with high interest in political news, left-wingers, and Spaniards.

In response to RQ3, participants in the three countries also react similarly regarding the actions taken to protect themselves against disinformation. They indicate the same order of measures to combat this phenomenon, and with similar frequency in Argentina, Chile, and Spain—although to a lesser extent in Spain. The checking of different sources is the main measure that respondents from the three countries take against disinformation, followed by avoiding viralizing news that is not entirely reliable and discussing news stories with interpersonal contacts when its accuracy is unclear. The importance given to fact-checking partially matches previous research (Wagner & Boczkowski, 2019). It works as an 'external act of authentication' (Tandoc, Ling, et al., 2018) complementing other measures such as the consulting of trusted peers. It is also important to note that surveyees gave importance to not being a propagator of fake news. Even if this is due to a possible social desirability bias, it is relevant as this constitutes one of the basic actions of media literacy against disinformation.

Considering how the variables of age, gender, and political position influence perceptions and actions to combat disinformation (RQ2), we pointed out that the concern about disinformation increases as the participants age. This is a finding that concurs with other research (Guess et al., 2019). As might be expected, those who are more interested in political news are also more concerned about disinformation (Chadwick & Vaccari, 2019). Regarding gender, there are few differences (as can be seen in more detail in Supplementary File, Tables A and B). Men and women share similar concerns about disinformation and ranked them in the same

Table 6. Actions against disinformation by age.

	Argentina						Chile						Spain					
	18–24 (N = 297)	25–34 (N = 413)	35–44 (N = 397)	45–54 (N = 305)	55+ (N = 594)	p *	18–24 (N = 271)	25–34 (N = 413)	35–44 (N = 357)	45–54 (N = 355)	55+ (N = 609)	p *	18–24 (N = 164)	25–34 (N = 283)	35–44 (N = 399)	45–54 (N = 399)	55+ (N = 760)	p*
I checked a number of different sources to see whether a news story was reported in the same way	57%	59%	61%	64%	60%	.337	58%	60%	60%	67%	68%	< .001	55%	52%	56%	57%	57%	1.83
I decided not to share a news story because I was unsure about its accuracy	49%	52%	54%	51%	54%	.282	51%	53%	47%	49%	49%	.310	37%	36%	45%	40%	39%	.923
I discussed a news story with a person I trust because I was unsure about its accuracy	39%	40%	36%	30%	36%	.080	44%	40%	32%	43%	40%	.956	42%	41%	35%	37%	35%	.014
I stopped paying attention to news shared by someone because I am unsure whether I trust that person	34%	39%	39%	40%	41%	.084	35%	34%	38%	44%	45%	< .001	34%	24%	24%	27%	30%	.500
I stopped using certain news sources because I was unsure about the accuracy of their reporting	36%	37%	33%	35%	35%	.737	33%	35%	34%	36%	38%	.145	26%	30%	27%	29%	30%	.407
I started relying more on sources of news that are considered more reputable	26%	26%	24%	22%	31%	.073	31%	24%	25%	23%	30%	.270	20%	24%	19%	18%	19%	.481

Notes: Question "Have you done any of the following in the last year? Please select all that apply." * p-value of the Spearman correlation between each type of action and age (within each country). Source: Adapted from the Reuters Institute Digital News Report survey 2019, conducted by YouGov.

Table 7. Actions against disinformation by political position.

	Argentina				Chile				Spain			
	Left (N = 179)	Centre (N = 897)	Right (N = 177)	p*	Left (N = 267)	Centre (N = 787)	Right (N = 248)	p*	Left (N = 547)	Centre (N = 867)	Right (N = 218)	p*
I checked a number of different sources to see whether a news story was reported in the same way	68%	66%	62%	.641	72%	70%	63%	.118	58%	58%	55%	.461
I decided not to share a news story because I was unsure about its accuracy	56%	58%	53%	.558	54%	55%	55%	.965	41%	43%	37%	.364
I discussed a news story with a person I trust because I was unsure about its accuracy	44%	41%	32%	.025	44%	42%	43%	.920	38%	39%	36%	.781
I stopped paying attention to news shared by someone because I am unsure whether I trust that person	43%	44%	37%	.203	49%	44%	44%	.276	29%	30%	29%	.859
I stopped using certain news sources because I was unsure about the accuracy of their reporting	49%	39%	37%	0.14	47%	38%	38%	.035	33%	30%	34%	.226
I started relying more on sources of news that are considered more reputable	30%	30%	38%	.083	25%	32%	33%	.064	18%	22%	22%	.098

Notes: Question "Have you done any of the following in the last year? Please select all that apply." * p-value of the Chi-squared test. Source: Adapted from the Reuters Institute Digital News Report survey 2019, conducted by YouGov.

Table 8. Predictors associated with concern about misinformation and with actions against disinformation.

	Concern about misinformation [a]				Actions against disinformation [b]			
	N	Mean (SD)	p [c]	B (95% CI) [d]	N	Mean (SD)	p [c]	B (95% CI) [d]
Gender								
Male	2362	21.88 (4.98)	.004	(ref)	2160	2.58 (1.63)	.745	(ref)
Female	2250	22.30 (4.83)		0.59 (0.31 to 0.87)	2052	2.60 (1.53)		0.03 (−0.06 to 0.13)
Age								
18–44	2191	21.44 (4.84)	< .001	(ref)	1978	2.59 (1.54)	.956	(ref)
45+	2421	22.67 (4.89)		1.00 (0.72 to 1.28)	2234	2.59 (1.62)		0.00 (−0.09 to 0.10)
Post-secondary education								
No	1358	22.00 (5.08)	.436	(ref)	2387	2.50 (1.59)	< .001	(ref)
Yes	3254	22.12 (4.84)		−0.05 (−0.35 to 0.26)	1825	2.70 (1.56)		0.17 (0.08 to 0.27)
Access to news								
Once per day or less	1699	21.71 (5.03)	< .001	(ref)	1580	2.28 (1.55)	< .001	(ref)
More than once per day	2913	22.31 (4.82)		0.27 (−0.03 to 0.56)	2632	2.77 (1.57)		0.38 (0.29 to 0.48)
Interest in political news								
Little	1885	20.93 (5.21)	< .001	(ref)	2187	2.40 (1.55)	< .001	(ref)
Much	2727	22.88 (4.52)		1.75 (1.45 to 2.05)	2025	2.79 (1.60)		0.34 (0.24 to 0.44)
Political self-position								
Left	1032	22.54 (4.90)	< .001	(ref)	984	2.53 (1.58)	.241	(ref)
Centre	2992	22.05 (4.80)		−0.27 (−0.61 to 0.08)	2589	2.62 (1.57)		−0.01 (−0.12 to 0.11)
Right	588	21.47 (5.40)		−0.61 (−1.10 to −0.11)	639	2.54 (1.64)		−0.11 (−0.26 to 0.04)
Country								
Argentina	1447	21.92 (4.88)	< .001	(ref)	1259	2.79 (1.57)	< .001	(ref)
Chile	1387	21.61 (5.41)		−0.03 (−0.39 to 0.33)	1297	2.83 (1.62)		0.07 (−0.05 to 0.19)
Spain	1778	22.60 (4.46)		0.44 (0.10 to 0.78)	1656	2.24 (1.50)		−0.57 (−0.68 to −0.45)

Notes: [a] Index (range 6–30) obtained from the sum of 6 questions regarding the extent to which participants were concerned about different types of disinformation. [b] Index (range 0–6) obtained from the sum of 6 questions regarding whether participants had done any of different actions against disinformation. [c] p-value for the Student's t-test in dichotomous variables ('Age' through 'Interest in political news') or for the ANOVA in trichotomous variables ('Political self-position' and 'Country'). [d] B coefficient (and 95% confidence intervals) of the multiple linear regression of each variable, adjusted for all variables in the first column. Adjusted R-squared = .06 (for the model predicting 'Concern about misinformation') and .07 (for the model predicting 'Actions against disinformation').

order of importance. This also seems to agree with previous studies (Wolverton & Stevens, 2019), which found no significant differences in this regard. However, and despite this relative uniformity, there are different reactions against disinformation. Men and women reported taking different measures to protect themselves from misinformation, although they are concerned about the same elements. While men rely more on the reputation of the source as a preventive action, women appear to be more active in taking steps to protect themselves against fake news, checking different sources to compare, and/or avoiding sharing a news story if they are not sure of its adequacy. In other words, men and women agree more in the diagnosis of the problem, than in the path followed to solve it. Following Giglietto et al. (2019), it could be said that when women receive disinformation, they avoid becoming propagators of it to a greater degree than men.

Finally, regarding ideological self-positioning as a variable for interpreting the phenomenon, we see that those who define themselves as 'right-wing' or 'centre' tend to worry slightly less about the problems of disinformation than those that claim to be 'left-wing.' This was clearly observed in the Argentinian and Chilean cases. This is not a finding against what previous research has found regarding US voters, as one might think on first impulse: Those studies that analysed the 2016 US presidential election and found that American right-wing voters were more vulnerable to disinformation than the left-wing voters, were focused on content and reception (Allcott & Gentzkow, 2017; Grinberg et al., 2019; Mckay & Tenove, 2020). Instead, our findings are based on users´ perceptions. They show that those who see themselves as leftists are more aware of the problem of disinformation, and that circumstances might make them better protected against the phenomena.

