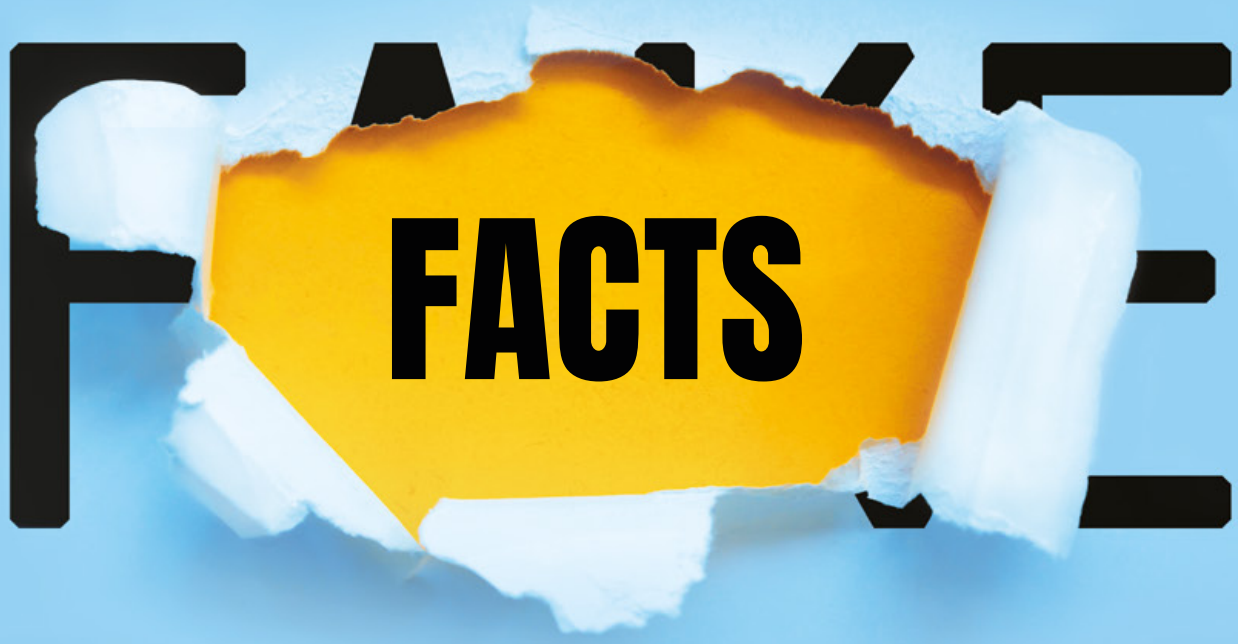

FIGHTING FAKE FACTS

Edited by Peter Seitz, Mark Eisenegger and Manfred Max Bergman

MDPI

Fighting Fake Facts

Fighting Fake Facts

Editors

Peter Seitz
Mark Eisenegger
Manfred Max Bergman

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Editors
Peter Seitz
Vice Presidency
Swiss Academy of Engineering Sciences
Zurich, Switzerland

Mark Eisenegger
Department of Communication and Media Research
University of Zürich
Zürich, Switzerland

Manfred Max Bergman
Department of Sociology
University of Basel
Basel, Switzerland

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

For citation purposes, cite each chapter independently as indicated below:

Lastname, A.A.; Lastname, B.B.; Lastname, C.C. Chapter Title. In *Book Title*; Editor 1, Editor 2, Eds. Series Title (optional). MDPI: Basel, Switzerland, Year, Page Range.

ISBN 978-3-0365-1348-5 (Hbk)
ISBN 978-3-0365-1347-8 (PDF)
doi.org/10.3390/books978-3-0365-1347-8

Contents

About the Editors

Peter Seitz is a Professor Emeritus of Optoelectronics at the Swiss Federal Institute of Technology (EPFL). He received his M.Sc. degree in experimental physics in 1980 from the Federal Institute of Technology (ETH) in Zurich, Switzerland, and in 1984, he obtained his Ph.D. degree from ETH for research work on X-ray Computer Tomography at the Institute for Biomedical Engineering. From 1984 to 1987, he was a staff member at the RCA research laboratories in Princeton, New Jersey, (David Sarnoff Research Center) and in Zurich, Switzerland, performing applied research in optics and image processing. In 1987, he joined the Paul Scherrer Institute in Zurich, where he created and led the Image Sensing research group. From 1997 to 2012, he was working for CSEM, first as a group leader and then as the head of CSEM's Photonics division in Zurich. From 2006 to 2011, he was CSEM's Vice President Nanomedicine, building and heading the Research Center for Nanomedicine in Landquart, Switzerland. Since 1998, he has also been a professor of Optoelectronics at the Institute for Microtechnology of the University of Neuchâtel, Switzerland, and in 2009, he transferred as an adjunct professor to EPFL. He was the Managing Director of the ETH Innovation and Entrepreneurship Lab, concurrently building and heading the Hamamatsu Photonics Europe Innovation Center. Peter Seitz has authored and co-authored more than 200 publications in the fields of applied optics, semiconductor image sensing, machine vision, optical metrology, and in the MedTech domain. He holds 75 patents, and he has won more than 20 national and international awards together with his teams, of which the most prestigious is the IST Grand Prize 2004 of the European Commission. He is a Fellow of the European Optical Society (EOS), the Vice President of the Swiss Academy of Engineering Sciences (SATW), and the Vice President of the European PPP Photonics21.

Mark Eisenegger is a full professor in the Department of Communication and Media Research (IKMZ) at the University of Zurich and the director of the Research Institute for the Public Sphere and Society (fög), also at the University of Zurich. From 2014 to 2017, he was a full professor at the Paris-Lodron University of Salzburg. He is a board member of the Swiss Association of Communication and Media Research (SACM), President of the Kurt-Imhof Foundation for Media Quality, and was Editor of "Communication Theory" from 2019 to 2023. His main research areas are organizational communication, media change, the digital transformation of the public

sphere and quality in journalism. He is especially interested in how digitization is transforming the public sphere and how this process affects organizations and journalism.

Manfred Max Bergman is a professor of social research and methodology at the University of Basel, Switzerland, and an adjunct professor at the University of Michigan. He is a graduate of the University of California, holds MA degrees from the Universities of Geneva and Cambridge, and completed his Ph.D. at the latter university. He chairs the World Sustainability Forum and the Basel Sustainability Forum, and he is a member of the Research Council of the Swiss National Science Foundation and the United Nation's Sustainable Development Solutions Network initiative. He is the Founding Editor and Editor-in-Chief of World and until recently was the Section Editor-in-Chief of Sustainability. A film he recently produced with Zinette Bergman on three case studies on sustainability in Pittsburgh, PA, has received three awards. He previously held the position of Head of Research and Methodology at the Swiss Information and Data Archive Service for the Social Sciences and co-directed the South African Social Sciences Research Methodology Winter School. He has also served as the editor of the Swiss Journal of Sociology and co-editor of the Journal of Mixed Methods Research. His research focuses on sustainability, particularly the interdependence of society, business, and government in a globalized world, and he is currently developing new research methods for policy-relevant and change-oriented research related to societal sustainability transformations.

Introduction: The Interdependence of Interests, Managed Truths, News, and Facts

Manfred Max Bergman

1. Types of Fake News

Fake news, often based on fake facts, is usually understood as misleading information presented as news. Fake news is frequently divided into misinformation (fake news without harmful intent) and disinformation (fake news with harmful intent) [1]. Intent, however, is a problematic classifier: a newscaster may initially use a story without harmful intent (for example, as a parody), but upon discovering its serendipitous effect on their audience, they may subsequently repeat the story with the intent to mislead. According to this dichotomy, the same story represents both misinformation and disinformation. Conveying the story, now classified as disinformation, to unsuspecting members of the newscaster's audience, who then repeat the story as news but without harmful intent, transforms the same story yet again from disinformation into misinformation. Recently, the term malinformation (genuine news with harmful intent) was introduced (Wardle, undated). Examples of malinformation are revenge porn or the publication of studies on the potential health risks of the COVID-19 vaccine, which are intended to harm the reputation of a person or the rollout of a health program, respectively. However, Mir [2] raises important concerns about the classification of malinformation, i.e., "wrong truths", as fake news, referring to the "Censorship-Industrial Complex", especially in contexts where persons or organizations have the power to shape informational content in their favor through sponsorship or censorship, irrespective of factual content.

Dozens of typologies of fake news have been proposed. For example, Wardle [3] suggests seven "information disorders": satire or parody, misleading content, imposter content, fabricated content, false connection, false context, and manipulated content. Reviewing

34 academic articles on fake news, Tandoc, Lim, and Ling [4] propose an alternative typology: news satire, news parody, news fabrication, photo manipulation, advertising and public relations, and propaganda. It appears that older typologies tend to focus on different ways to diverge from the truth, while recent approaches include the manipulation of truth itself. As such, an understanding of what fake news is and what its consequences are requires a clearer understanding of truth, from which fake news and fake facts diverge.

2. Types of Truth

Truth has been a scholarly subject for thousands of years, and the following does not do justice to this scholarship. Instead, what is presented here is an abbreviated reflection on truth from a social science perspective, which owes a great debt to philosophy, social theory, and science studies.

Factual truth: Possibly the most classical and dominant proposition about truth is that truth statements (sometimes referred to as truthful propositions or true beliefs) directly relate to facts. In other words, this type of truth necessitates the existence of an objective reality made up of facts that exists irrespective of how humans think of or act toward it. The problem here is that this proposition requires the definition of a fact, which leads to an unavoidable tautology, namely, that a fact, here an organizing principle of reality, is that which is true. According to the World Health Organization, for example, "Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus" [5]. Five facts are necessary to render this statement factual: the factual existence of COVID-19, the factual existence of infectious diseases, the factual classification of COVID-19 as an infectious disease, the factual existence of the SARS-CoV-2 virus, and the factual existence of the causal link between the virus and the disease through infection. If all these are factually established, then the statement by the WHO is factually true. But from a factual truth perspective, this statement can only be true if all elements in the statement are factual. I have no reason to doubt the veracity of this statement, given that I have read countless scientific and journalistic articles on COVID-19. I have been infected with the virus multiple

times and observed it infecting others, and I have listened to scholarly exchanges between virologists on this matter. Furthermore, this statement also makes sense to me, as it does to the great majority of members of my peer group. However, none of these reflections and experiences renders my belief in the truth of this statement a factual truth.

Data-centric truth: There is a large toolbox of methods with which to investigate truth through evidence. While data-centric truth is often equated with factual truth, especially since evidence is incorrectly understood as indicative of truth, it is at best an approximation. Sometimes, data are imperfect indications of truth (and, sometimes, they lead away from truth or are nothing but "noise"), and even the "right" data can be (mis)interpreted in a variety of ways. Empirical methods themselves are subject to multiple biases that thus further obfuscate direct access to factual truth. Data-centric or evidence-based truth statements are limited by multiple assumptions and limitations imposed through the very activity of empirical inquiry. Truth statements based on rigorously applied empirical methods ***may*** come close enough to factual truth, but it is rarely possible to measure the exact difference between data-centric truth and factual truth. For example, many studies have examined the national or global prevalence of COVID-19 infections. Understandably, these numbers vary widely depending on how COVID-19 prevalence is defined, identified, and measured (e.g., extrapolation from sample populations with antibodies or via statistical modeling) and the inferences drawn from subpopulations assessed with different degrees of rigor. Based on a specified set of data and methods and within their corresponding limitations, empirical evidence will provide nothing but an estimation of factual truth based on imperfect data and imperfect methods. The resulting prevalence cannot be understood as factual truth but merely an approximation.

Deliberative truth: Expertise has an interesting status in relation to truth. While most believe that experts have privileged access to knowledge and thus truth, we also have developed a healthy doubt about expertise. Furthermore, in some applications, we are increasingly recognizing the expertise of lay persons. Accordingly,

deliberation among experts, among lay persons and experts, or among lay persons is often used in think tanks or transdisciplinary research approaches to improve on the limits imposed by conventional researchers or experts. As part of co-designing research and co-creating knowledge, deliberation, participation, and, to some extent, democratic decision making during the investigative search for truth become part of the production of truth statements. In this sense, truth is not merely extracted from data using scientific methods but instead negotiated with significant stakeholders by employing various participatory techniques, which may or may not include data and methods. From a deliberative truth perspective, the "best" mode of a COVID-19 vaccine rollout may be identified based on what a heterogeneous group of relevant stakeholders judges to be best. While this approach to truth makes, in particular, knowledge-to-practice economically viable, politically defensible, and culturally acceptable, we may never know if what the group decided at a specific place and time was indeed the "best" method for the rollout. In other words, how many Katalin Karikós and Drew Weissmans, the discoverers of nucleoside base modifications that enabled the development of mRNA vaccines against COVID-19, need to be in the room to negotiate the right approach to mRNA vaccine development?

Systemic truth: Truth may also be understood as system-specific such that a truth statement is true if it coherently relates to other truth statements within a particular system. In this variant, truth is derived from the coherence it forms with other truth statements within a system, and no external references to facts are necessary to maintain the veracity of systemic truth. As stated earlier, the majority of my peers and I believe in the infectious properties of SARS-CoV-2 and the adverse health outcomes that this virus can cause, especially in vulnerable populations. There must be hundreds of overlapping and confirmatory beliefs that I hold that make me consider the WHO's statement a statement of truth. However, the architecture of truth statements shared by other communities may be as well supported as mine, except that their system-specific truth statements make members of this community doubt that such a virus exists or believe that if it exists, it is no more harmful than the common cold.

Holistic truth: Instead of multiple truth systems that may overlap or even compete with each other, a wider understanding of truth would include the entirety of all existing truth statements, irrespective of their relation to so-called facts or reality and regardless of their degrees of coherence with each other. Truthfulness in this sense is not attributable to singular statements but to the entirety of all existing statements. Truth statements associated with the existence or inexistence of COVID-19 are part of a discursive environment that does not necessitate evaluation criteria beyond their existence or their coexistence with other truth statements. Thus conceptualized, holistic truth is generated as a cacophony of truth statements, and its product is the entire collection of truth statements, as illustrated by the content of various social media channels.

Egocentric truth: Egocentric truth systems do not refer to the entirety of possible truth statements within a system. Egocentric truth may be nothing more than truth statements uttered by an individual in a specific time and place. Such truth statements may or may not form relations with each other or coalesce with truth statements made by significant others. Instead, they may merely reflect an agential insistence of the self as the ultimate judge of truth. Someone believing in the power of numerology could claim that COVID-19 emerged in 2019, numerically corresponding to 3 (2 + 0 + 1 + 9 = 12, i.e., 1 + 2 = 3), which, to this person's understanding, represents the number of divinity. Given the perceived upheaval associated with COVID-19 in 2019, this person may realize, based on evidence gleaned from his environment in conjunction with numerological exegesis, that the spiritual power behind the pandemic is the Antichrist. Facial masks, according to this person, are ineffective attempts to ward off Satan, while a COVID-19 vaccination is nothing less than the "mark of the beast" or formal allegiance to Satan. This individual supports this truth with evidence gleaned from the Book of Revelation. While this example seems far-fetched, it is nevertheless striking how egocentric truth statements about COVID-19′s origins, physiological effects, or consequences have mutated even in relatively homogeneous populations.

We can draw three insights from this brief and incomplete sketch of truth from a social science perspective: First, this concept of truth is more complex (composed of many parts) and complicated (associated with many extraneous elements) than the one implied by factual truth. Second, if we could identify factual truth, then mitigating fake news would be relatively easy. Third, all types of truth outlined here tend toward making factual claims about truth, whether they are produced by means of data, deliberation, systemic coherence, etc. In the end, it is no longer the fakeness of news or facts, i.e., the distance between a truth claim and factual truth, that is the main cause of concern; instead, it is the factualness of facts.

3. The Interdependence of Interests, Managed Truth, News, and Facts

Two of the most important institutions charged with investigating, protecting, and even defending truth are science and the news media. In my faculty, for example, doctoral candidates have to swear an oath, wherein they must promise "to always regard the scientific investigation of the truth as a sincere and fundamental task, to promote this goal to the bests of my abilities, and to always act responsibly, conscientiously and impartially in all scientific activities" [6]. Only after this oath is taken will the successful doctoral candidate receive their degree. While this oath is not in any way representative, many institutions of higher learning would subscribe to norms and values similar to those implied here.

High professional norms and values are also embedded in PBS News Anchor Jim Lehrer's Harvard commencement address from 2006. He proposes the following nine rules for good journalism: "(1) Do nothing I cannot defend. (2) Cover, write, and present every story with the care I would want if the story were about me. (3) Assume there is at least one other side or version to every story. (4) Assume the viewer is as smart and caring and good a person as I am. (5) Assume the same about all people on whom I report. (6) Assume personal lives are a private matter until a legitimate turn in the story mandates otherwise. (7) Carefully separate opinion and analysis from straight news stories and clearly label them as such. (8) No one should

ever be allowed to attack another anonymously. (9) I am not in the entertainment business" [7].

As different as both institutions may be in terms of their institutional histories, purpose, function, funding, and societal roles, we can nevertheless find many overlapping lofty goals associated with both science and journalism, many of which relate to an uncompromising commitment to the exploration and unfettered reporting of truth. But how does this truth fare beyond decontextualized promises and mission statements? While science and the news media indeed present themselves as the guardians of truth (even as the guardians of multiple and sometimes incommensurable truths), their interdependencies with other institutions and the thus developing institutional arrangements make it difficult to maintain the standards to which these institutions aspire (see Figure 1).

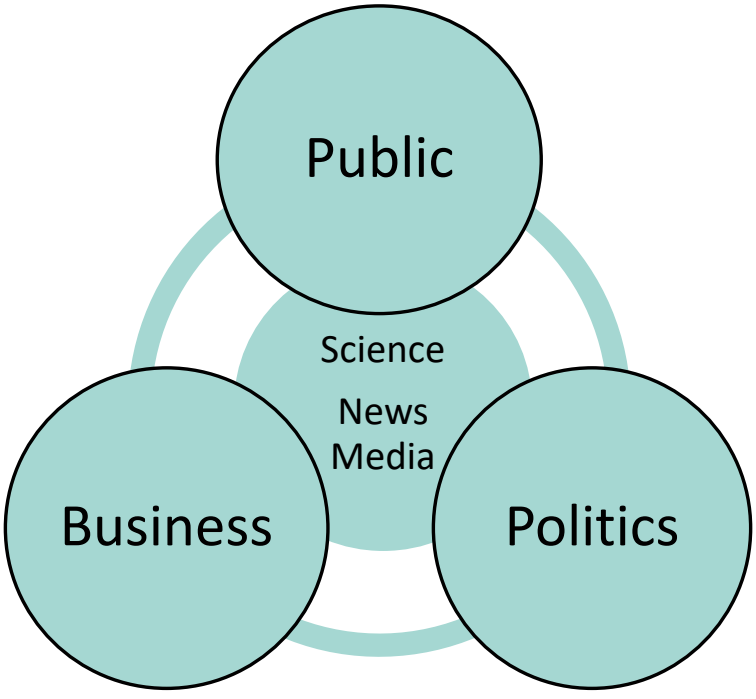

Figure 1. Interdependencies between science, the news media, the public, politics, and the business sector.

Most universities, public or private, are funded in ways that require them to be accountable to their funders. However, funding dependencies also place increasing pressures on universities to acquiesce to demands and conditions that may have implications for their once lofty missions. For example, a professor may have

received some bad press due to the content of a lecture or interview, putting political or public pressure on the institution to act in a way that signals to the public, business, or political actors the rigor with which the professor has been dealt with. Or, for similar transgressions, a university may be forced to rid itself of a university president in order to prevent a private donor from withdrawing a significant bequest. While these two examples seemingly deal with only two individuals, the involved universities are sending a strong and clear signal about acceptable and unacceptable work and conduct, thus creating an expectation filter that will influence truth work and truth statements regarding the respective subject and conduct arenas.

In a similar vein, a state-funded university may be judged ideologically inadequate by a political body. By withholding funds or guiding funding toward appointments or units that rectify this perceived inadequacy, the university and its members receive a strong and clear signal about the kind of truth statements that are incentivized or discouraged. Such (dis)incentivizing is not an invitation to create fake news or fake facts. Instead, it is an invitation to adjust discourses and activities associated with potentially harmful or pleasing truth statements. Simply deleting courses from syllabi, cancelling lectures and debates, rescinding invitations to collaborate, or keeping quiet when certain topics arise may already achieve the desired result. Similar mechanisms exist in association with the funding of scientific units, personal chairs, or institute directors, especially if such funding comes with ideological strings attached. Even in these situations, universities and their members may be free to search for truth, as long as they do not jar with relevant public, political, or business interests. Going one step further, many successful universities may elect to arrange their pursuit of truth in line with public, political, and especially business interests. Merely a disinvestment in the humanities in favor of more life science funding may find stronger support among the three institutions. Accordingly, the lofty mission statements of universities are bounded by the interdependencies that provide opportunities to some and impose limits on others. Overall, however, such interdependencies signal to the institution and its members the value of different types and pursuits of truth.

The news media, despite its many differences from universities, is in a similar position. Ruled by complex management and ownership structures and dependent on an often declining readership and advertisement base, conventional news outlets have difficulty maintaining research and reporting staff to cover news as envisioned by Lehrer. Furthermore, considerable competition from social media channels and influencers, increasing legal threats from powerful opinion leaders or the billionaire class, and an increasing pressure to adjust the degree and kind of reporting of personae, organizations, and events in ways that are conducive to news makers, advertisers, or politically dominant ideologies may not necessarily lead to producing fake news or facts, but these factors do create incentives to adjust truth statements to either please or not jar unduly relevant members of the public, political system, or business.

Attenuating, embellishing, negotiating, aligning, and carefully managing truths as a function of interdependencies can translate into opportunities for some and burdens for others. Such interdependencies have contributed to what Wardle refers to as information disorders. The pursuit of any type of truth has never been easy and unconstrained. However, we are not only experiencing a massive surge in fake news and fake facts, which are increasingly difficult to detect and mitigate; more importantly, we have entered an era in which facts and news no longer need to be fake to detract from the truth. According to Mir [2], we are experiencing an "epistemological shift from absolute truth to negotiated truth… The cultural and generational consequences will be much deeper. The next generation will not challenge or bypass the absolute truth; they simply will not know what it is."

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. APA. Topic: Misinformation and Disinformation. Available online: https://www.apa.org/topics/journalism-facts/misinformation-disinformation (accessed on 28 December 2023).

2. Mir, A. 'Malinformation' and the Wrong Truth. 16 March 2023. Available online: https://www.discoursemagazine.com/p/malinformation-and-the-wrong-truth (accessed on 28 December 2023).
3. Wardle, C. The Age of Information Disorder. Available online: https://datajournalism.com/read/handbook/verification-3/investigating-disinformation-and-media-manipulation/the-age-of-information-disorder (accessed on 28 December 2023).
4. Tandoc, E.C., Jr.; Lim, Z.W.; Ling, R. Defining "Fake News". *Digit. J.* **2018**, *6*, 137–153. [CrossRef]
5. World Health Organization. Coronavirus Disease (COVID-19). Available online: https://www.who.int/health-topics/coronavirus (accessed on 28 December 2023).
6. University of Basel. Ordnung der Philosophisch-Historischen Fakultät der Universität Basel für die Promotion, 446.540 9. 2017. Available online: https://www.unibas.ch/dam/jcr:73d7f4b6-8a5d-4ea5-af79-5727917155ec/446_540_04.pdf (accessed on 28 December 2023).
7. Jones, J. Jim Lehrer's 16 Rules for Practicing Journalism with Integrity, Open Culture. 2020. Available online: https://www.openculture.com/2020/01/jim-lehrers-16-rules-for-being-a-journalist-with-integrity.html (accessed on 28 December 2023).

A Psychological Perspective on Misinformation: Why Do Individuals Trust Fake News?

Mariela Jaffé and Rainer Greifeneder

Abstract: Abstract: Fake news is not a new phenomenon. The impact of misinformation, however, has increased massively in recent years, as the internet and social media have allowed the spread of true and false information at exceptional rates to an unprecedented number of recipients. We offer a psychological perspective on misinformation and discuss reasons why people may believe and share false information. When people process information, they rely on cognitive shortcuts which may increase the likelihood of believing information that, for example, has been heard before (repetition–truth effect), is associated with non-probative pictures (truthiness bias), is presented in a negative frame (negativity bias), and is in line with personal opinions (confirmation bias). Understanding these tendencies allows mitigation measures to be designed that enable news providers, social media platforms, and individuals to combat misinformation. Here, we integrate selected research to discuss options for social media platforms to reduce the sharing of misinformation. Moreover, we compile a set of recommendations for individuals, allowing them to create their own truth toolbox. Using this toolbox may enable them to critically approach and navigate information in everyday life.