There were remarkable similarities in how users handle problematic information in the three countries analysed. Hasty conclusions should be avoided as they could lead to generalizations about the phenomenon of disinformation in Iberoamerica; it would be worth conducting a more in-depth investigation into the differences by country and their reasons, considering previous studies. For example, in the Venezuelan case (Serrano-Puche et al., 2021), a different hierarchization and intensity was found in the modalities of 'problematic information,' where among other differences, poor journalism was users' main concern. In short, it is important to continue researching the geographic and demographic variables that best allow the identification of information vulnerability, since the clear identification of the groups most vulnerable to disinformation could suggest effective measures to mitigate its negative effects.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Aufderheide, P. (1993). *Media literacy: A report of the national leadership conference on media literacy*. Washington, DC: The Aspen Institute.

Bail, C. A., Guay, B., Maloney, E., Combs, A., Hillygus, D. S., Merhout, F., . . . & Volfovsky, A. (2020). Assessing the Russian Internet Research Agency's impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proceedings of the National Academy of Sciences*, *117*(1), 243–250.

Chadwick, A., & Vaccari, C. (2019). *News sharing on UK social media: Misinformation, disinformation, and correction*. Loughborough: Loughborough University—Online Civic Culture Centre.

Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, *43*(2), 97–116. https://doi.org/10.1080/23808985.2019.1602782

Fernández, C. B., & Rodríguez-Virgili, J. (2019). Electors are from Facebook, political geeks are from Twitter: Political information consumption in Argentina, Spain and Venezuela. *KOME*, *7*(1), 42–62. https://doi.org/10.17646/KOME.75698.62

Figueira, J., & Santos, S. (2019). Percepción de las noticias falsas en universitarios de Portugal: Análisis de su consumo y actitudes [Perceptions about fake news in Portuguese higher education students: An analysis of consumption and attitudes]. *El Profesional de la Información*, *28*(3).

Freelon, D., Bossetta, M., Wells, C., Lukito, J., Xia, Y., & Adams, K. (2020). Black trolls matter: Racial and ideological asymmetries in social media disinformation. *Social Science Computer Review*. Advance online publication. https://doi.org/10.1177/0894439320914853

Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, *37*(2), 145–156. https://doi.org/10.1080/10584609.

2020.1723755

Giglietto, F., Iannelli, L., Valeriani, A., & Rossi, L. (2019). 'Fake news' is the invention of a liar: How false information circulates within the hybrid news system. *Current Sociology*, *67*(4), 625–642. https://doi.org/10.1177/0011392119837536

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378. https://doi.org/10.1126/science.aau2706

Guallar, J., Codina, L., Freixa, P., & Pérez-Montoro, M. (2020). Desinformación, bulos, curación y verificación: Revisión de estudios en iberoamérica 2017–2020 [Misinformation, hoaxes, curation and verification: Review of studies in Iberomerica 2017–2020]. *Telos: Revista de Estudios Interdisciplinarios en Ciencias Sociales*, *22*(3), 595–613. www.doi.org/10.36390/telos223.09

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

Ha, L., Perez, L. A., & Ray, R. (2019). Mapping recent development in scholarship on fake news and misinformation, 2008 to 2017: Disciplinary contribution, topics, and impact. *American Behavioral Scientist*, *65*(2), 290–315. https://doi.org/10.1177/0002764219869402

High Level Expert Group on Fake News and Disinformation. (2018). *A multi-dimensional approach to disinformation: Report of the independent high level group on fake news and online disinformation*. Brussels: European Commission. Retrieved from https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation

Hjorth, F., & Adler-Nissen, R. (2019). Ideological asymmetry in the reach of pro-Russian digital disinformation to United States audiences. *Journal of Communication*, *69*(2), 168–192. https://doi.org/10.1093/joc/jqz006

Howard, P. N., Ganesh, B., Liotsiou, D., Kelly, J., & Francois, C. (2018). *The IRA, social media and political polarization in the United States, 2012–2018*. Oxford: University of Oxford.

Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, *25*(3), 493–516. https://doi.org/10.1177/1940161219900126

Internet World Stats. (2020). World Internet users and population stats. *Internet World Stats*. Retrieved from www.internetworldstats.com

Iyengar, S., & Hahn, K. S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, *59*(1), 19–39.

Jack, C. (2017). *Lexicon of lies: Terms for problematic information*. New York, NY: Data & Society Research Institute. Retrieved from https://datasociety.net/library/lexicon-of-lies

Khaldarova, I., & Pantti, M. (2016). Fake news. *Journalism Practice*, *10*(7), 891–901. https://doi.org/10.1080/17512786.2016.1163237

Klurfeld, J., & Schneider, H. (2014). *News literacy: Teaching the internet generation to make reliable information choices*. Washington, DC: The Brookings Institution. Retrieved from https://www.brookings.edu/research/news-literacy-teaching-the-internet-generation-to-make-reliable-information-choices

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., . . . & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096.

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369.

Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *The Communication Review*, *7*(1), 3–14. https://doi.org/10.1080/10714420490280152

Lord, C. G., & Taylor, C. (2009). Biased assimilation: Effects of assumptions and expectations on the interpretation of new evidence. *Social and Personality Psychology Compass*, *3*(5), 827–841.

McKay, S., & Tenove, C. (2020). Disinformation as a threat to deliberative democracy. *Political Research Quarterly*. Advance online publication. https://doi.org/10.1177/1065912920938143

Meckstroth, T. W. (1975). Most different systems and most similar systems, a study in the logic of comparative inquiry. *Comparative Political Studies*, *8*(2), 132–157.

Miller, M. L., & Vaccari, C. (2020). Digital threats to democracy: Comparative lessons and possible remedies. *The International Journal of Press/Politics*, *25*(3), 333–356. https://doi.org/10.1177/1940161220922323

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2019). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*, *65*(2), 180–212. https://doi.org/10.1177/0002764219878224

Morais, N. S., & Cruz, M. (2020). Gender perception about fake news and disinformation: Case study with Portuguese higher education students. In L. Gómez Chova, A. López Martínez, & I. Candel Torres (Eds.), *EDULEARN20 Proceedings: 12th international conference on Education and New Learning Technologies* (pp. 7746–7753). Valencia: International Academy of Technology, Education and Development. https://doi.org/10.21125/edulearn.2020.1957

Nielsen, R. K., & Graves, L. (2017). *"News you don't*

believe": Audience perspectives on fake news. Oxford: Reuters Institute for the Study of Journalism.

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A., & Nielsen, R. (2018). *Reuters Institute digital news report 2018*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf

Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. (2019). *Reuters Institute digital news report 2019*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR_2019_FINAL.pdf

Reuter, C. A., Hartwig, K., Kirchner, J., & Schlegel, N. (2019). Fake news perception in Germany: A representative study of people's attitudes and approaches to counteract disinformation. In V. Pipek & T. Ludwig (Eds.), *Proceedings of the International Conference on Wirtschaftsinformatik* (pp. 1069–1083). Siegen: Ais.

Rojas, H., & Valenzuela, S. (2019). A call to contextualize public opinion-based research in political communication. *Political Communication*, *36*(4), 652–659. https://doi.org/10.1080/10584609.2019.1670897

Scheufele, D., & Krause, N. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, *116*(16), 7662–7669.

Serrano-Puche, J., Fernández, C. B., & Rodríguez-Virgili, J. (2018). Political information and incidental exposure in social media: The cases of Argentina, Chile, Spain and Mexico. *Doxa Comunicación*, *27*, 19–42. https://doi.org/10.31921/doxacom.n27a1

Serrano-Puche, J., Fernández, C. B., & Rodríguez-Virgili, J. (2021). Disinformation and news consumption in a polarized society: An analysis of the case of Venezuela. In D. Ramírez Plascencia, B. Carvalho Gurgel, & A. Plaw (Eds.), *The politics of technology in Latin America volume 2: Digital media, daily life and public engagement* (pp. 161–178). London: Routledge.

Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2018). Defining "fake news": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tandoc, E. C., Jr., Ling, R., Westlund, O., Duffy, A., Goh, D., & Zheng Wei, L. (2018). Audiences' acts of authentication in the age of fake news: A conceptual framework. *New Media & Society*, *20*(8), 2745–2763. https://doi.org/10.1177/1461444817731756

Valenzuela, S., Halpern, D. l., Katz, J. E., & Miranda, J. P. (2019). The paradox of participation versus misinformation: Social media, political engagement, and the spread of misinformation. *Digital Journalism*, *7*(6), 802–823. https://doi.org/10.1080/21670811.2019.1623701

Wagner, M. C., & Boczkowski, P. B. (2019). The reception of fake news: The interpretations and practices that shape the consumption of perceived misinformation. *Digital Journalism*, *7*(7), 870–885.