1. Fake News in the Past and Today

News is and has always been part of our daily lives. When we get up, we check news headlines on our phone, listen to updates on the radio, or enjoy a cup of coffee while browsing through the newspaper. Over the last years, more and more of our news consumption has begun to take place online. People read the newspaper online, watch the news whenever they have a break, and listen to a podcast that they

can start whenever they like. News, moreover, is not only consumed online, but also more and more via social media. Instead of looking at the website of a newspaper or public television channel to learn about what is happening in the world, people open Facebook, Twitter, Instagram, or other social media applications and see a curated mix of news items, texts, and pictures from organizations or people they are connected with, as well as advertised posts. Oftentimes it is on these platforms that people access information about political and social events and news of any kind [1].

Whenever people are in contact with news, they likely encounter misinformation, too. In the months up to the 2016 US election, researchers estimated that each voting age participant in their study saw an average of 5.5 fake news articles within 1.5 months [2,3]. Fake news is defined as "false, often sensational, information disseminated under the guise of news reporting" [4]. One may further want to distinguish between misinformation (meaning information that is false but not necessarily intended to harm) and disinformation (information that is false and was designed to mislead the recipient) [5]. As, however, the intention cannot be identified with certainty from the outside, we use the term misinformation throughout the chapter. While it may be tempting to think that misinformation is solely associated with the widespread use of social media, this is not necessarily true see, e.g., [6].

A misinformation campaign has been documented as far back as 1274 BCE: Ramses II wanted to pretend that his campaign on the city of Kadesh was successful (which it was not). To spread his version of the story, however, he had to rely on a commissioned poem that was circulated on papyrus as well as on texts that were carved into the stones of royal architectural sites, eventually accompanied by reliefs for the illiterate viewer [7]. Back then, spreading misinformation was a costly and slow endeavor.

The invention of the printing press by Johannes Gutenberg in the 15th century made the distribution of text much easier, but it was still far away from today's reach via social media, which enables individuals and organizations to spread information (and therefore also misinformation) to a large audience with a few clicks to a large

audience. Today, even non-human agents such as algorithms, social bots, etc., can spread information.

Misinformation, therefore, represents a long-lasting problem for past, current, and presumably future generations. As the opportunities to spread misinformation increase to unprecedented levels, the consequences can be serious [6]. Misinformation contradicts individuals' preferences to make well-informed and reason-based decisions [8]. Misinformation can change actions with grave implications. For instance, after being confronted with misinformation, individuals reported lower intentions to use tracking apps to combat the spread of SAR-COV-2 [9]. Fake news also decreases interpersonal trust and trust in democratic institutions [10,11], which are the basis for peaceful human interactions and societal prosperity.

To discuss potential measures to counteract the problem, in a first step we provide selected findings from psychological research to illustrate why people believe in misinformation and may contribute to its distribution by liking and sharing it on social media. While the creation of misinformation is sometimes associated with malintent (this is particularly true, by definition, when we discuss disinformation), this is not necessarily the case when recipients like or share a news item.

2. Why Do People Believe and Share False Information?

When consuming news, people may operate under the assumptions that their communication partners follow the maxim of quality [12]. This maxim holds that communication partners share what they themselves believe to be true. Traditional media channels such as newspapers reassured this belief by self-imposing ethical norms for journalism see, e.g., [13]. However, the maxim of quality and the norms of ethical journalism do not apply to agents that create fake news and harm. As people expect truth i.e., have a truth bias, see [14], they generally do not anticipate lies and deceit and can therefore become easy prey for those who want to manipulate.

Next to the problem of operating with a truth bias, individuals might also be ill equipped to detect misinformation due to the nature of information transmission in social media and everyday life. As articles in newspapers are long, errors and inconsistencies are more easily

detectable. On social media platforms, in contrast, articles are mostly displayed as text snippets that leave less room for inconsistencies to occur, which makes it more difficult for users to assess veracity [10].

Another reason for peoples' difficulties in detecting misinformation is the way information is processed. When investigating human information processing, dual-process theories in psychology argue that people have two parallel and independent processing modes. Epstein et al. [15], for example, refer to a rational system and an experiential system. The rational system operates on a conscious level and is primarily analytical. The experiential system is instead assumed to be more automatic, affect-driven, and reliant on heuristics [15]. Kahneman and Frederick (2002) and Kahneman (2011) [16,17] refer to the decision modes as System 1, which is intuitive, automatic, and effortless, and System 2, which is reflective, slow, and effortful.

Whether people rely more strongly on System 1 or System 2 depends on situational and personal variables (e.g., time pressure and/or individual preferences). However, when it comes to the processing of everyday information such as news, the experiential and intuitive System 1 may be faster in proposing reactions [16] and System 2 may be too slow and effortful in eventually correcting the wrong impression e.g., [18]. Consistent with this assumption, Bago et al. [19] show that participants who engaged in more reflective reasoning (i.e., where the rational System 2 was dominant) were also better at telling true from false news. In a different line of work, the authors showed that reliance on emotion, associated with System 1 processing, increased the likelihood that misinformation is believed to be true [20]. If people (a) do not have a preference for reflective reasoning, (b) are not prompted by situational cues to engage in effortful processing, or c) prefer to rely on emotions or are nudged to do so, it appears likely that they will use System 1 to process news, meaning that they rely on cognitive shortcuts (i.e., a rule of thumb). Oftentimes, these shortcuts may lead to outcomes that would have resulted from System 2 processing, too, but in a context where deceit and manipulation are present, certain shortcuts may impede truth judgments. By way of example, we here discuss four shortcuts and their potential shortcomings in the context

of truth judgments: the repetition–truth effect, the truthiness bias, the negativity bias, and the confirmation bias or partisanship.

2.1. Repetition–Truth Effect

Prior exposure to news may critically influence how we perceive it during a subsequent encounter. Indeed, repeated exposure to a piece of information (compared to seeing it only once) increases the belief that the information is true see e.g., [21]. This phenomenon is called the repetition–truth effect [22] or the illusory truth effect [11,23].

Hasher et al. [24] asked participants to listen to a set of true and false but always plausible statements on three successive occasions. Some of the statements were presented repeatedly, whereas the rest of the statements were presented only once. Participants were then asked for truth ratings, and results show that repeated statements were more likely to be judged as true compared to statements presented only once. Presumably, hearing or reading a piece of information again increases the ease with which the item can be processed, making it appear familiar. Familiarity (or the ease of processing itself), in turn, can be used to judge whether a statement is true. Research shows that individuals consistently and strongly rely on the ease of processing to judge truth even in the presence of valid declarative information that speaks to the contrary [25]. This heuristic strategy may result in valid judgments as long as true things are repeated again and again, whereas false things are not [12]. However, in post-truth times where disinformation is created to manipulate beliefs and is, as research shows, eventually distributed faster and further compared to true information [26], the ease of processing may falsely indicate truth.

Pennycook et al. [11] tested the repetition–truth effect in the realm of misinformation. Working with news headlines, the authors showed that headlines that participants had seen before and that were therefore easier to process were subsequently more likely to be judged as true. This was the case for both true and false news, and also when participants were explicitly warned and not warned about fake news. This means that if news is repeatedly presented, it is more likely to be judged as true—even if it contains a low level of overall believability or is labeled as contested by fact checkers [11,25].

2.2. Truthiness Bias

News is often accompanied by pictures Newman and Zhang 2020 [27]. These pictures may further illustrate or explain the content, and are thus probative, or they may only be decorative and are thus non-probative (e.g., a picture of Windsor Castle in a report on Queen Elizabeth II's health). Interestingly, even though it is non-probative, the presence of a decorative picture may impact judgments of truth. Showing participants a picture of a famous person accompanied with the statement that this person was dead or alive, for example, and then asking participants whether this was true resulted in higher truth ratings [28]. This truthiness bias is not contingent on positive messages (e.g., the famous person is alive), but works for negative messages as well [28].

Pictures may increase perceived truth, especially when people do not know whether something is true or false. In such situations of ambiguity, individuals may rely on associations, feelings, and their processing through System 1. Non-probative, decorative pictures may also facilitate conceptual processing of the information [27], thus increasing fluency. Furthermore, pictures may help people to generate pseudo evidence, meaning that whatever statement people see, they might selectively interpret information in the picture as consistent proof [28]. Combining a misinformation headline with a picture [27] may therefore increase perceived truthfulness and deceive recipients.

Aside from truthiness effects, adding pictures to statements may also boost the intention to like and share the information, thereby distributing the information further on social media [29].

2.3. Negativity Bias

Individuals are more likely to attend to, process, and remember negative information compared to positive (negativity bias) [30,31]. Moreover, negative information is also more salient and diagnostic compared to positive information [32]. All of these aspects may impact judgments of truth if information is framed negatively rather than positively.

Hilbig [33,34] showed that framing an identical piece of information negatively compared to positively results in a higher

likelihood of the statement being judged as true. For instance, after learning that "XX% of marriages are divorced within the first 10 years"—compared to the equivalent information of "100—XX% of marriages last for at least 10 years"—participants perceived the statement as more likely to be true. Focusing on the downsides compared to the upsides might be particularly compelling and could therefore result in people being more likely to trust negatively framed information. Further research has replicated and extended the negativity bias in judgments of truth e.g., [35,36]; for instance, the bias appears to be more pronounced in the here and now, meaning when statements refer to a spatially close place (the country people live in) rather than to a more spatially distant place [36].

A particularly effective way to create a negative frame is through the use of negation. Negation is a semantic operator that frames a conceptually positive concept in a negative way [37]. Statements such as "XX% of marriages do not last for at least 10 years" are more likely judged as true compared to the equivalents of "100—XX% of marriages last for at least 10 years". Additional studies explain this effect by showing that negation may increase the realm of possible states in which a statement could be true [37]. Rephrased in the words of Tolstoi in his 1877 novel Anna Karenina, every unhappy family is unhappy in their own way, whereas all happy families are alike [38]. As negating a positive concept (e.g., happy) increases the number of situations to which a statement may pertain (e.g., there are so many different ways to be not happy), the base rate likelihood of being true may be higher for negated compared to non-negated content.

Creators of misinformation may aim to intentionally increase the chances that statements appear as true by intuitively or deliberately framing content negatively compared to positively. They may highlight the number of disagreements versus agreements, or mention the percentage of people who do not have access to health insurance versus the percentage of people who have access.

2.4. Confirmation Bias, Partisanship, and Cognitive Reflection

People like to selectively expose themselves to information, news, and people that confirm their attitudes and beliefs [39,40]. People

are more likely to spend time and engage with content if it is in line with their preferences; they prefer information that confirms what they already believe (see [40,41]).

When it comes to news and truth judgments, a message is more likely to be perceived as true when it is compatible with other things a person knows [42], and people may, furthermore, be motivated to persuade themselves to believe what they want to be true [18]. People have a "tendency to seek out, notice, and favor information that is consistent with one's attitudes, beliefs, and behaviors" [39], and if people encounter information that is in line (or concordant) with their political positions, they are presumably also more motivated to believe this information to be true partisan bias [43]. Concordant news items might therefore create blind spots for deception detection. People may continue to process concordant information intuitively and without effort (System 1), and if this processing is not interrupted (e.g., because something seems incoherent or out of place), falseness may be difficult to detect and truthfulness simply assumed [42].

Pennycook and Rand [23] investigated the impact of motivated reasoning and cognitive reflection on the detection of misinformation. They compiled data from 14 studies to investigate to what extent partisanship impacts truth ratings for concordant versus discordant actual news posts from social media (true and false). In addition, across studies, participants' cognitive reflection was assessed to measure analytical thinking. The authors show that partisanship is associated with overall belief in the news headlines (meaning belief in both true and in false headlines). Furthermore, concordant (versus discordant) news items are more likely to be judged as true. However, this effect is smaller than the effect of veracity: "true but politically discordant news is typically believed much more than false but politically concordant news—politics does not trump truth" [23,44]. The authors furthermore looked not only at overall belief but also at truth discernment. Truth discernment reflects the difference between the number of true items correctly identified as true (hits) and the number of false items incorrectly identified as true [45]. Described differently, truth discernment reflects people's ability to distinguish between true and false news [23]. Surprisingly, results show that

people are better at discernment when judging concordant versus discordant information.

Studies, furthermore, have shown that people with high (versus low) scores on the cognitive reflection test also show higher levels of truth discernment. Cognitive reflection seems important in telling truth from tales, and the authors argue that people's ability to detect misinformation might be endangered in situations where they are more likely to process information using System 1. Strengthening or activating reasoning, making people "think better even without thinking more" [46], can help people to pause and reflect on their prior knowledge, which may result in more informed and better judgment when it comes to telling truth from tale [45,47].

3. What Can News Providers, Social Media Platforms, and Individuals Do to Combat Misinformation?

Distinguishing fake from true news is difficult when people operate under the assumptions that mostly true information is communicated and when they process information rather effortlessly and intuitively. Some of the research reviewed here so far suggests that misinformation may be more likely detected when people process information under System 2 (compared to System 1). How can news providers and social media platforms create an environment in which people are more likely to stop and reason before they engage with news items?

One potential change in the environment could be to ask truth questions before users engage with news items. Pennycook et al. [48,49] tested this strategy in several studies. They assigned participants to an accuracy or a sharing condition and showed all participants true and false news items. In the accuracy condition, participants were asked to judge the truthfulness of the news items. Results suggest that participants perceived true statements as more likely to be true when compared to false statements. In the sharing condition, participants were asked to indicate their intention to share the news item. When sharing news, participants differentiated less between true and false news. This may suggest that judgments of truthfulness and sharing

are disconnected; put differently, people may share independently of judging truth.

This disconnection can be explained by inattention to accuracy (presumably due to intuitive and effortless information processing). The authors conducted further studies in which they created a combined condition: they asked participants to rate perceived accuracy for one news item before deciding whether or not to share it. Here, participants were less likely to consider sharing false news compared to the control group that was not prompted to evaluate accuracy before sharing [48]. The authors conclude that making accuracy more salient by asking people to focus on it reduces the likelihood that false information is shared.

This strategy can be applied in multiple ways. Users could be asked occasionally to rate the accuracy of a single, non-political headline before engaging with online content [48]. This approach would also allow platforms to collect crowd ratings on news items or news outlets that they could then use to find and fight misinformation [50,51]. Another strategy to increase more careful information processing could be to attach warnings to news items that fact-checkers have found to be false [23]—again, this would have the potential to increase the likelihood of reason-based decision making, as it would increase the salience of accuracy [50].

If these interventions could reduce the sharing of misinformation, much would be won—for news providers, social media platforms, and users alike. However, individuals might not even engage with the majority of news that they are exposed to on social media, but simply skim it, without clicking on the link to the full article [23]. Further strategies [23] are therefore required to generally reduce the amount of misinformation on these platforms.

4. Building a Truth Toolbox to Navigate Information in Everyday Life

In the previous section, we highlighted strategies that news providers and social media platforms could use to create a better environment to curtail the spread of misinformation. Individuals may also apply some of the recommendations to increase their own

abilities to critically approach and navigate information in the 21st century. Here, we illustrate potential strategies based on the research discussed above. We provide recommendations by means of four different questions, and further research may further test these and add to collections of potential measures to combat misinformation [52].

4.1. Question 1: Do I Believe the Piece of News to Be Accurate?

Before engaging with news content on social media, people may create a small boost, (see [53]) for themselves by actively asking whether they believe that a piece of news is accurate. This measure might prevent them from falling for a misinformation item, but also from liking or sharing it [48].

4.2. Question 2: Is There a Reason for Undue Familiarity?

When coming across a news headline that seems true, people may ask themselves why they think it is true. Does it feel familiar? Have they heard the message before? And if yes, do they remember where they came across that piece of information? In general, familiarity can serve as valid cue for truth, but having read a message over and over again can increase felt truth even though the message is false [11,25]. People may test against their impression of familiarity by checking the sources [5] and seeing whether the news is corroborated by different news outlets. They could further engage in lateral reading, meaning that they check other websites to test the credibility of the source of the news item that they are trying to evaluate [5,54].

4.3. Question 3: Is the News Plausible Even When Disregarding Potential Pictures?

Decorative but non-probative pictures may increase a news item's truthiness. To counteract, people may double check the news' content whenever it is accompanied by a non-probative picture.

The truthiness effect, furthermore, seems to be attenuated when people have high prior knowledge [27]. If no prior knowledge is available, people could check the source of both the news and the picture and investigate whether they can find similar support for the claim on different platforms to increase their own knowledge.

Warnings could seem a plausible strategy to prevent people falling for misinformation, but research indicates that while warnings may increase the accuracy when judging claims, they may not reduce the truthiness effect happening due to pictures [27]. Instructing people to ignore pictures may work, but this may be difficult to implement when reading news online [27]. Hopefully, the awareness of the truthiness effect of pictures may nevertheless help people to look at ambiguous information twice and will thereby increase people's chances to detect misinformation.

4.4. Question 4: Is the News Framed in a Way That Increases Truthiness?

A news headline may further be framed in a negative way to grab the readers' attention or to deliberately increase the chances that the messages is judged as true [55]. When encountering a headline that may seem very compelling, people may choose to reframe the message as a test. When reading "XX% of people are not satisfied with their looks", they might translate this to "100—XX% are satisfied with their looks" and investigate whether such a change in framing would still sound compelling. Furthermore, they could again test the information by looking at additional and different sources.

5. Conclusions

Societies have faced, are facing, and will face misinformation. While the distribution of misinformation has been slow and effortful in the past, the internet and social media have accelerated the speed with which news can be distributed and increased the number of people it can reach. This tendency is worrisome, as misinformation may lead to health-endangering behavior and erode trust in society and democracy. Measures to reduce misinformation and its spread are required. Furthermore, people's abilities to detect true and false information need to be strengthened so that they can better navigate online platforms and the landscape of news in the 21st century.

Author Contributions: M.J. conceptualized and wrote the original draft of the manuscript. R.G. supported the revision and editing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work received no external funding.

Acknowledgments: We thank Loris Jeitziner for his help with the literature search.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rathje, S.; Van Bavel, J.J.; Van der Linden, S. Out-group animosity drives engagement on social media. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2024292118. [CrossRef]
2. Guess, A.; Nyhan, B.; Reifler, J. Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign. 2018. Available online: https://www.american.edu/spa/ccps/politics-of-truth/upload/nyhan-american.pdf (accessed on 1 July 2021).
3. Lyons, B.A.; Merola, V.; Reifler, J. How bad is the fake news problem? The role of baseline information in public perceptions. In *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*; Greifeneder, R., Jaffé, M.E., Newman, E.J., Schwarz, N., Eds.; Routledge: London, UK, 2020; pp. 11–26.
4. Collins Dictionary. Fake News. In Collins Dictionary. 2021. Available online: https://www.collinsdictionary.com/dictionary/english/fake-news (accessed on 1 July 2021).
5. Lewandowsky, S.; Cook, J.; Ecker, U.K.H.; Albarracín, D.; Amazeen, M.A.; Kendeou, P.; Lombardi, D.; Newman, E.J.; Pennycook, G.; Porter, E.; et al. The Debunking Handbook 2020. *Databrary* **2020**, 1–19. [CrossRef]
6. Burkhardt, J.M. History of fake news. Combatting fake news in the digital age. *Libr. Technol. Rep.* **2017**, *53*, 5–9.
7. Loktionov, A. Ramesses II, victor of Kadesh: A kindred spirit of Trump? *The Guardian*, 5 December 2016.
8. Elster, J. Taming chance: Randomization in individual and social decisions. *Tann. Lect. Hum. Values* **1987**, 105–179.
9. Greene, C.M.; Murphy, G. Quantifying the effects of fake news on behavior: Evidence from a study of COVID-19 misinformation. *J. Exp. Psychol. Appl.* **2021**, *27*, 773–784. [CrossRef]
10. Ackland, R.; Gwynn, K. Truth and the dynamics of news diffusion on twitter. In *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*; von Rainer, G., Jaffé, M.E., Newman, E.J., Schwarz, N., Eds.; Routledge: London, UK, 2020; pp. 27–46. [CrossRef]

11. Pennycook, G.; Cannon, T.D.; Rand, D.G. Supplemental Material for Prior Exposure Increases Perceived Accuracy of Fake News. *J. Exp. Psychol. Gen.* **2018**, *147*, 1865–1880. [CrossRef]
12. Grice, H.P. Logic and conversation. In *Syntax and Semantics 3: Speech Arts*; Cole, P., Morgan, J.L., Eds.; Academic Press: New York, NY, USA, 1975; pp. 41–58.
13. Schweizer Presserat. *Journalistenkodex*, Schweizer Presserat: Bern, Switzerland.
14. Brashier, N.M.; Marsh, E.J. Judging Truth. *Annu. Rev. Psychol.* **2020**, *71*, 499–515. [CrossRef]
15. Epstein, S.; Pacini, R.; Denes-Raj, V.; Heier, H. Individual differences in intuitive–experiential and analytical–rational thinking styles. *J. Personal. Soc. Psychol.* **1996**, *71*, 390–405. [CrossRef]
16. Kahneman, D.; Frederick, S. Prepresentativeness revisited: Attribute substitution in intuitive judgment. In *Heuristics of Intuitive Judgment: Extensions and Applications*; Gilovich, T., Griffin, D., Kahneman, D., Eds.; Cambridge University Press: New York, NY, USA, 2002; pp. 1–30.
17. Kahneman, D. *Thinking, Fast and Slow*; Farrar, Straus and Giroux: New York, NY, USA, 2011.
18. Pennycook, G.; Rand, D.G. Why do people fall for fake news? *The New York Times*, 19 January 2019.
19. Bago, B.; Rand, D.G.; Pennycook, G. Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *J. Exp. Psychol. Gen.* **2020**, *149*, 1608–1613. [CrossRef]
20. Martel, C.; Pennycook, G.; Rand, D.G. Reliance on emotion promotes belief in fake news. *Cogn. Res. Princ. Implic.* **2020**, *5*, 47. [CrossRef]
21. Reber, R.; Schwarz, N. Effects of perceptual fluency on judgments of truth. *Conscious. Cogn.* **1999**, *8*, 338–342. [CrossRef] [PubMed]
22. Dechêne, A.; Stahl, C.; Hansen, J.; Wänke, M. The truth about the truth: A meta-analytic review of the truth effect. *Personal. Soc. Psychol. Rev.* **2010**, *14*, 238–257. [CrossRef] [PubMed]
23. Pennycook, G.; Rand, D.G. The Psychology of Fake News. *Trends Cogn. Sci.* **2021**, *25*, 388–402. [CrossRef] [PubMed]
24. Hasher, L.; Goldstein, D.; Toppino, T. Frequency and the conference of referential validity. *J. Verbal Learn. Verbal Behav.* **1977**, *16*, 107–112. [CrossRef]

25. Unkelbach, C.; Greifeneder, R. Experiential fluency and declarative advice jointly inform judgments of truth. *J. Exp. Soc. Psychol.* **2018**, *79*, 78–86. [CrossRef]
26. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false online. *Science* **2018**, *359*, 1146–1151. [CrossRef]
27. Newman, E.J.; Zhang, L. Truthiness: How non-probative photos shape belief. In *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*; Greifeneder, R., Jaffé, M.E., Newman, E.J., Schwarz, N., Eds.; Routledge: London, UK, 2020; pp. 90–114.
28. Newman, E.J.; Garry, M.; Bernstein, D.M.; Kantner, J.; Lindsay, D.S. «Nonprobative photographs (or words) inflate truthiness». *Psychon. Bull. Rev.* **2012**, *19*, 969–974. [CrossRef]
29. Fenn, E.; Ramsay, N.; Kantner, J.; Pezdek, K.; Abed, E. Nonprobative Photos Increase Truth, Like, and Share Judgments in a Simulated Social Media Environment. *J. Appl. Res. Mem. Cogn.* **2019**, *8*, 131–138. [CrossRef]
30. Baumeister, R.F.; Bratslavsky, E.; Finkenauer, C.; Vohs, K.D. Bad is stronger than good. *Rev. Gen. Psychol.* **2001**, *5*, 323–370. [CrossRef]
31. Rozin, P.; Royzman, E.B. Negativity bias, negativity dominance, and contagion. *Personal. Soc. Psychol. Rev.* **2001**, *5*, 296–320. [CrossRef]
32. Alves, H.; Koch, A.; Unkelbach, C. Why good is more alike than bad: Processing implications. *Trends Cogn. Sci.* **2017**, *21*, 69–79. [CrossRef] [PubMed]
33. Hilbig, B.E. Sad, thus true: Negativity bias in judgments of truth. *J. Exp. Soc. Psychol.* **2009**, *45*, 983–986. [CrossRef]
34. Hilbig, B.E. Good things don't come easy (to mind): Explaining framing effects in judgments of truth. *Exp. Psychol.* **2012**, *59*, 38–46. [CrossRef]
35. Jaffé, M.E.; Greifeneder, R. Less than I expected and oh so true? On the interplay between expectations and framing effects in judgments of truth. *J. Lang. Soc. Psychol.* **2019**, *38*, 735–755. [CrossRef]
36. Jaffé, M.E.; Greifeneder, R. Negative is true here and now, but not so much there and then. On the impact of psychological distance on the negativity bias. *Exp. Psychol.* **2020**, *67*, 314–326. [CrossRef] [PubMed]
37. Jaffé, M.E.; Greifeneder, R. Negative or Negated, thus True? An Investigation of Concept Valence and Semantic Negation as Drivers of Framing Effects in Judgments of Truth. *Soc. Cogn.* **2021**, *39*, 687–716. [CrossRef]
38. Tolstoi, L. *Anna Karenina*; Carl Hanser Verlag: München, Germany, 2009.