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Strasbourg: Council of Europe. Retrieved from https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

Wolverton, C., & Stevens, D. (2019). The impact of personality in recognizing disinformation. *Online Information Review*, *44*(1), 181–191. https://doi.org/10.1108/OIR-04-2019-0115

About the Authors

Jordi Rodríguez-Virgili is Associate Professor of Political Communication at the University of Navarra, where he serves as Associate Dean of Teaching Staff at the School of Communication. He is a researcher at the Center for Internet Studies and Digital Life at the UNAV. He has been Visiting International Scholar at The Graduate School of Political Management (George Washington University). Currently, his main inquiry interests are focused on media consumption, political communication, and disinformation.

Javier Serrano-Puche is Associate Professor and Vice Dean for Academic Affairs at the School of Communication of the University of Navarra. He serves as a research fellow for the Center for Internet Studies and Digital Life at the same university and has performed extensive research on news consumption, media literacy and the relations between media and emotions. His recent publications include papers in *Communications*, *International Journal of Communication*, or *Comunicar*, and in publishing houses as Routledge, Palgrave MacMillan, and Springer.

Carmen Beatriz Fernández is Guest Professor and Doctoral Candidate of the School of Communication at the University of Navarra, and Researcher at the Center for Internet Studies and Digital Life. She earned a MA in Business Administration (IESA, Venezuela) and a MA in Political Science, concentration on Electoral Campaigning (University of Florida, USA). Mrs Fernandez is CEO of DataStrategia consulting, and Co-Founder of the Organization of Latin American Political Consultants OCPLA.

Article

Journalism Students and Information Consumption in the Era of Fake News

Santiago Tejedor *, Marta Portalés-Oliva, Ricardo Carniel-Bugs and Laura Cervi

Department of Journalism and Communication Sciences, Autonomous University of Barcelona, 08193 Barcelona, Spain; E-Mails: santiago.tejedor@uab.cat (S.T.), marta.portales@e-campus.uab.cat (M.P.-O.), ricardo.carniel@uab.cat (R.C.-B.), laura.cervi@uab.cat (L.C.)

* Corresponding author

Abstract
Technological platforms, such as social media, are disrupting traditional journalism, as a result the access to high-quality information by citizens is facing important challenges, among which, disinformation and the spread of fake news are the most relevant one. This study approaches how journalism students perceive and assess this phenomenon. The descriptive and exploratory research is based on a hybrid methodology: Two matrix surveys of students and a focus group of professors (n = 6), experts in Multimedia Journalism. The first survey (n = 252), focused on students' perception of fake news, the second (n = 300) aims at finding out the type of content they had received during the recent confinement caused by the Covid-19 pandemic. Results show that most of the students prefer online media as a primary source of information instead of social media. Students consider that politics is the main topic of fake news, which, according to the respondents, are mainly distributed by adult users through social networks. The vast majority believe that fake news are created for political interests and a quarter of the sample considers that there is a strong ideological component behind disinformation strategies. Nonetheless, the study also reveals that students do not trust in their ability to distinguish between truthful and false information. For this reason, this research concludes, among other aspects, that the promotion of initiatives and research to promote media literacy and news literacy are decisive in the training of university students.

1. Introduction

In recent years, fake news and misinformation have become a recurrent object of study both because of their rapid growth and because of the problems and threats, they generate. Since the emergence of this concept (Love, 2007; Tally, 2011), the first works have focused on political phenomena (Blanco Alfonso, García Galera, & Tejedor Calvo, 2019) and have rapidly given way to research on the impact of fake news in other areas of

our daily lives, such as education, entertainment, health and journalism, among others.

The popularization of fake news has taken place in a very particular informative and communicative situation, characterized by digital noise, 'infoxication' and information disorder (Wardle & Derakhshan, 2017). Infoxication is a term popularized by Alvin Toffler (1970) that describes nowadays information overload derived from our constant connection to Internet. In this context, two very particular phenomena have occurred. On

the one hand, the rapid distribution of this type of message (Jang & Kim, 2018) and, on the other hand, the accelerated generation of a wide variety of distorted content (Ireland, 2018; Southwell, Thorson, & Sheble, 2018). The studies by Fletcher, Cornia, Graves, and Rasmus (2018) and Musgrove, Powers, Rebar, and Musgrove (2018) emphasize the capacity and easiness by online platforms to spread fake news in comparison with conventional media. Thus, the so-called information society has given way to a society of infoxication or digital noise where fake news are disseminated 10% more than real news (Vosoughi, Roy, & Aral, 2018).

Generation Z, composed of individuals born between the years 1996 and 2010 (Dimock, 2019), is the first generation that has never known a world without the Internet. Their lives are moulded by the Internet, which has been converted in a natural part of their lives. Nonetheless, numerous studies (Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010; Wineburg & McGrew, 2016) have highlighted that they are the most vulnerable to fake news.

Among today's Generation Zers, we find the journalists of tomorrow. Journalism and communication students represent a category of special interest, since besides belonging to the broader category of Generation Zers, soon will be in charge of taking up the responsibilities involved in the task of being a professional journalist and/or communicator.

Despite the interest that this subgroup of young people should arise, very few studies (Bhaskaran, Mishra, & Nair, 2019) have focused on journalism students' understanding and perception of 'fake news.'

Herrero-Diz, Conde-Jiménez, Tapia-Frade, and Varona-Aramburu (2019), analyzing students of Communication from the Spanish region of Andalusia, conclude that university students have difficulty differentiating the veracity of the sources. Another study based in the Basque country (Mendiguren, Pérez Dasilva, & Meso Ayerdi, 2020) observe that students are mainly informed in the online environment and a high percentage confesses to having fallen into the trap of fake news at some point.

Besides the mentioned studies, fake news and the way journalism students react to them are still an understudied topic, specifically, as Blanco Alfonso et al. (2019) point out, in the Ibero-American context. Ibero-America should be intended as a 'space' that goes beyond a mere geo-linguistic area, tracing cultural, socio-political and socioeconomic relationships, within it (Tejedor, Ventín, Cervi, Pulido, & Tusa, 2020). Along the same lines, these authors have stressed the need to deepen on the knowledge of this field in the area. Based on this, the present study presents a diagnostic analysis of the consumption of information by university students in the field of communication.

Accordingly, our descriptive and exploratory research aims at understanding the informative habits of journalism students and their position towards fake news, answering the two following research questions:

RQ1: How do young journalism students inform themselves.

RQ2: How do they position themselves towards fake news?

2. Theoretical Framework

2.1. Fake News

'Fake news' in the media is not a new phenomenon. On the one hand, there is no consensus on the origin of it. Some scholars consider that disinformation started with the earliest writing systems (Tandoc, Lim, & Ling, 2017), others (Posetti & Matthews, 2018) recall that early record dates back to ancient Rome, when Octavian launched a smear campaign against Antony made of short slogans written upon coins, comparing it to a sort of archaic Tweets. Most observers (Molina, Sundar, Le, & Lee, 2019), however, trace it back to World War II, specifically to the Russian word *dezinformatsiya*, used by Soviet planners in the 1950s to describe the dissemination (in the press, on the radio, etc.) of false reports intended to mislead public opinion.

On the other hand, there is a wide consensus that whereas the use of disinformation is not new, the digital revolution has greatly enhanced public vulnerability. In particular, in an informative scenario dominated by the emergence of content designed for rapid viralization (Romero-Rodríguez, de-Casas-Moreno, & Torres-Toukoumidis, 2016), the risk of disinformation increases. In other words, what is new is the speed, scale and massive proliferation and consumption of false information, in the current context of the destabilization of the mainstream media (Cervi, 2019) and information disorder (Wardle & Derakhshan, 2017).

In particular, as pointed out by Giglietto, Righetti, Rossi, and Marino (2020), two recent events—2016 Presidential elections in the US and Brexit referendum in UK—showed how the antagonist online participatory practices of sharing, collaborating and organizing collective actions, considered the prerogative of democratizing forces fighting established powers, proved to be just as effective in supporting the spread of extremisms, hate speech, violence and fake news.

Since then, 'fake news' has become a buzzword (Tandoc et al., 2017), thus, like other buzzwords, semantically confusing. Although since then there has been an explosion of both academic and journalistic work on this topic, defining 'fake news' is not an easy task.

Tandoc et al. (2017), analysing how the term has been used by scholars, reveal up to six types of definition: news satire, news parody, fabrication, manipulation, advertising, and propaganda. Earlier studies, in fact, have applied the term to define related but distinct types of content, such as news parodies or political satires (Blanco Alfonso et al., 2019), or to define information that adopted conventional news formats to

make satirical commentary, as in the case of late-night TV shows. Other approaches pointed at tabloid journalism, that walked a fine line between reporting reality and making wild claims (Molina et al., 2019) and news propaganda (Pérez-Tornero, Tayie, Tejedor, & Pulido, 2018). Currently the term is mostly used to generically describe false stories spreading on social media.

Acknowledging the complexities and ambiguity of the term, and stressing out the need for further definitory work, we understand 'fake news' within the broader phenomenon of disinformation, as a deliberate effort to mislead, deceive, or confuse an audience in order to promote their personal, religious, or ideological objectives (Cervi & Carrillo-Andrade, 2019; Fetzer, 2004; Pérez-Tornero et al., 2018; Thompson, 2016; Turkle, 2015).