39. Galdi, S.; Gawronski, B.; Arcuri, L.; Friese, M. Selective exposure in decided and undecided individuals: Differential relations to automatic associations and conscious beliefs. *Personal. Soc. Psychol. Bull.* **2012**, *38*, 559–569. [CrossRef]
40. Stroud, N.J. Selective exposure theories. In *The Oxford Handbook of Political Communication*; Edited by Kenski, K., Jamieson, K.H., Eds.; Oxford University Press: Oxford, UK, 2014; pp. 1–21. [CrossRef]
41. Waldrop, M.M. The genuine problem of fake news. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12631–12634. [CrossRef]
42. Schwarz, N.; Jalbert, M. When (fake) news feels true: Intuitions of truth and the acceptance and correction of misinformation. In *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*; Greifeneder, R., Jaffé, M.E., Newman, E.J., Schwarz, N., Eds.; Routledge: London, UK, 2020; pp. 73–89.
43. Van Bavel, J.J.; Pereira, A. The Partisan Brain: An Identity-Based Model of Political Belief. *Trends Cogn. Sci.* **2018**, *22*, 213–224. [CrossRef]
44. Pennycook, G.; Rand, D.G. Lack of Partisan Bias in the identification of fake (versus real) news. *Trends Cogn. Sci.* **2021**, *25*, 725–726. [CrossRef]
45. Batailler, C.; Brannon, S.M.; Teas, P.E.; Gawronski, B. A Signal Detection Approach to Understanding the Identification of Fake News. *Perspect. Psychol. Sci.* **2022**, *17*, 78–98. [CrossRef] [PubMed]
46. Lin, H.; Pennycook, G.; Rand, D.G. Thinking more or thinking differently? Using drift-diffusion modeling to illuminate why accuracy prompts decrease misinformation sharing. *Cognition* **2023**, *230*, 105312. [CrossRef] [PubMed]
47. Gawronski, B. Partisan bias in the identification of fake news. *Trends Cogn. Sci.* **2021**, *25*, 723–724. [CrossRef] [PubMed]
48. Pennycook, G.; Epstein, Z.; Mosleh, M.; Arechar, A.A.; Eckles, D.; Rand, D.G. Shifting attention to accuracy can reduce misinformation online. *Nature* **2021**, *592*, 590–595. [CrossRef] [PubMed]
49. Pennycook, G.; McPhetres, J.; Zhang, Y.; Lu, J.G.; Rand, D.G. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychol. Sci.* **2020**, *31*, 770–780. [CrossRef] [PubMed]
50. Pennycook, G.; Rand, D.G. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 2521–2526. [CrossRef]

51. Allen, J.; Arechar, A.A.; Pennycook, G.; Rand, D.G. Scaling up fact-checking using the wisdom of crowds. *Sci. Adv.* **2021**, *7*, eabf4393. [CrossRef]
52. Kozyreva, A.; Lorenz-Spreen, P.; Herzog, S.; Ecker, U.; Lewandowsky, S.; Hertwig, R.; Basol, M.; Berinsky, A.J.; Betsch, C.; Cook, J.; et al. Toolbox of Interventions Against Online Misinformation and Manipulation. *PsyArXiv Preprints* **2022**, 1–24. [CrossRef]
53. Hertwig, R.; Grüne-Yanoff, T. Nudging and Boosting: Steering or Empowering Good Decisions. *Perspect. Psychol. Sci.* **2017**, *12*, 973–986. [CrossRef]
54. Breakstone, J.; Smith, M.; Connors, P.; Ortega, T.; Kerr, D.; Wineburg, S. Lateral reading: College students learn to critically evaluate internet sources in an online course. *Harv. Kennedy Sch. Misinformation Rev.* **2021**, 2. [CrossRef]
55. Jaffé, M.E.; Greifeneder, R. Can that be true or is it just fake news? New perspectives on the negativity bias in judgments of truth. In *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*; Greifeneder, R., Jaffé, M.E., Newman, E.J., Schwarz, N., Eds.; Routledge: London, UK, 2020; pp. 115–130. [CrossRef]

Doctor Facebook: Social Media and Health (Mis)Information

Jorge Revez

Abstract: Abstract: Using Facebook and other social media tools for information consumption is a pervasive habit. Health information is a relevant part of information activities and behavior, and its importance has been highlighted during the COVID-19 pandemic. The Facebook news feed has arisen as an information channel for many individuals who read and share health information, including fake news and material from sources of disinformation, for different purposes, disregarding basic assessment criteria or fact-checking strategies. This chapter intends to relate the current knowledge about online health information behavior through social media, focusing on Facebook, COVID-19, and the misinformation phenomenon. A literature review was performed using 51 recent works (published between 2020 and 2021). The main goals were to understand online information behavior in the social media setting; assess the landscape of fake or misinformed health information transmitted through Facebook; and evaluate the relationship between Facebook and health information during the COVID-19 pandemic. The results show two dominant research approaches: information behavior analysis, mostly acquisition and sharing activities, and online content analysis, focusing on infodemics, social behavior, and public health authorities' communication strategies. Despite serious concerns about misinformation, research has revealed a yet uncertain scenario regarding solutions to counter this public health issue.

1. Introduction

Health information is a relevant part of information-seeking behavior [1–4]. Among other subjects, health represents a major concern for citizens, but in most countries, healthcare access is not universal. Even a small piece of advice or a simple conversation with a healthcare professional is not easily achievable. In developed countries,

there is also the phenomenon of looking for health information on social media despite having better health systems. Therefore, many rely on online information to satisfy their information needs and are exposed to a diversity of health-related information, often without sufficient quality or accuracy. This scenario worsened with the emergence and development of the coronavirus pandemic: "The digital divide has become much more obvious as a result of the COVID-19 pandemic in 2020. With doctors' surgeries closed or operating only through telephone communication, and hospitals rescheduling even serious conditions because of the need to deal with the pandemic, more people turn to online sources of health information" [5] (p. 35).

These online sources reveal the relationship between information behavior and social media and their strong connection [6]. However, social media exposure may have a significant relationship with information overload as well as information anxiety [7]. This might help to explain the return to traditional mass media during the COVID-19 crisis [8]; an example is television, which has mostly been used by public authorities and health experts to convey hours' worth of what could be considered health information.

In the social media ecosystem, Facebook (founded in 2004) is the leading platform. According to Statista, "With roughly 2.89 billion monthly active users as of the second quarter of 2021, Facebook is the biggest social network worldwide" [9]. In the last ten years, Facebook has increased its monthly active users from 739 million (Q2 2011) to almost 3 billion. Recently, Facebook Transparency Center published a first quarterly report (Q2 2021) entitled Widely Viewed Content Report: What People See on Facebook [10]. This report "provides an overview of the content that reaches the most people on Facebook" [11] but only "captures views of public content in News Feed, including recommended content, seen in the United States" [10]. Not surprisingly, "most content views in News Feed during Q2 2021 came from posts shared by the friends, Groups and Pages people were connected to (for example, people they had chosen to add as a friend, Groups they had joined, Pages they followed and so on). Only about 9.5% of all News Feed content views in the US during Q2 2021 came from content that was not shared by Groups people joined, friends or Pages people

follow" [10]. In this report, Facebook reveals a somehow closed setting for information consumption, as most of the content viewed did not include a link to a source outside the platform: "the majority of News Feed content views in the US were on posts without links and were from content viewers' friends or from Groups they were connected to" [10]. Despite these findings, news sharing is an important social media behavior, as receiving valuable comments on shared news content leads to a sense of influence, and seeking others' opinions increases involvement more than sharing one's own opinion [12].

During the COVID-19 pandemic, online health information behavior became a major issue for governments, public health authorities, and healthcare professionals. The Facebook news feed (personal profiles, pages followed, and group activities) rose as an information channel for many individuals who read and share health information, including fake news and materials from disinformation sources, for different purposes, disregarding basic assessment criteria or fact-checking strategies. Fake news, disinformation, misinformation, infodemics, and other concepts, have emerged or re-emerged as an alert for authorities and academics. A previous review found that "social media platforms, although providing immense opportunities for people to engage with each other in ways that are beneficial, also allow misinformation to flourish" and that "there is broad consensus that misinformation is highly prevalent on social media and tends to be more popular than accurate information, while its narrative often induces fear, anxiety and mistrust in institutions" [13] (p. 8). The rising pandemic and global lockdowns, together with an increasing degree of online information consumption [14], brought together the ingredients for the degradation of the social media arena, as many users started producing and sharing low-quality content with obvious perils for public health, revealing the "co-existence and interaction of two domains: technology and health literacy" [15] (p. 525).

One of the strongest reactions against misinformation came from the World Health Organization (WHO) [16], which warned of an 'infodemic', mostly consisting of fake news [17]. An infodemic, "simply put, is an overabundance of information, good and bad. Together, it forms a virtual tsunami of data and advice that makes it hard

for people in all walks of life to find clear messages, trustworthy sources and reliable guidance when they need them. Some of it is merely confusing, but some of the misinformation can be actively harmful to life" [18] (p. 1). In April 2020, the WHO held a global online consultation on managing the COVID-19 infodemic. Among several specialists, Facebook representatives stated that the company's goal was "to support global public health work and keep people safe through a twofold strategy of (1) connecting people to accurate information from credible sources; and (2) stopping the spread of misinformation and other harmful content" [18] (p. 10). The Facebook strategy consisted of two parts: first, Facebook developed a coronavirus information center connecting user experience with credible sources of information; second, false claims, like cures or conspiracy theories, started being removed. In February 2021, after abundant social debate and a lot of criticism, Facebook announced additional efforts to remove more false claims on Facebook and Instagram about COVID-19 [19]. The COVID-19 policy updates and protections statement announced that the company would "remove misinformation when public health authorities conclude that the information is false and likely to contribute to imminent violence or physical harm" [20].

Besides these efforts, the major problem of misinformation circulation is that despite being false information, it "was not created with the intention of hurting others. Misinformation is often started by someone who genuinely wants to understand a topic and cares about keeping other people safe and well. It is then shared by others who feel the same. Everyone believes they are sharing good information—but unfortunately, they are not. And depending on what is being shared, the misinformation can turn out to be quite harmful" [21]. Much more dangerous is disinformation, that is, "false information created with the intention of profiting from it or causing harm. That harm could be to a person, a group of people, an organization or even a country. Disinformation generally serves some agenda and can be dangerous. During this pandemic, we are seeing it used to try to erode our trust in each other and in our government and public institutions" [21].

The development of fact-checking strategies and tips for navigating the infodemic [21] is important but faces a major challenge,

as "beliefs and values do occur in information behavior research, mainly in relation to health information" [5] (p. 36). The problem is not only about the truth but what we want (or do not want) to be the truth. Therefore, truth is no longer related to authority, expertise, or real facts but to interpretation, perception, emotions, and sentiments [22]. Post-truth [23] has emerged as a new information environment and behavior: "The overconsumption of information fueled by the internet has produced a so-called "post-truth" society in which people consume information that reaffirms their pre-existing beliefs and ideologies rather than attempting the difficult task of identifying the truth" [24] (p. 1).