2.2. Fake News, the Audience and Young People

As previously mentioned, since 2016 there has been an explosion of academic work that fixes its subject matter using the terms 'fake news,' this topic of fake news has become one of the most talked about during Covid-19 lockdown, to the extent that the World Health Organization warned of the risk of the growth of the generation and circulation of this type of content.

This topic has received attention in a variety of fields, with scholars investigating the antecedents, characteristics, and consequences of its creation and dissemination. It is possible to divide these studies into two macro branches: Those who focus on the supply side, that is to say, primarily interested in the nature and construction of fake news, and those interested in the consumption side, geared at understanding why people consume, share and/or believe 'fake news.'

A recent research by Pennycook, Cannon, and Rand (2018), for example, warns of an 'illusory truth effect' linked to fake news headlines. The authors argue that social networks enhance the attraction and belief in this type of false content, while pointing out that the categorisation of this content as unreliable is not an effective solution to the problem of fake news. In this sense, they mention examples of implausible contents and stories labelled as controversial that reached important credibility rates among Internet users. Accordingly, the works of Marcom, Murdoch, and Caulfield (2017) and those of Peters, Tartari, Lotfinejad, Parneix, and Pittet (2018) and Guess, Nagler, and Tucker (2019) in the field of health, warn that in the current scenario, marked by fake news, even a correctly conducted investigation could be distorted to make people believe something is false.

Most of the works concentrating on the user side deal with how people assess the messages they receive and how they establish criteria of credibility. In this vein, Tandoc et al. (2018) discover that people rely on both their own judgment of the source and the message, and when this does not adequately provide a definitive answer, they turn to external resources to authenticate news items. Along the same lines, Samuel-Azran and Hayat (2019) have pointed out that the social link between the user who distributes the content and the users who receive it impacts on the effect of the perception of credibility with respect to the news source, as well as on the credibility conferred on the message received. In another study, Pennycook and Rand (2019) claim that users resort to analytical thinking to assess the plausibility of potential fake news headlines. Thus, the authors conclude that the vulnerability of citizens to fake news is more a result of the inertia of lazy thinking than of partisan bias.

As previously mentioned, numerous studies (Hargittai et al., 2010; Wineburg & McGrew, 2016) have highlighted that young people, especially Generation Zers, are the most vulnerable to fake news.

Although there is no absolute consensus about the precise boundaries of Generation Z, most literature (Dimock, 2019) considers that it is composed of individuals born between the years 1996 and 2010. Their most important characteristic is that they are the first generation that has never known a world without the Internet. Their lives are moulded by the Internet, that has been converted in a natural part of their lives. In this vein, Marchi (2012) discusses how teenagers get informed about current events and why they prefer certain news formats to others. The results reveal changing ways news information is being accessed, new attitudes about what it means to be informed, and a youth preference for opinionated rather than objective news.

Generation Z will also be the most educated generation ever (Dimock, 2019). Notwithstanding, other studies (Chen, Sin, Theng, & Lee, 2015) have proven that the acquisition of a high level of education for an individual does not mean that this individual will stop accessing and distributing fake news. Other resources are needed.

2.3. Fake News, Future Journalists and Media Literacy

Many works focusing on the susceptibility of users and the reason why people might fall for false news, make efforts to propose concrete initiatives able to provide citizens with the necessary tools and skills to protect themselves from this vulnerability. All these studies agree that media literacy is the key for providing Internet users with a set of skills and abilities that will enable them to navigate with confidence, criteria and ethical parameters in cyberspace.

From an institutional perspective, the European Union defines 'media literacy' as the "capacity to access, have a critical understanding of, and interact with the media" (European Political Strategy Centre, 2017) and defends that media literacy is more important than ever, mainly in relation to young people's poor ability to distinguish fake news from true news, to perceive the influence of algorithms on social networks or to recognize bots. Until now, the European Commission's efforts have focused on promoting and measuring those more

technical skills (Durán-Becerra, 2016; Guess et al., 2019). UNESCO, on its side gives a special importance to the information within its proposed definition of 'media and information literacy' (UNESCO, 2013, 2018).

Media and information literacy share conceptual terrain and often overlap. Livingstone, Van Couvering, and Thumim (2008, p. 107) propose the following differentiation: "Media literacy sees media as a lens or window through which to view the world and express oneself, while information literacy sees information as a tool with which to act upon the world." We align with Pérez-Tornero and Varis' (2010) holistic approach, understanding media literacy as a concept embracing all the fields and all the competences related to media, that include news literacy.

Thus, news literacy can be defined a series of competences related to news, within the broader concept of media literacy. Malik, Cortesi, and Gasser (2013, pp. 8–9) propose a definition based on what it is meant to achieve: "An understanding of the role news plays in society; motivation to seek out news; the ability to find/identify/recognize news; the ability to critically evaluate news; the ability to create news."

Many studies stress out the benefit of media literacy (Spratt & Agosto, 2017) and news literacy in providing people the competences to protect themselves against fake news (Vraga & Tully, 2015; Vraga, Tully, Kotcher, Smithson, & Broeckelman-Post, 2015). Lotero-Echeverri, Romero-Rodríguez, and Pérez-Rodríguez (2018), within the framework of a study that analyses the relationship between media competition and fake news, stress the importance of this set of skills in tackling the problem of disinformation in its different variants and contexts. Kahne and Bowyer (2016) demonstrate that young people who had exposure to media literacy education were significantly more likely than young people without such exposure to be guided by accuracy motivation when making judgments about controversial political claims; media literacy essentially helped young people to override the pull of prior beliefs, or directional motivation, in making such judgments. Media literacy training is also linked with increased perceptions of credibility and trust in news media and was found to help reduce perceptions of media bias (Vraga, Tully, Akin, & Rojas, 2012).

Recent studies (Middaugh, 2019) focus specifically on critical skills as the most valuable resources to equip media users with. In an experiment related to coverage of biofuels, specific news media literacy training was found to reduce hostile interpretations of media, increase perceptions of news story credibility, and increase trust in the media generally and the news specifically (Vraga & Tully, 2015, 2016).

As previously mentioned, Herrero-Diz et al. (2019) and Mendiguren et al. (2020) point out how journalism do not feel protected against 'fake news.' Accordingly, the works of Romero-Rodríguez and Aguaded (2016) have warned that the media literacy and information filtering capacities of journalism students in Latin American countries denote a lack of critical and reflective analysis in relation to their deontological role in the processes of production of contents of a journalistic nature.

3. Methods

The research takes an exploratory perspective (Vilches, 2011), based on a hybrid methodology developed from a matrix survey of journalism students and a subsequent survey to find out the type of content they had received on their digital devices during the lockdown, especially focusing on the reception and sending of fake news. The survey samples have been selected for convenience and are composed of university students from the field of journalism at the Autonomous University of Barcelona (UAB). A total of 252 students participated in the surveys in 2019, comprising 143 women (57%) and 109 men (43%) aged 18–21. The survey conducted during the confinement in 2020 obtained 300 responses and was composed of 71% women and 29% men. The participants were informed of the study and their consent was requested to participate in it.

By focusing on a sample of students from a specific university, we do not intend to make any generalization, rather to offer an approximation. The UAB, thereby, has been selected for convenience (the researchers could conveniently have access to a representative sample of journalism for this educational centre) and because it is one of the better known and ranked universities at Spanish level in the Journalism field.

We complemented our research with students, with a focus group of teachers (n = 6), from the same university and department, selected for being responsible of teaching different disciplines around Multimedia Journalism. In this article, they are identified as Professor 1, Professor 2 and, consecutively, up to 6. The results of the focus group allow to assume a more qualitative perspective in the reading and interpretation of the students' answers. In order to further explore the information consumption and the position of journalism students towards fake news, the proposed method should be scaled to a broader sample, as well as to other cases.

The research combines both quantitative and qualitative data. Both the survey and the focus group have been designed based on a structured form of questions, both closed and open, mostly designed to encourage qualitative reflection.

The 2019 questionnaire, validated by a panel of experts (n = 10) in journalism, is composed of 22 open and closed questions about the identification and attitudes towards fake news, summarized in Table 1.

In addition, as an experimental manner, questions have been incorporated with examples of news that respondents should identify as true or fake news (see Figure 1).

The 2020 survey was not identical to the previous one, mainly focusing about disinformation during Covid-19 crisis, however it shared five questions with the

![Cogitatio logo] COGITATIO

Table 1. Thematic variables.

Research questions	Variables	Categories of analysis
How do journalism students get information?	Information consumption	Habits of access and consumption of informative content. Frequency, sources and objectives
	Sources of information	Identification of information sources typologies
	Information validation and contrasting	Routines of verification, contrast and validation of the consulted content
What is their attitude towards fake news?	Authorship of the fake news	Subjects responsible for the creation of false news
	Reason to create fake news	Reflection on the reasons of their production
	Audience	Typology of profiles and degrees of vulnerability to fake news. Analysis of the intergenerational component
	Methods to spot fake news	Identifying features/traits of fake news
	Importance and reach	Assessment of the seriousness and repercussions of the spread of false news

previous survey (see Table 1), that have been taken into consideration in order to complement and enrich 2019's study, addressing our RQ1.