The fake news phenomenon also represents a deeper crisis, as a crisis of truth is principally a crisis of trust [25,26]. Greater problems lie beneath the surface, like the politicization and weaponization of information, the traditional media crisis, and the technological incapacity to control the spread of misinformation [27]. One literature review provided the following typology of fake news: news satire, news parody, fabrication, manipulation, advertising, and propaganda [28]. Anstead refers to three forms of fake news—fake news as satire, fake news as misleading content, and fake news as populist rhetoric—but observes them as distinct responses to an ongoing and evolving crisis of democratic and media legitimacy [8]. These traditional and authoritative elements were challenged by the consequences of pluralism, which were somehow exacerbated by social media and its different fora of (free) speech. This is a paradox, as fake news is part of democratic life but is profoundly anti-democratic. The future management of fake news on social media sites will probably combine two approaches, namely, high-profile takedowns and discrete changes, thereby adjusting how content is presented and consumed by users [8].

Adopting the user's perspective and shifting away "from the structured information system and toward the person as a finder, creator, and user of information" [29] (p. 6), information behavior studies are useful for assessing this new landscape. Information behavior is "a shortened form of the behavior of humans in relation to information. It denotes how we act towards information, how we seek it or discover it, how we use it, how we exchange it with others, how we

may choose to ignore it, and, by extension, how we learn from it and act upon it" [5] (p. 14). It "encompasses information seeking as well as the totality of other unintentional or passive behaviors (such as glimpsing or encountering information), as well as purposive behaviors that do not involve seeking, such as actively avoiding information" [29] (p. 5).

Following previous research [13], this chapter intends to delineate the current knowledge about online health information behavior through social media, focusing on Facebook, COVID-19, and the misinformation phenomenon. Therefore, the main goals of this chapter are to understand online information behavior in the social media setting; assess the landscape of fake or misinformed health information transmitted through Facebook; and evaluate the relationship between Facebook and health information during the COVID-19 pandemic. To achieve these goals, the materials and methods used in this research are outlined in Section 2. Section 3 reports the main findings of the literature review, and Section 4 provides a discussion, followed by the conclusions of this research.

2. Materials and Methods

To analyze current research about online health information behavior exhibited on social media, a non-systematic literature review was performed. The first retrieval used a combination of search terms encompassing the primary concepts of health, Facebook, COVID-19, and fake news, disinformation, and misinformation. An attempt to include the concept of information-gathering behavior in a unique string yielded no results. Therefore, a screening process and analysis were afterward used to refine the dataset. Later, during full-text reading, the citations and references found pointed to other information sources, and the most relevant were included and reviewed.

2.1. Data Extraction

The literature was first retrieved on 6 September 2021, using the *Scopus* database, with regard to its importance, relevance, quality, and inclusiveness, for example, in comparison with *Web of Science* [30]. The scope of the literature review was not limited to one region. Also,

there were no limitations concerning the languages of publication, though the search terms were only used in English. The search strategy intended to combine the different research topics. The following query was applied to titles, abstracts, and keywords: health AND Facebook AND (fake OR disinform* OR misinform*) AND (COVID-19 OR coronavirus). This query returned 51 results, which corresponded to papers published between January 2020 and September 2021. The results were refined to include only articles, reviews, and conference papers, as there were no book chapters within the results, thus yielding 48 results. These were exported to an MS Excel spreadsheet. The 48 results included 39 articles, 5 conference papers, and 4 reviews. Afterward, titles and abstracts were reviewed, and the following exclusion criteria were applied: exclude all material not specifically related to Facebook or online information behavior. One of the papers reviewed was withdrawn by the author and therefore also excluded.

2.2. Final Dataset

The core set was reduced to 36 works (Table 1), published between 2020 and 2021, corresponding to empirical studies from the USA, the UK, Australia, Japan, Nigeria, Denmark, Brazil, Philippines, Iraq, and Gaza, among other countries. A global perspective is reflected in this core set; however, during full-text reading, other references emerged, forming a final dataset of 51 works reviewed. Nevertheless, the overall sample is only a part of all the literature published about these subjects, and it was hardly conditioned by the keywords initially used to form the core set of publications.

Table 1. Core set of publications for analysis (n = 36).

Title	Publication	Reference
"Ask a doctor about coronavirus": How physicians on social media can provide valid health information during a pandemic	*Journal of Medical Internet Research*	[31]
A Survey of COVID-19 Information Dissemination Behavior of Library and Information Professionals in Nigeria	*Library Philosophy and Practice*	[32]
Adaptation to SARS-CoV-2 under stress: Role of distorted information	*European Journal of Clinical Investigation*	[33]
Addressing COVID-19 misinformation on social media preemptively and responsively	*Emerging Infectious Diseases*	[34]
Artificial intelligence-enabled analysis of public attitudes on facebook and twitter toward COVID-19 vaccines in the United Kingdom and the United States: Observational study	*Journal of Medical Internet Research*	[35]
Awareness on spread of misinformation and its effect on public with regard to COVID-19	*International Journal of Current Research and Review*	[36]
COVID-19 misinformation: Accuracy of articles about coronavirus prevention mostly shared on social media	*Health Policy and Technology*	[37]
COVID-19 on Facebook Ads: Competing Agendas around a Public Health Crisis	*COMPASS 2020—Proceedings of the 2020 3rd ACM SIGCAS Conference on Computing and Sustainable Societies*	[38]

Table 1. *Cont.*

Title	Publication	Reference
COVID-19 vaccine rumors and conspiracy theories: The need for cognitive inoculation against misinformation to improve vaccine adherence	*PLoS ONE*	[39]
COVID-19, a tale of two pandemics: Novel coronavirus and fake news messaging	*Health Promotion International*	[15]
COVID-19-Related infodemic and its impact on public health: A global social media analysis	*American Journal of Tropical Medicine and Hygiene*	[17]
Detecting fake news on Facebook: The role of emotional intelligence	*PLoS ONE*	[40]
Disinformation and COVID-19: Quantitative analysis through the hoaxes debunked in Latin America and Spain [Desinformación y COVID-19: Análisis cuantitativo a través de los bulos desmentidos en Latinoamérica y España]	*Estudios Sobre el Mensaje Periodistico*	[41]
Exploring Sub-Saharan Africa's Communication of COVID-19-Related Health Information on Social Media	*Libri*	[42]
Fact or fake? An analysis of disinformation regarding the COVID-19 pandemic in Brazil [Fato ou Fake? Uma análise da desinformação frente à pandemia da COVID-19 no Brasil]	*Ciencia e Saude Coletiva*	[43]
Health information seeking behaviors on social media during the COVID-19 pandemic among american social networking site users: Survey study	*Journal of Medical Internet Research*	[44]

Table 1. *Cont.*

Title	Publication	Reference
How do Canadian public health agencies respond to the COVID-19 emergency using social media: A protocol for a case study using content and sentiment analysis	*BMJ Open*	[45]
How social media comments inform the promotion of mask-wearing and other COVID-19 prevention strategies	*International Journal of Environmental Research and Public Health*	[46]
In the midst of the Coronavirus pandemic in Brazil, watch out for snowmen in the north and northeast regions! Post-truth under discussion [Na pandemia brasileira, tá tendo boneco de neve no norte e nordeste do país! Pós-verdade em debate]	*Praksis*	[47]
Infodemic, Misinformation and Disinformation in Pandemics: Scientific Landscape and the Road Ahead for Public Health Informatics Research	*Studies in health technology and informatics*	[48]
Knowledge about COVID-19 in Brazil: Cross-sectional web-based study	*JMIR Public Health and Surveillance*	[49]
Learning about COVID-19: a qualitative interview study of Australians' use of information sources	*BMC Public Health*	[50]
Misinformation on social networks during the novel coronavirus pandemic: a quali-quantitative case study of Brazil	*BMC Public Health*	[51]

Table 1. *Cont.*

Title	Publication	Reference
Paying SPECIAL consideration to the digital sharing of information during the COVID-19 pandemic and beyond	*BJGP Open*	[52]
Peer influence, risk propensity and fear of missing out in sharing misinformation on social media during the COVID-19 pandemic	*ICCE 2020—28th International Conference on Computers in Education, Proceedings*	[53]
Public engagement and dialogic accounting through social media during COVID-19 crisis: a missed opportunity?	*Accounting, Auditing and Accountability Journal*	[54]
Reasons for rejecting official recommendations and measures concerning protection against SARS-CoV-2—a qualitative study of social media posts [Gründe für die Ablehnung behördlicher Empfehlungen und Maßnahmen zum Schutz vor SARS-CoV-2—eine qualitative Studie auf Basis von Beiträgen in sozialen Medien]	*Bundesgesundheitsblatt—Gesundheitsforschung—Gesundheitsschutz*	[55]
ReOpen demands as public health threat: a sociotechnical framework for understanding the stickiness of misinformation	*Computational and Mathematical Organization Theory*	[56]
Social media and the COVID-19 pandemic: Observations from Nigeria	*Cogent Arts and Humanities*	[57]
Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases	*Human Vaccines and Immunotherapeutics*	[58]

Table 1. *Cont.*

Title	Publication	Reference
The bright and dark sides of social media usage during the COVID-19 pandemic: Survey evidence from Japan	*International Journal of Disaster Risk Reduction*	[59]
The impact of COVID-19-related changes in media consumption on public knowledge: results of a cross-sectional survey of Pennsylvania adults	*Current Medical Research and Opinion*	[60]
The impact of social media on panic during the COVID-19 pandemic in iraqi kurdistan: Online questionnaire study	*Journal of Medical Internet Research*	[61]
The information-seeking behavior and levels of knowledge, precaution, and fear of college students in Iloilo, Philippines amidst the COVID-19 pandemic	*International Journal of Disaster Risk Reduction*	[62]
The role of social media in spreading panic among primary and secondary school students during the COVID-19 pandemic: An online questionnaire study from the Gaza Strip, Palestine	*Heliyon*	[63]
User motivation in fake news sharing during the COVID-19 pandemic: an application of the uses and gratification theory	*Online Information Review*	[64]

Source: Table by author.

3. Results

The core dataset revealed two main thematic research lines. On one side, there were studies concerning information behavior analysis, focusing on individual or group behavior; on the other side, there was research on online content analysis, focusing on web content, posts, commentaries, or other kinds of online publications. Some works

could be related to both perspectives. This two-fold approach was also used as a framework for the inclusion of other literature beyond the core set of publications.

3.1. Information Behavior Analysis

3.1.1. Information Acquisition

Has online information acquisition increased since the beginning of the COVID-19 pandemic? The previous trend of the prevalence of social media sources seems to coexist with a trend of returning to more traditional sources, like television, during a health crisis [7]. In Australia, analysis of information acquisition behavior revealed "that participants were active users of information sources rather than passively accepting news accounts, government spokespeople or social media content as authoritative". Increasing levels of trust in official information contrast with feelings of anger and frustration "about the extent of misinformation that was circulating in the community and online and the potential for it to contribute to the spread of the coronavirus and pose a risk to others" [50] (pp. 8–9). This suggests some sort of Facebook avoidance, as one participant reported that the "initial joking on Facebook was countered by the dramatic television news reports of the growing threat posed by COVID-19 to Australians", and another mentioned that "I must admit I've become quite careful about reading conspiracy-type theories on Facebook. Yeah, it's a platform for everyone to have their say, but I've discovered that in my own opinion, some theories are quite farfetched. People can be sincerely wrong" [50] (pp. 4–6). One of the mentioned conspiracy theories was a relationship between coronavirus and 5G [65].

With different results, a US study showed a heavy reliance on social media during COVID-19 with respect to information acquisition behavior together with fewer fact-checking actions. These findings "highlight the increasing importance of social media in health information seeking and thus highlight its potential value to health professionals as a conduit for personal and public health communications". The authors recommend ensuring "more active engagement between health professionals and patients and

consumers" [44] (pp. 6–7). Another US study confirmed an increase in the consumption of news from the internet versus television, but "adults whose most trusted information source is government health websites are more likely to correctly answer questions about COVID-19 than those with another most trusted source. Individuals whose most trusted source is television news and those who use Facebook as an additional source of news are less likely to correctly answer COVID-19 questions" [60] (p. 4). Despite having a scope limited to literate people who use social networks, a study conducted in Brazil recruited participants on several social media platforms, including Facebook, and showed "satisfactory knowledge about COVID-19 when true information and fake news are mixed" and an ability "to differentiate the two types of information", suggesting that the impact of fake news on the knowledge of COVID-19 was limited [49] (p. 18).