4. Results

4.1. Information Consumption by Young Journalism Students

The first part of the 2019 survey deals with young people's information habits and relationship with different sources. To begin with, 90% of the respondents answered that they do use social networks for getting information. However, when asked about their favourite sources of information, more than 67% prefer online newspapers.

They consume updates on Instagram or Facebook from media accounts such as *Código Nuevo*, *El País*, *VICE*, just as they follow posts from friends or family, but they do not deliberately access them to be informed, they follow certain social accounts of online newspapers that sporadically feed them with information pills.

When it comes to topic that interest the young journalism students, 89% of respondents identify politics as the top thematic issue in their access to cyberspace platforms, with events (at 60%) and sports (at 23%) ranking second and third, respectively.

When it comes to one of the fundamental practices of the profession, contrasting information, the majority (77.1%) assures that they do, although only 43.3% say that they do so to verify the first information received; 30.6% say that they consult other sources to expand the information with more data, and 3.2% cite other reasons. Among those who assume that they are not in the habit of contrasting the information received, 12.3% say that they usually trust the first medium through which they receive the news, 7.5% say that they do not have time to verify the information, and 3.2% cite other reasons.

Figure 1. Example of question based on a case. Source: Authors.

Besides the habit of verifying or not the information, the students were also asked if they usually comment or discuss the news with their personal environment, as it can be seen in Figure 2. In this line, 50% assure that sometimes a week they discuss with their family or friends about this type of content; while 31% indicate that they do it daily.

Teachers, on their side, doubt about the quality of the debate they generate. According to Professor 3:

Most of them stay with the headline, this is not enough to have a deeper debate. Such an epidermal reading of information does not allow a subsequent debate to be generated when questioning information, let alone being able to glimpse the certainty of whether an item of information is false, real or true.

4.2. Journalism Students' Position towards Fake News

When asked how they would define so-called fake news, the most frequent responses were 'fake news,' 'fake fact' or similar, clearly associated with a literal translation of the English expression. But as this is an open question, we can gather a series of concepts and ideas, which are part of the conception of future journalists with the fake news. In Table 2, we present some of these recurrent ideas, which are basically associated with three elements or categories: the sender (his interests or professional practice), the message (its characteristics) and the objective (or intentionality) of the information.

There is a predominance of ideas associated with 'cheating,' 'manipulating,' 'confusing,' 'benefiting,' 'harming,' i.e., mainly focused on aims or objectives that are incompatible with the ethics and professional deontology of journalism.

When it comes to the topic mostly affected by fake news, Figure 3 shows that Culture, Politics and Events are the most quoted.

During Covid-19 pandemic, according to the respondents of the 2020 survey, fake news accounted for 4.5% of all the content received. Although they were not the most recurrent, 91.1% of the students acknowledge that they received this type of content, especially during the quarantine period.

Although most experts recognize that most of fake news are circulated for political reasons, students in the 2019 survey do not seem to be aware of this, precisely the politicians are in third place, they consider that the Internet users are the main distributors of fake news (see Figure 4).

As for the digital platforms on which they believe in which more fake news were circulating during the 2020 pandemic, more than half of the students indicated Facebook as the first (28.6%) or second choice (26.2%) closely followed by Twitter (24.6% and 23%, respectively). At the other end of the scale, online newspapers were considered to have the least amount of fake news (30% of those interviewed placed them at the top of the scale).

As Professor 4 states: "Social networks are another means to manipulate people and show or facilitate the information that interests them at a political or economic level…they do not have a critical vision to listen to other points of view." Therefore, teachers consider critical skills in students essential, especially media literacy. In spite of being journalism students, they "do not consult other sources or evaluate the journalist's information." The personification of information influences the belief of truth: "[Students] believe it and from there they don't move, like extremist positions, as this person said I don't believe the version that you say." And highlighting the phenomenon of echo chambers, Professor 4 explains a recent conversation with an AI specialist: "I was saying that we like to be given our ears, we believe the news we like." Professor 5 points out as an example of this in the political arena the case of Brexit in 2016 and the US elections in Trump. Professor 4 concludes by stressing the importance of democratic journalism, quoting the recognized journalist Martín Caparrós, who says: "In journalism it is not only necessary to tell people what they have the right to hear, but also to report what they do not want to hear and know."

Regarding the age groups that, in their opinion, contributed most to the misinformation, the students indicated: adults (43.3%), almost tied with adolescents (42.9%), and followed more closely by the elderly (12.7%) and children (1.2%). And on the motivations for contributing to disinformation, 24.6% consider that ideological interests predominate, 11.1% highlight economic
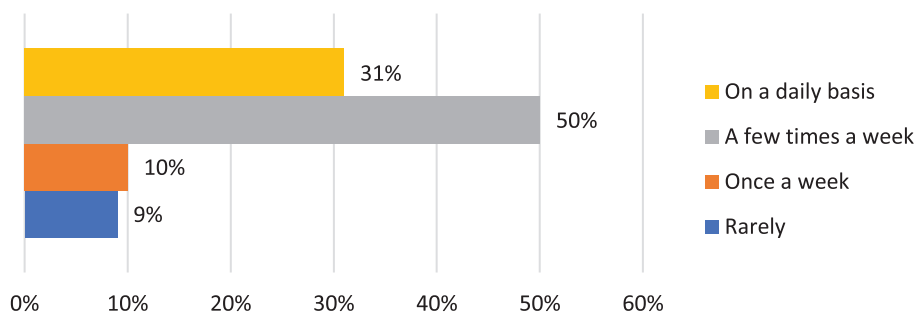

Figure 2. How often do you discuss the news with your family and friends?

Table 2. Concept and ideas about fake news.

| Sender	Message	Intentionality
"News published by sensationalist media…"	"…news with false data or completely invented…"	"…to attract the attention of the audience."
"…might be influenced by interests unrelated to the transmission of events."	"…news based on totally or partially fictitious facts…"	"…in order to make believe something or create an opinion about some subject or person."
"News not based on solid arguments or reliable sources…"	"…news in which is included untrue information or part of it is hidden…"	"…to damage another person's image, to see the effect it has and for the author's own interest…"
"…coming from doubtful sources…"	"…fake news that can be understood as true."	"…seeks to harm someone or a collective."
"…with interests unrelated to veracity, ethics and civic responsibility…"	"…false information spread by error or by lack of foundation and investigation."	"…are intended to create a social alarm."
"…made up by unreliable media…"	"News that have a high percentage of falseness…"	"…the intention to deceive the reader about an event."
"…often published by the country's leading newspapers and television stations."	"…tries to appear of being true but the information it offers cannot be contrasted."	"…whose author has no intention of informing, but of gaining an easy click or influencing the public."
	"It's often alarmist, ambiguous, unrealistic and even ridiculous news."	"…created with the will to influence the reader by manipulating information."
	"…nobody knows where it comes from (source), everybody believes it and shares it very quickly."	"…wants to spread a rumor of false news to confuse people…"
	"They decontextualize an event, make it seem more alarmist."	"…with the aim of confusing society or supporting a specific ideology."
		"…in order to satisfy or attack an ideology."
		"…is intended to disinform."
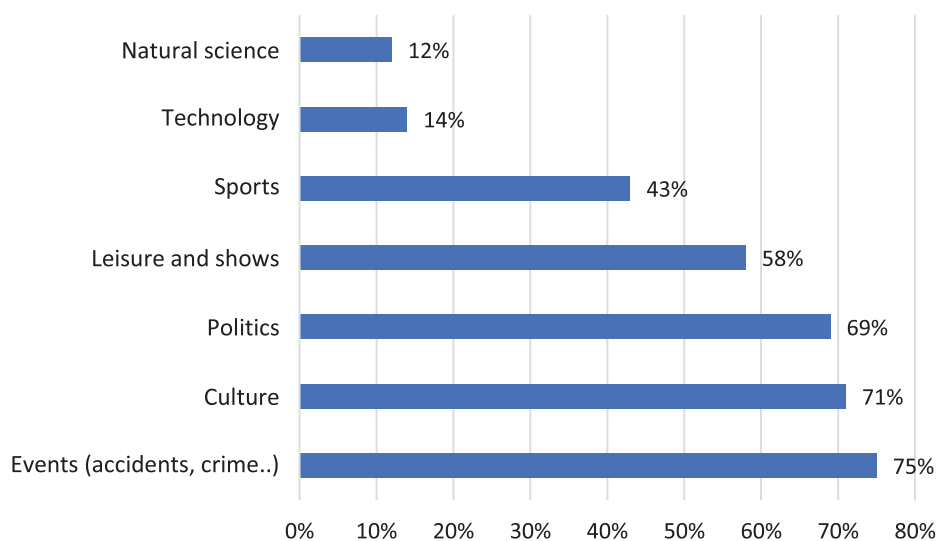
Figure 3. Which is the topic mostly affected by fake news?

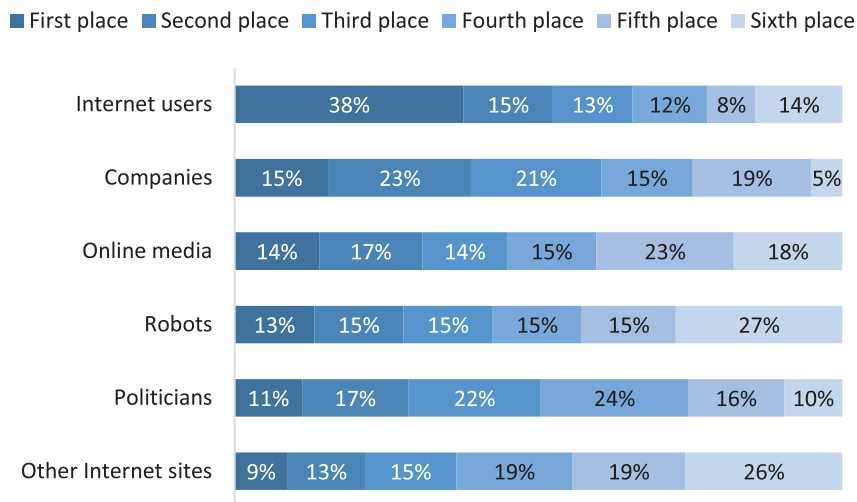■ First place ■ Second place ■ Third place ■ Fourth place ■ Fifth place ■ Sixth place

Figure 4. Who/which is the main source of fake news?

interests, but the majority consider that both ideological and economic interests motivate the creation and dissemination of false information in the same way. Regarding the ideology of the users who disseminate the most fake news, it should be noted that the vast majority of students agree that it is the right wing that shares the falsest information.

According to Professor 4, this response from the students "is not ideologically motivated, but rather in line with society." And as Professor 3 rightly comments:

> It makes no sense to ask and consider the survey sample if they think that the fake news is mostly distributed by the left, the right, etc….The students in the survey show in their response that they are starting from

a previous mistake, a product of ideological and cultural warfare, which is to assume that the falsified news is the product of a specific political ideology, without contrasting.

On the other hand, they were also asked about their own behaviour as Internet users, and whether they share or have shared fake news for any reason. The vast majority of the sample shared fake news, either deliberately or by mistake. Less than half (44.1%) said they had never shared a piece of fake news. Among the majority who do admit to having done so, they justify themselves on different grounds as shown in Figure 5.

The data leads us to question whether journalism students really know how to recognize fake news. Likewise,
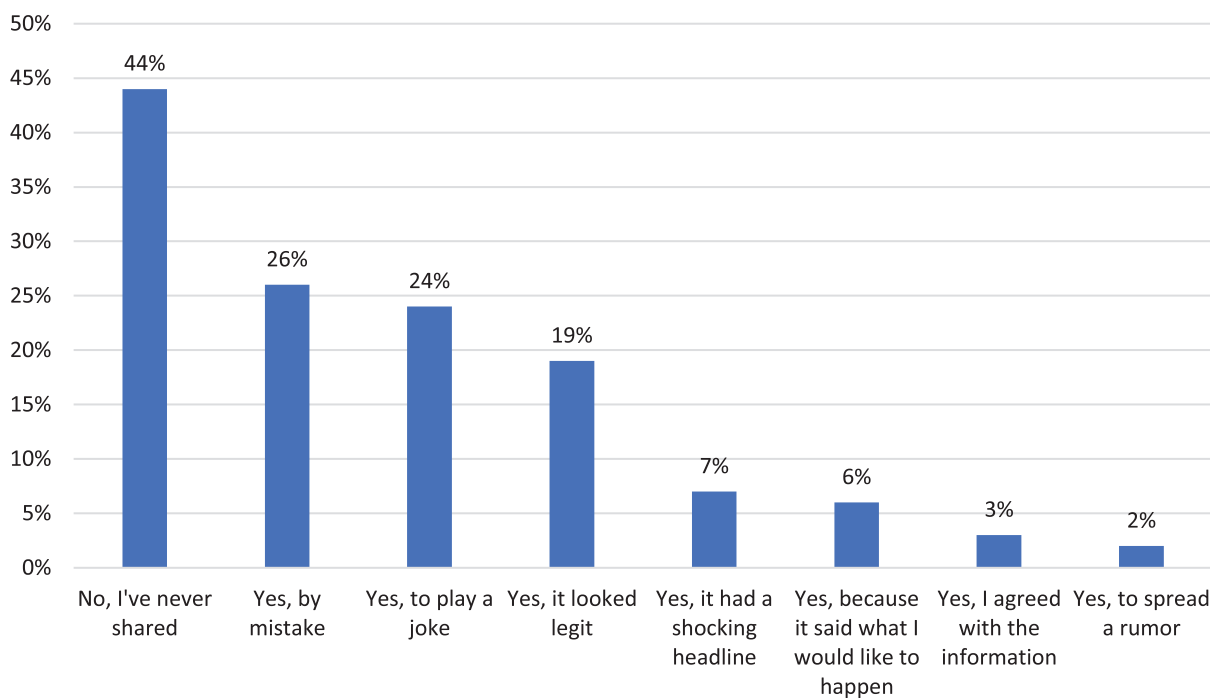

Figure 5. Have you ever shared fake news?

taking up the definitions above (Table 2), it is striking that some students connect fake news with the perception of news that cannot be contrasted and that is credible to most of the public.

Teachers perspective is particularly interesting: they insist that: "When it comes to fake news students are as helpless as everyone else" (Professor 2), however they recognized, as expressed by Professor 1, that "by end the semester they learn that it is necessary to verify, contextualize and check more than one source."

In this sense, when asked if they know when they are reading fake news, just over half (51.2%) answered 'sometimes.' 48% said they know 'almost always' when they read fake news, and 0.8% said they 'never' know. But no student claimed to 'always' know when they have false information in front of them. These answers somehow clash with another similar question. When asked if they think they know how to distinguish fake news from true information, less than half of the students (46.4%) said they think they know; 6.8% think they do not know how to make this distinction, and another 46.8% do not know what to answer. In order to contrast the self-perceptive information of whether they were able to detect fake news from true ones, a small evaluation test was elaborated, following the methodology of the media competence tests and 82% of the students correctly identified the fake news.

Students were also asked about the main elements or features that, in their opinion, allow the reader to detect fake news. As shown in Figure 6 most of them point to the headline style and the sources.

In conclusion, to answer RQ1, students consume information on social networks. Nonetheless, their main source of information, are online mainstream media and traditional media (TV or print). Interestingly, students consider that it is precisely social networks that spread the most fake news: firstly, Facebook, followed in second place by Twitter and in third place by Instagram. While cybermedia or other Internet sites are in fourth and fifth place respectively. However, despite identify-

ing these networks as the most active in disseminating fake news, they are the platforms most consulted in their information habits.

To answer RQ2, students consider that fake news are: a) False news published by a specific issuer; b) news with untruthful data, therefore, the message is wrong; and c) carrying an intentionality behind their dissemination. In this case, they think that it is mainly generated from political actors. They consider that the profile of the disseminator of false news is a user of social networks, mainly adolescent or adult. In addition, most students know that fake news is spread because of ideological and economic interests. Along these lines, students consider that politics is the subject that generates the most of false news, an aspect that is surely influenced by the country's political instability and the daily lives of young people. In relation to this, the students assure that those who spread the most fake news are the centre-right users.

Furthermore, no student in the sample knows for sure whether he or she is reading a false story. In this sense, a little more than half recognize detecting them only sometimes. This conclusion was reached even though they answered the evaluation question about detecting a fake news item correctly. Most students admit to sharing fake news (a quarter were by mistake or legitimate appearance).

In conclusion, it is necessary to take up the opinion of Professor 3, who states that we are witnessing "an infantilization of society and a cultural war for de-democratization." According to the interviewee fake news are nothing new, "All the reception studies and Walter Lippmann's theorization on public opinion has already warned us, but the consequences in the society of fake news are going to be catastrophic."