In Japan, a social media user study demonstrated the bright and dark sides of information acquisition through digital platforms, including Facebook: "The bright side is that it encourages users to take protective measures officially endorsed by the government based on scientific evidence, such as social distancing and use of disinfectants"; simultaneously, "users take measures which are not grounded in scientific evidence, such as eating fermented soybeans" [59] (p. 5). In developed countries, the problem is not the lack of reliable information but the exposure "to rumors from unreliable sources that contain misinformation, presenting an obstacle to appropriate responses [to emergencies]" [59] (p. 5).

Social media use appears to relate to serious mental health issues. Facebook has been associated with fear levels regarding contracting COVID-19. In the Philippines, "fear levels were higher (severe or extreme) among college students who preferred Facebook as a source of COVID-19 information", thus revealing "the adverse effects of social media on the mental and psychological well-being of individuals facing threats caused by epidemics, such as worry, depression, anxiety, anger and fear" [62] (p. 11). Social media consumption seems to have a direct impact on mental health. In Iraq, Facebook was the most used social media network for spreading panic about the COVID-19 outbreak, causing high levels of psychological anxiety and uncertainty

about the true/false information disseminated online [61]. In Gaza, similar results indicated that "social media has a significant effect on spreading panic about the COVID-19 pandemic among students, with a potential negative impact on their mental health and psychological well-being" [63] (p. 9). In India, similar mental health problems were reported [36].

Social media seems also to harm health knowledge acquisition, but are Facebook groups' activities mitigating the problem? In an attempt to counterbalance mental health impacts and the negative consequences of misinformation and foster feelings of calmness, trust, and safety, a Facebook group was created in March 2020 in Denmark to provide a direct relationship between more than 200 volunteer physicians and the public. Entitled "Ask a Doctor About Coronavirus", the group has a rigorous moderation policy, managing inappropriate questions and the dissemination of fake or misinformed facts. This group, with 30.000 daily active users, "provides a proof of concept of a new way for health professionals to communicate and interact with the general public on social media platforms" and reveals "unique insights into the potential of Facebook in health communication; however, we cannot ignore the possibility of the distinctive information-seeking environment of the COVID-19 pandemic providing a favorable foundation for dissemination and upscaling of information" [31] (p. 4). Several studies analyzed Facebook groups' activities, and one identified four key information activities: posting, monitoring, commenting, and searching. As a small network within a larger one, "people report that they are more likely to find information that is more relevant to their specific information needs through these groups than through personal networks" [66] (p. 213).

Another strategy for fighting misinformation was the publication of infographics to debunk coronavirus myths. These shareable materials could provide a good prevention/reaction tool in the social media setting. A test of the efficacy of these infographics—namely, one produced by the WHO about hot baths that raise body temperature and prevent coronavirus infection—indicated that "preemptively sharing these graphics can be effective. Users and organizations can debunk misinformation circulating in society by sharing high-quality

information on social media emphasizing the facts without waiting to see it shared directly in their feeds, which expands the opportunities for observational correction to occur" [34] (p. 402). The authors suggest the development of a WHO bot that can respond directly to pieces of misinformation, with a capacity beyond the previous experience of a Facebook Messenger chatbot, that is, a version of the "WHO Health Alert platform—offering instant and accurate information about COVID-19—via Facebook's global reach" [67].

3.1.2. Information Sharing

Sharing social media content is nowadays a behavior similar to content creation. The decision to share is often quick, and its motivation is not easily acknowledged in information behavior research. The decision to share fake news or misinformed content represents another level of complexity, but a US study found that discernment and a simple accuracy judgment are key factors that can interrupt a sharing chain [68]. University students from the Philippines revealed that peer influence and fear of missing out have a positive influence on the behavioral intention to share misinformation on social media: on "Facebook, content posted by people within a student's social network can influence the decision process to re-share such information" [53] (p. 355). Sharing is not only an individual decision; it is also a social act. To prevent misinformation sharing, authors from the University of Oxford in the UK proposed "a framework to help us be strategic and choose wisely, by paying SPECIAL consideration to the information we share" [52]. This framework recommends attention to source, privacy, evaluation, contribution, intention, audience, and legacy aspects to improve the quality of information shared by researchers and health professionals, but it could also be adopted by a more general audience.

In Nigeria, Facebook and WhatsApp users' motivation for sharing fake news was also analyzed. The social dimension of sharing information in order to help others was again highlighted, as "altruism and instant news sharing had more effect on fake news sharing behaviour compared to socialisation and self-promotion" and "entertainment had no association with fake news sharing on COVID-19" [64] (p. 13). Instant news-sharing behavior is a

consequence of the rapid publication that is enabled by digital technologies, causing inadvertent fake news dissemination. Also in Nigeria, research on library and information professionals indicated that this group "disseminated COVID-19 information, fact-checked, corrected misinformation and provided update information majorly through Facebook and WhatsApp channels" [32] (p. 24), reinforcing the informational role of librarians in times of emergency [69]. Another study involving Nigerian undergraduate students identified a positive effect of information literacy competence on curtailing the spread of fake news about the COVID-19 pandemic [70].

Regarding the psychological dimension of misinformation-reacting behavior, a UK-based study explored the association between fake news detection on Facebook, emotional intelligence, and educational attainment, concluding "that individuals who are high in emotional intelligence and who are in receipt of a university education are less likely to fall for fake news than low EQ/School-College educated individuals". These findings suggest that if a user's level of emotional intelligence was rated simply by looking at their Facebook data, "then Facebook could alert low EQ scores that they should be more vigilant about misinformation and fake news that might appear on their platform" [40] (p. 10). In Finland, researchers proposed a model for understanding the effect of information seeking, information sources, and information overload on information anxiety, resulting in information avoidance. Their findings suggest that "individuals who have more exposure to social media sources were more likely to feel information overload and information anxiety during health crisis", and "an individual's level of information anxiety has a significant positive impact on the level of information avoidance" [7] (p. 12). The development of health literacy or eHealth literacy skills, together with information literacy skills, is proposed as one of the possible solutions [71].

3.2. Online Content Analysis

3.2.1. Infodemic

The infodemic situation is probably one of the most interesting phenomena of the COVID-19 period. In Russia, a study combined online content analysis and a survey of healthcare professionals to draw attention to an informational problem surrounding COVID-19 crisis management. The stress, fear, and anxiety experienced by medical personnel were related to a social background of nervousness, mostly derived from the infodemic phenomenon. A content analysis of the personal accounts of a Russian-speaking audience provided on social networks "exhibited high degrees of anxiety and even panic brought about by distorted, exaggerated or false information on COVID-19 disease and its causative agent" [33] (p. 3). The author concluded that the "exaggerated and distorted information on COVID-19 has big negative influence upon Russian society and healthcare system" [33] (p. 5), thus undermining medical workers' performance.

One of the dangerous components of an infodemic is misinformation. In an extensive analysis, an international team identified 2.311 reports of rumors, conspiracy theories, and stigma related to the COVID-19 infodemic in 25 languages from 87 countries. A large majority of the reports were false, which means that these global waves of misinformation had serious consequences for public health. Rumors were the largest category detected, and some examples are provided in the following: "eating garlic, keeping the throat moist, the need to avoid spicy food, and the importance of taking vitamins C and D to help prevent the disease" and "so-called treatments such as miracle mineral solutions that involved mixing sodium chlorite solution with citric acid or drinking bleach or alcohol for immunity and cures" [17] (p. 1622). People died or became sick following the dissemination of these pieces of misinformation; consequently, the pervasive nature of social media consumption has become a significant public health issue. In addition, protective measures to prevent the COVID-19 pandemic were also undermined by misinformation, fueling rejective attitudes through the dissemination of doubts and mistrust in public authorities or science expertise. Following the economic

consequences of these protective measures, e.g., lockdowns and social distancing, a German study found that social media misinformation may contribute to underestimating the pandemic [55]. A Nigerian study found a relevant role of religious practices that "significantly influence the spread of false preventive measures of the coronavirus disease" [57] (p. 7). This influence, together with the increasing use of social media, worried public authorities and led to a strong strategy for information dissemination, including daily updates through SMS.

Regarding rumors and conspiracy theories about COVID-19 vaccines, for almost the entirety of 2020, Facebook-only content represented about half of the 637 pieces of misinformation identified in 24 languages from 52 countries, which reveals the social relevance of the platform and the amount of misinformation in circulation [39]. Misinformation content categories, as a product of human actions, adapt to the course of events: "During the initial months of COVID-19, most of the COVID-19 vaccine claims were related to pre-pandemic vaccine and conspiracy theories. More recent claims were related to efficacy and effectiveness of the vaccine, morbidity, and mortality due to participation in the vaccine trial" [39] (p. 5). An early-pandemic literature review identified social media as a terrain for anti-vaccination messaging possibly leading to COVID-19 vaccine hesitancy, comparing previous anti-vaxxer movements and this ideology's online dissemination [58]. However, a study developed using an artificial-intelligence-based approach to analyze public sentiments on social media in the UK and the US toward COVID-19 vaccines showed mostly positive attitudes [35].

As Facebook remains the leading platform for sharing content, information accuracy is naturally a relevant issue. The 30 most frequently shared articles in April 2020 about COVID-19 prevention were analyzed, and Facebook accounted for the most shares [37]. Most of the articles were considered accurate yet less likely to be shared, which may strengthen the argument that false information is spreading online more easily than truth [72]. Facebook ads are also a relevant means for information dissemination. Despite some efforts to control this content [20], researchers have found traces of misinformation "ranging from bioweapons conspiracy theories to unverifiable claims

by politicians, to the sale of face masks which may not necessarily protect the wearer" [38] (p. 22).

Many of the misinformed stories on social media are fact-checked by media outlets or independent organizations. A quantitative approach analyzing Spain and Latin America concluded that most of the examined stories were false, and Facebook accounted for half of the share of dissemination. Half of the fact-checked sample was health-related information [41]. Fact-checking systems or platforms remain one of the technological solutions to combatting misinformation and fake news. In Brazil, an application entitled Eu Fiscalizo (I check) received users' notifications of circulating fake news about COVID-19. The main dissemination channel was WhatsApp, followed by Facebook. Regarding content, "65% of them taught homemade methods to prevent the spread of COVID-19; 20% showed homemade methods to cure the disease; 5.7% referred to banking scams; 5% mentioned scams on fundraising for a research institution; and 4.3% concerned the use of the New Coronavirus as a political strategy" [43] (p. 4204). A similar study observed misinformation data from the Brazilian fact-checking service Lupa Agency, most of which was disseminated through Facebook. The research found that 92.9% of the misinformation classified as fabricated content consisted of health tips, and "43.8% of the pieces of misinformation classified as misleading/imposter/manipulated have a scientific/epidemiological content" [51] (p. 8).

Infodemics will be an important research field in the next years. In an attempt to determine the structural components of infodemics, a research team proposed a taxonomy of fake news. First, two separate groups were identified: health- and non-health-related fake news. Health-related fake news (more than 60%) "included fake prevention, fake medicines and treatments, fake SARS-CoV-2 information, hoax or fake pandemic and anti-vaccine messages, and home remedies recommendations". The non-health-related group "included messages generating confusion, phishing/scams, conspiracy theories, political propaganda, pseudo-science and 5G theories along with anti-Bill Gates messages" [15] (p. 527). A bibliometric study depicted a concept mapping of infodemic literature and proposed future research directions, namely, infodemic fabrication context and evaluation,

digital tools and the agency of actors, infodemic crisis management, infodemic knowledge and mental health impacts, and the study and modeling of infodemic trends and topics [48].

3.2.2. Social Behavior

The use of social media to convey messages against public authorities or science expertise is not new, but the COVID-19 pandemic was an optimal terrain for its development. One example is The Reopen the States Movement (US). Observing their coronavirus-skeptical Facebook groups, one study revealed a strong connection between health misinformation and political action, namely, conservative values: "Since members did not trust mainstream media, government, or public health reporting on the COVID pandemic, many returned to the decidedly hermeneutical and conservative scriptural inference process to "do their own research" on the numbers and then share their own individualized, unverified, and often contradictory interpretations of COVID data with their ideological social networks" [56] (p. 11). A qualitative analysis of the content shared in the groups demonstrated the assumption of a serious public health threat.