5. Discussion and Conclusions

Besides the geographical and sample limitations, our findings perfectly align with other studies realized in Spain (Herrero-Diz et al., 2019; Mendiguren et al., 2020),
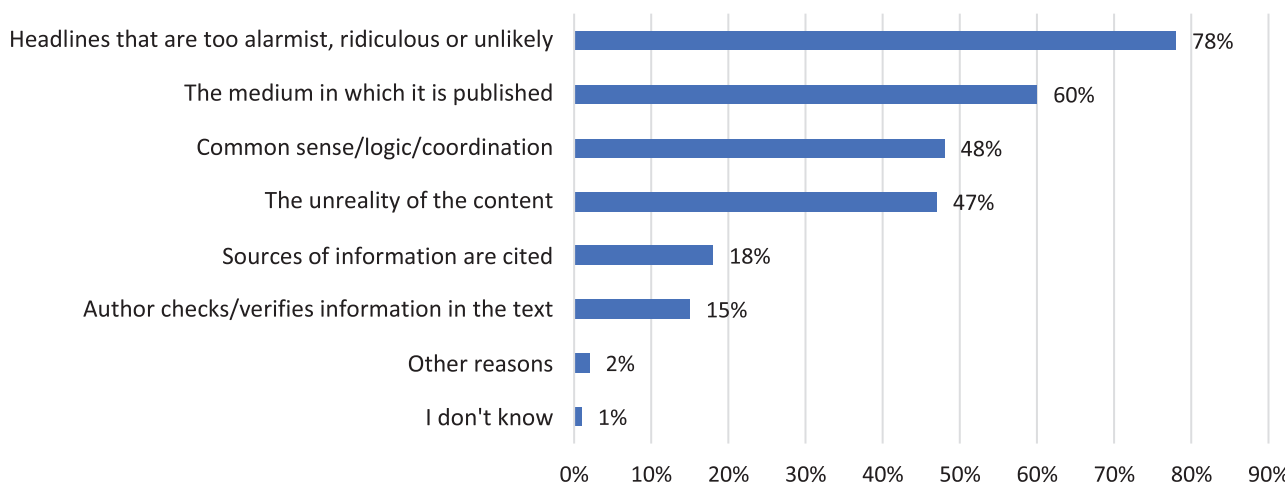
Figure 6. How can you spot fake news?

346

allowing to argue that certain trends are retrievable, at least at Spanish level: Journalism students get informed through social media, even if when looking for information they rely more online mainstream media. As for 'fake news,' they are well informed about what these are, they are concerned about this phenomenon, but they do not feel to have the necessary ability to spot one.

Our results also align with the findings on other studies in other contexts (Bhaskaran et al., 2019). To confirm this assumption, more meticulous comparative research is needed, not to overstate what could be context-specific and to allow some extent of generalization. Thereby, comparative and international research, besides assessing their perceptions, should go further into the specific assessment of the critical abilities of the students to spot fake news.

However, our results stress out the need to foster media literacy skills. Actions that encourage critical thinking must be implemented constantly. Education at the general level has an undisputed role in ensuring a media-literate society. This milestone encompasses not only digital competencies and technical skills, but also the importance of critical thinking in the face of increasingly changing information environments. In addition, issues such as AI, content creation robots and newsroom automation are taking a predominant role in communication.

In particular, fake news represent a threat not only for all the citizens, but for the future of journalism itself. Specifically, as the former editor of *The Guardian*, Alan Rusbridger (2018, p. 4), observes:

> Journalism is facing an existential economic threat in the form of a tumultuous recalibration of our place in the world. And on both sides of an increasingly scratchy debate about media, politics, and democracy, there is a hesitancy about whether there is any longer a common idea of what journalism is and why it matters.

Journalism Studies need to adapt accordingly and so Journalism Faculties and Schools. As Silvio Waisbord (2018) notes, it is necessary to weave the study of journalism with the rapidly, constantly changing communication ecology.

Various studies (Cervi, Pérez-Tornero, & Tejedor, 2020; Cervi, Simelio, & Tejedor Calvo, 2020) have already pinpointed that most Journalism School's educational programs are not ready to adapt to the important structural changes that the profession have undergone in recent years. Others (Waisbord, 2018) have highlighted the need to develop transversal actions for instructing both university professors and students in media competences to face an ecosystem dominated by fake news and disinformation. Thereby, our findings, together with the positive assessment of media literacy initiative (Vraga & Tully, 2015; Vraga et al., 2015), allow to suggest the need to reform journalism curricula focusing on the promotion of media literacy among students (Tejedor &

Cervi, 2017). In particular, news literacy (Spratt & Agosto, 2017), understood of a sub-branch of media literacy, is seen as a crucial competence for journalism students, but also as a necessary tool to recover the value of professional journalism and its foundational values, fundamental for the development and maintenance of a healthy public sphere.

Conflict of Interests

The authors declare no conflict of interests.

References

Bhaskaran, H., Mishra, H., & Nair, P. (2019). Journalism education in post-truth era: Pedagogical approaches based on Indian journalism students' perception of fake news. *Journalism & Mass Communication Educator*, *74*(2), 158–170. https://doi.org/10.1177%2F1077695819830034

Blanco Alfonso, I., García Galera, C., & Tejedor Calvo, S. (2019). El impacto de las fake news en la investigación en Ciencias Sociales: Revisión bibliográfica sistematizada [The impact of fake news on Social Science Research: Systematized bibliographic review]. *Historia y Comunicación Social*, *24*(2), 449–469. https://doi.org/10.5209/hics.66290

Cervi, L. (2019). Citizen journalism and user generated content in mainstream media: New dialogic form of communication, user-engagemente technique or free labor exploration? *Revista de Comunicação Dialógica*, *1*, 120–141. https://doi.org/10.12957/rcd.2019.41871

Cervi, L., & Carrillo-Andrade, A. (2019). Post-truth and disinformation: Using discourse analysis to understand the creation of emotional and rival narratives in Brexit. *ComHumanitas: Revista Científica de Comunicación*, *10*, 125–149. https://doi.org/10.31207/rch.v10i2.207

Cervi, L., Pérez-Tornero, J. M., & Tejedor, S. (2020). The challenge of teaching mobile journalism through MOOCs: A case study. *Sustainability*, *12*(13). https://doi.org/10.3390/su12135307

Cervi, L., Simelio, N., & Tejedor Calvo, S. (2020). Analysis of journalism and communication studies in Europe's top ranked universities: Competencies, aims and courses. *Journalism Practice*. https://doi.org/10.1080/17512786.2020.1762505

Chen, X., Sin, S. C. J., Theng, Y. L., & Lee, C. S. (2015). Why students share misinformation on social media: Motivation, gender, and study-level differences. *Journal of Academic Librarianship*, *41*(5), 583–592. https://doi.org/10.1016/j.acalib.2015.07.003

Dimock, M. (2019, January 17). Defining generations: Where Millennials end and Generation Z begins. *Pew Research Center*. Retrieved from https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins

Durán-Becerra, T. (2016). *AMI en Latinoamérica: Aproximación, análisis y propuesta de medición sobre el contexto de la Alfabetización Mediática e Informacional en América Latina* [AMI in Latin America. Aproximation, analysis and measurement proposal on the context of media and information literacy in Latin America] (Doctoral dissertation). Autonomous University of Barcelona, Barcelina, Spain. Retrieved from https://ddd.uab.cat/record/173980

European Political Strategy Centre. (2017). 10 Trends transforming education as we know it. *European Political Strategy Centre.* Retrieved from https://op.europa.eu/en/publication-detail/-/publication/227c6186-10d0-11ea-8c1f-01aa75ed71a1

Fetzer, J. H. (2004). Disinformation: The use of false information. *Minds and Machines*, *14*, 231–240. https://doi.org/10.1023/B:MIND.0000021683.28604.5b

Fletcher, R., Cornia, A., Graves, L., & Rasmus, K. (2018). *Measuring the reach of fake news and online disinformation in Europe*. Oxford: Reuters Institute for the Study of Journalism.

Giglietto, F., Righetti, N., Rossi, L., & Marino, G. (2020). It takes a village to manipulate the media: Coordinated link sharing behavior during 2018 and 2019 Italian elections. *Information, Communication & Society*, *23*(6), 867–891. https://doi.org/10.1080/1369118X.2020.1739732

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1). https://doi.org/10.1126/sciadv.aau4586

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. (2010). Trust online: Young adults' evaluation of web content. *International Journal of Communication*, *4*, 468–494.

Herrero-Diz, P., Conde-Jiménez, J., Tapia-Frade, A., & Varona-Aramburu, D. (2019). The credibility of online news: An evaluation of the information by university students. *Culture and Education*, *31*(2), 407–435. https://doi.org/10.1080/11356405.2019.1601937

Ireland, S. (2018). Fake news alerts: Teaching news literacy skills in a meme world. *The Reference Librarian*, *59*(3), 1–7. https://doi.org/10.1080/02763877.2018.1463890

Jang, S. M., & Kim, J. K. (2018). Third person effects of fake news: Fake news regulation and media literacy interventions. *Computers in Human Behavior*, *80*, 295–302. https://doi.org/10.1016/j.chb.2017.11.034

Kahne, J., & Bowyer, B. (2016). Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation. *American Educational Research Journal*, *54*(1), 3–34. https://doi.org/10.3102%2F0002831216679817

Livingstone, S., Van Couvering, E., & Thumim, N. (2008). Converging traditions of research on media and information literacies: Disciplinary, critical, and methodological issues. In J. Coiro, M. Knobel, C. Lankshear, & D. Leu (Eds.), *Handbook of research on new literacies* (pp. 103–132). New York, NY: Lawrence Erlbaum Associates.

Lotero-Echeverri, G., Romero-Rodríguez, L. M., & Pérez-Rodríguez, M. A. (2018). Fact-checking vs. Fake news: Periodismo de confirmación como componente de la competencia mediática contra la desinformación [Fact-checking vs. fake news: confirmation journalism as a tool of media literacy against misinformation]. *Index Comunicación*, *8*(2), 295–316.

Love, R. (2007). Before Jon Stewart. *Columbia Journalism Review*, *45*(6), 33–37.

Malik, M., Cortesi, S. C., & Gasser, U. (2013). The challenges of defining 'news literacy.' *Berkman Center Research Publication*. http://dx.doi.org/10.2139/ssrn.2342313

Marchi, R. (2012). With Facebook, blogs, and fake news, teens reject journalistic "objectivity." *Journal of Communication Inquiry*, *36*(3) 246–262.

Marcom, A., Murdoch, B., & Caulfield, T. (2017). Fake news portrayals of stem cells and stem cell research. *Regenerative Medicine*, *12*(7), 765–775. https://doi.org/10.2217/rme-2017-0060

Mendiguren, T., Pérez Dasilva, J., & Meso Ayerdi, K. (2020). Actitud ante las fake news: Estudio del caso de los estudiantes de la Universidad Del País Vasco [Facing fake news: The case of the students of the University of the Basque Country]. *Revista de Comunicación*, *19*(1), 171–184. https://doi.org/10.26441/RC19.1-2020-A10

Middaugh, E. (2019). More than just facts: Promoting civic media literacy in the era of outrage. *Peabody Journal of Education*, *94*(1), 17–31. https://doi.org/10.1080/0161956X.2019.1553582

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2019). 'Fake news' is not simply false information: A concept explication and taxonomy of online content. *The American Behavioral Scientist*. https://doi.org/10.1177%2F0002764219878224

Musgrove, A., Powers, J., Rebar, L., & Musgrove, G. (2018). Real or fake? Resources for teaching college students how to identify fake news. *College & Undergraduate Libraries*, *25*(3), 243–260. https://doi.org/10.1080/10691316.2018.1480444

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865–1880.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50, https://doi.org/10.1016/j.cognition.2018.06.011

Pérez-Tornero, J. M., Tayie, S., Tejedor, S., & Pulido, C. (2018). ¿Cómo afrontar las noticias falseadas mediante la alfabetización periodística? Estado de la cuestión [How to confront fake news through news literacy? State of the art]. *Doxa Comunicación:*

Revista Interdisciplinar de Estudios de Comunicación y Ciencias Sociales, *26*, 211–235.

Pérez-Tornero, J. M., & Varis, T. (2010). *Media literacy and new humanism.* Moscow: UNESCO Institute for Information Technologies in Education.

Peters, A., Tartari, E., Lotfinejad, N., Parneix, P., & Pittet, D. (2018). Fighting the good fight: The fallout of fake news in infection prevention and why context matters. *Journal of Hospital Infection*, *100*(4), 365–370. https://doi.org/10.1016/j.jhin.2018.08.001

Posetti, J., & Matthews, A. (2018). *A short guide to the history of 'fake news' and disinformation*. Washington, DC: International centre for Journalists. Retrieved from https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf

Romero-Rodríguez, L. M., & Aguaded, I. (2016). Consumo informativo y competencias digitales de estudiantes de periodismo de Colombia, Perú y Venezuela [Consumption of information and digital competencies of journalism students from Colombia, Peru and Venezuela]. *Convergencia: Revista de Ciencias Sociales*, *23*(70), 35–57.

Romero-Rodríguez, L. M., de-Casas-Moreno, P., & Torres-Toukoumidis, A. (2016). Dimensiones e indicadores de la calidad informativa en los medios digitales [Dimensions and indicators of informative quality in digital media]. *Comunicar*, *24*(49), 91–100. https://doi.org/10.3916/C49-2016-09

Rusbridger, A. (2018, August 31). Who broke the news? *The Guardian*. Retrieved from https://www.theguardian.com/news/2018/aug/31/alan-rusbridger-who-broke-the-news

Samuel-Azran, T., & Hayat, T. (2019). Online news recommendations credibility: The tie is mightier than the source. *Comunicar*, *60*, 71–80. https://doi.org/10.3916/C60-2019-07

Southwell, B., Thorson, E., & Sheble, L. (Eds.). (2018). *Misinformation and mass audiences*. Austin, TX: University of Texas.

Spratt, H. E., & Agosto, D. (2017). Fighting fake news because we all deserve the truth: Programming idea for teaching teens media literacy. *Young Adult Library Services*, *15*(4), 17–21.

Tally, R. (2011). I am the mainstream media (and so can you!). In A. Amarasingam (Ed.), *The Stewart/Colbert effect: Essays on the real impacts of fake news* (pp. 149–163). Jefferson, NC: McFarland.

Tandoc, E. C., Lim, Z. W., & Ling, R. (2017). Defining 'fake news': A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tandoc, E. C., Ling, R., Westlund, O., Duffy, A., Goh, D., & Zheng Wei, L. (2018). Audiences' acts of authentication in the age of fake news: A conceptual framework. *New Media & Society*, *20*(8), 2745–2763. https://doi.org/10.1177/1461444817731756

Tejedor, S., & Cervi, L. (2017). Análisis de los estudios de Periodismo y Comunicación en las principales universidades del mundo: Competencias, objetivos y asignaturas [Analysis of the journalism and communication degree programmes of the world's top-ranked universities: Competences, objectives and subjects]. *Revista Latina de Comunicación Social*, *72*, 1626–1647. http://dx.doi.org/10.4185/RLCS-2017-1238

Tejedor, S., Ventín, A., Cervi, L., Pulido, C., & Tusa, F. (2020). Native media and business models: Comparative study of 14 successful experiences in Latin America. *Media and Communication*, *8*(2), 146–158. http://dx.doi.org/10.17645/mac.v8i2.2712

Thompson, M. (2016). *Enough said: What's gone wrong with the language of politics?* London: The Bodley Head.

Toffler, A. (1970). *Future shock*. New York, NY: Random House.

Turkle, S. (2015). *Reclaiming conversation: The power of talk in a digital age*. New York, NY: Penguin Press.

UNESCO. (2013). *Global media and information literacy assessment framework: Country readiness and competencies*. Paris: UNESCO Communication and Information Sector. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000224655

UNESCO. (2018). *Journalism, fake news, and disinformation*. Paris: UNESCO Communication and Information Sector. Retrieved from https://en.unesco.org/sites/default/files/brochure_-_journalism_fake_news_and_disinformation.pdf

Vilches, L. (2011). *La investigación en comunicación: Métodos y técnicas en la era digital* [Communication research: Methods and techniques in the digital age]. Barcelona: Gedisa.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/sciencie.aap9559

Vraga, E. K., & Tully, M. (2015). Effectiveness of a non-classroom news media literacy intervention among different undergraduate populations. *Journalism & Mass Communication Educator*, *71*(4), 440–452. https://doi.org/10.1177/1077695815623399

Vraga, E. K., & Tully, M. (2016). Effective messaging to communicate news media literacy concepts to diverse publics. *Communication and the Public*, *1*(3), 305–322. https://doi.org/10.1177%2F2057047316670409

Vraga, E. K., Tully, M., Akin, M., & Rojas, H. (2012). Modifying perceptions of hostility and credibility of news coverage of an environmental controversy through media literacy. *Journalism*, *13*(7), 942–959. https://doi.org/10.1177%2F1464884912455906

Vraga, E. K., Tully, M., Kotcher, J. E., Smithson, A., & Broeckelman-Post, M. (2015). A multi-dimensional approach to measuring news media literacy. *Journal of Media Literacy Education*, *7*(3), 41–53.

Waisbord, S. (2018). Truth is what happens to news. *Journalism Studies*, *19*(13), 1866–1878. https://doi.org/10.1080/1461670X.2018.1492881

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Strasbourg: Council of Europe.

Wineburg, S., & McGrew, S. (2016). Why students can't google their way to the truth. *Education Week*, *36*(11), 22–28.

About the Authors

Santiago Tejedor is Director and Associate Professor at the Department of Journalism and Communication Sciences of the Autonomous University of Barcelona (UAB). He holds a PhD in Journalism and Communication Sciences from UAB and a PhD in Project Engineering from the Polytechnic University of Catalonia (UPC). He is a Member of the Communication and Education Office research group, and researches in the areas of digital journalism, media convergence and new transmedia narratives. He is also a Member of the Chair UNESCO UAB on MIL for Quality Journalism.

Marta Portalés-Oliva holds a PhD at the Journalism Department of the Autonomous University of Barcelona (2019) on media literacy. Since 2014 she is a Researcher at the group Gabinete de Comunicación y Educación. She studied Bachelor of Audiovisual Communication at the University of Valencia (Spain), York University in Toronto (Canada) and at the Johannes Gutenberg University Mainz (Germany) between 2009 and 2013.

Ricardo Carniel-Bugs is a Journalist (University of Santa Cruz do Sul, Brazil) and Associate Lecturer at the Department of Journalism and Communication Sciences of the Autonomous University of Barcelona (UAB). He holds a PhD in Communication and Journalism from UAB, with recognition of European Doctor. Member of the Laboratory of Prospective and Research in Communication, Culture and Cooperation and the Mediterranean Observatory of the Communication.

Laura Cervi is Serra Hunter Lecturer at the Department of Journalism and Communication Sciences of the Autonomous University of Barcelona. She holds a PhD in Political Science from the University of Pavia (Italy, 2006). Her main research interests are journalism and technology.

COGITATIO