Facebook was also a space for reactions against preventive measures. When observing 615 Facebook comments drawn from Montana (US) news sources, it was revealed that 63% presented resistance to complying with the mask-wearing measure, as "barriers to compliance with COVID safety precautions are related to both low perceived risk of COVID and low perceived efficacy of the prevention measures" [46] (p. 11). These findings also suggest that such resistance arises not from an informational or misunderstanding-related problem but from conspiracy theories and other social or political perspectives, many of which were already evident before the pandemic. Public health communication should move beyond scientific evidence and adapt itself to fight against previous biases. However, social media echo chambers are not easily disrupted, and "the current atmosphere of distrust and anti-scientific sentiment have undermined the ability of health educators to use informational messages to fight the COVID-19 pandemic" [46] (p. 14).

3.2.3. Public Health Authorities

The COVID-19 crisis forced a rapid change in the way health authorities communicate and engage with the public. An analysis of the official Facebook pages of the leading public agencies for health crises in Italy, the United Kingdom, and New Zealand found that social media was extensively used: "Public agencies mainly released information on the evolution of the epidemic, the correct behaviour to empower citizens in protecting from the virus and the actions taken to face the situation" [54] (p. 11). However, a missed opportunity was depicted, as some countries were unable to establish dialogic communication, possibly to avoid organizational damages or because of a lack of competencies and resources. Public engagement through social media is not sufficient in a pandemic crisis, as "public agencies establishing an active dialogic communication have a higher level of the tenor of comments index, with a lower incidence of contestation or posts reporting fake news" [54] (p. 11).

For a different perspective, different reactions arose on Facebook when the Brazilian Minister of Health established a connection between winter and the increasing severity of the pandemic in the north and northeast of Brazil. Analyzing Facebook comments and memes, a study found that humor was used to counterbalance post-truth discourse and dismantle misinformed political positions, presenting social media as a space of questioning not only political decisions but also political discourse [47]. Further research is expected on public health authorities' performance regarding social media, e.g., Canadian agencies are being studied using content and sentiment analysis [45].

Most developing countries were not prepared to face a pandemic that demanded new forms of communication. Extensive research concerning the public health authorities of 23 anglophone Sub-Saharan African countries demonstrated that a presence on social media, mostly on Facebook, is not enough. To convey trusted information and fight against misinformation, authorities must have social media legitimacy (e.g., verified accounts), the capacity to reach audiences, strategies, and expertise [42].

4. Discussion

The literature reviewed shows the global nature of the health misinformation problem and its impact on a platform as popular as Facebook [13]. The various countries addressed in the different studies also reveal the efforts of scientific research in addressing the misinformation problem and the concerns raised by public health issues, which have been aggravated by the COVID-19 pandemic. Both user information behavior perspectives and content analysis approaches, focusing on published content as the result of users' interactions with social media, converge to indicate the relevance of research being conducted around the world.

This review has demonstrated a bright and a dark side of online health information [59]. From the results emerge certainties and uncertainties, perhaps more uncertainties than certainties, as the short space of time and the speed of events do not make it possible to have a completely clear view of the various social and communicational phenomena that are occurring.

As for certainties, the results show that the impacts of misinformation exposure and the associated risks in the social media context are multiple following the increase in online information consumption [5,14]. Information is abundant and satisfies most users' needs; however, social media exposes people to several risks. Despite the revival of traditional media [8], this review has shown social impacts that generate fear and panic as well as increase and worsen the social perception of the pandemic emergency. Myths, conspiracy theories, false cures, and rumors of vaccines' ineffectiveness are components of a global wave of low-quality information, often conflicting with the official perspective of the authorities (including the WHO), who have reacted and showed concern about the infodemic [16,17]. Studies confirm political impacts, as misinformation reveals previous tensions, and it has also fueled a feeling of revolt against the authorities and undermined the sense of trust that should unite society and its leaders, especially during a pandemic that has killed millions of people worldwide [8,25,27]. Also, there were impacts on health, as misinformation led to risky behavior due to following content without

any scientific basis, or, on the other hand, a disregard for the seriousness of the pandemic and the information transmitted by official channels. Impacts on mental health, with reports of information overload and information anxiety [7], were also pointed out.

As for the uncertainties, research shows that there are still many problems to acknowledge and to help solve. It is not yet clear what the most effective methods to fight misinformation are and how society can control the role of social media, in which people behave based on freedom of expression and in a post-truth environment, where feelings and emotions come up against truth and facts [22,23]. It is also still uncertain how public health authorities can improve their way of communicating with the public, especially with respect to making use of social media, thus coexisting with the public in the same digital ecosystem. It is also uncertain how platforms such as Facebook will improve their systems to combat misinformation. At this point, it should be noted that the various impacts on mental health are among the aspects that social media companies should consider and on which they should act in the future. Another deep crisis is in the process of becoming serious. It is also uncertain if health literacy is a solution for misinformation, because, on the one hand, information behavior studies show that the way people share information has emotional components, which are difficult to address in the development of literacy skills; on the other hand, misinformation is often consumed within an echo chamber, helping to confirm what people wanted to have confirmed [24]. Finally, it is also uncertain whether countries will be able to develop effective legal mechanisms to counter misinformation, a global phenomenon of unknown authorship that is viral and without borders.

Future research should discuss the role of user education and the ability to prepare individuals against these pitfalls. Other solutions to preventing disinformation and fake news should also be researched, as health information is critical to decision making and public health.

5. Conclusions

This chapter sought to explore the most recent research on the problem of misinformation in the context of social media. The

case of Facebook was chosen due to its social relevance and large community of users. The scope was also focused on the pandemic period, brought about by COVID-19, to try to isolate and observe the analyzed phenomenon's acceleration that, although not new, has worsened since the beginning of 2020. In this sense, health was the main theme analyzed.

Recent research has shown the danger of trusting social media as a source of health information. Nevertheless, the popularity and massive use of Facebook pose a challenge to health literacy, as the quantity of information channeled make some sort of human cognitive or automatic machine-based counter-reactions almost impossible. Even fact-checking strategies appear to have several limitations.

The results show two dominant research approaches: information behavior analysis, mostly acquisition and sharing, and online content analysis, focusing on infodemics, social behavior, and public health authorities' communication strategies. Despite serious concerns about misinformation, researchers have revealed an uncertain scenario regarding solutions to countering this public health issue.

The main research goals were achieved. Regarding online information behavior in the social media setting, the results demonstrated its key active features, like acquisition and sharing, along with the passive ones, such as being subject to mental health issues, like anxiety and fear. This review assessed the landscape of fake or misinformed health information transmitted through Facebook, showing its diversity and impact. The relationship between Facebook and health information during the COVID-19 pandemic was strong, as the most popular platform was largely used for health information activities (ranging from support groups promoted by healthcare professionals to groups of activists against preventive protection measures), revealing the bright and the dark sides of social media.

The main limitation of this study concerns the way the first information retrieval process was carried out since it is certain that a different query could have led the review along other paths. Thus, this review should be read with caution and attention paid to the methodological assumptions presented.

Online health (mis)information remains an open research topic and a global challenge, as it does not seem that the associated problems will diminish or be mitigated in the coming years.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Pang, P.C.-I.; Verspoor, K.; Chang, S.; Pearce, J. Conceptualising health information seeking behaviours and exploratory search: Result of a qualitative study. *Health Technol.* **2015**, *5*, 45–55. [CrossRef]
2. Cao, W.; Zhang, X.; Xu, K.; Wang, Y. Modeling Online Health Information-Seeking Behavior in China: The Roles of Source Characteristics, Reward Assessment, and Internet Self-Efficacy. *Health Commun.* **2016**, *31*, 1105–1114. [CrossRef]
3. Jacobs, W.; Amuta, A.O.; Jeon, K.C. Health information seeking in the digital age: An analysis of health information seeking behavior among US adults. *Cogent Soc. Sci.* **2017**, *3*, 1302785. [CrossRef]
4. Wang, X.; Shi, J.; Kong, H. Online Health Information Seeking: A Review and Meta-Analysis. *Health Commun.* **2021**, *36*, 1163–1175. [CrossRef]
5. Wilson, T.D. *Exploring Information Behaviour: An Introduction*; Preliminary Edition; Author Edition; 2020.
6. Mastley, C.P. Social Media and Information Behavior: A Citation Analysis of Current Research from 2008–2015. *Ser. Libr.* **2017**, *73*, 339–351. [CrossRef]
7. Soroya, S.H.; Farooq, A.; Mahmood, K.; Isoaho, J.; Zara, S. From information seeking to information avoidance: Understanding the health information behavior during a global health crisis. *Inf. Process. Manag.* **2021**, *58*, 102440. [CrossRef]
8. Anstead, N. *What Do We Know and What Should We Do About Fake News?* Sage: London, UK, 2021; ISBN 978-1-5297-1789-1.
9. Statista Facebook MAU Worldwide 2021. Available online: https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ (accessed on 8 September 2021).
10. Facebook Q2'21 Widely Viewed Content Report: What People See on Facebook|Transparency Center. Available online: https://transparency.fb.

com/pt-pt/data/widely-viewed-content-report/ (accessed on 8 September 2021).
11. Facebook Transparency Center. Available online: https://transparency.fb.com/pt-pt/data/widely-viewed-content-report/companion-guide/ (accessed on 8 September 2021).
12. Oeldorf-Hirsch, A.; Sundar, S.S. Posting, commenting, and tagging: Effects of sharing news stories on Facebook. *Comput. Hum. Behav.* **2015**, *44*, 240–249. [CrossRef]
13. Wang, Y.; McKee, M.; Torbica, A.; Stuckler, D. Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. *Soc. Sci. Med.* **2019**, *240*, 112552. [CrossRef]
14. Koetsier, J. Global Online Content Consumption Doubled in 2020. Available online: https://www.forbes.com/sites/johnkoetsier/2020/09/26/global-online-content-consumption-doubled-in-2020/ (accessed on 8 September 2021).
15. Atehortua, N.A.; Patino, S. COVID-19, a tale of two pandemics: Novel coronavirus and fake news messaging. *Health Promot. Int.* **2021**, *36*, 524–534. [CrossRef]
16. World Health Organization Call for Action: Managing the Infodemic. Available online: https://www.who.int/news/item/11-12-2020-call-for-action-managing-the-infodemic (accessed on 17 September 2021).
17. Islam, M.S.; Sarkar, T.; Khan, S.H.; Kamal, A.-H.M.; Murshid Hasan, S.M.; Kabir, A.; Yeasmin, D.; Islam, M.A.; Chowdhury, K.I.A.; Anwar, K.S.; et al. COVID-19-Related infodemic and its impact on public health: A global social media analysis. *Am. J. Trop. Med. Hyg.* **2020**, *103*, 1621–1629. [CrossRef]
18. World Health Organization. *An Ad Hoc WHO Technical Consultation Managing the COVID-19 Infodemic: Call for Action*; World Health Organization: Geneva, Switzerland, 2020.
19. Facebook an Update on Our Work to Keep People Informed and Limit Misinformation about COVID-19. Available online: https://about.fb.com/news/2020/04/COVID-19-misinfo-update/ (accessed on 16 September 2021).
20. Facebook Help Centre. COVID-19 Policy Updates and Protections. Available online: https://www.facebook.com/help/230764881494641/ (accessed on 16 September 2021).

21. World Health Organization. Let's Flatten the Infodemic Curve. Available online: https://www.who.int/news-room/spotlight/let-s-flatten-the-infodemic-curve (accessed on 13 September 2021).
22. Cooke, N.A. *Fake News and Alternative Facts: Information Literacy in a Post-Truth Era*; ALA Editions: Chicago, IL, USA, 2018; ISBN 978-0-8389-1751-0.
23. McIntyre, L.C. *Post-Truth*; The MIT Press Essential Knowledge; MIT Press: Cambridge, MA, USA, 2018; ISBN 978-0-262-53504-5.
24. De Paor, S.; Heravi, B. Information literacy and fake news: How the field of librarianship can help combat the epidemic of fake news. *J. Acad. Librariansh.* **2020**, *46*, 102218. [CrossRef]
25. Cosentino, G. *Social Media and the Post-Truth World Order: The Global Dynamics of Disinformation*; Palgrave Pivot: Cham, Switzerland, 2020; ISBN 978-3-030-43005-4.
26. Lazer, D.M.J.; Baum, M.A.; Benkler, Y.; Berinsky, A.J.; Greenhill, K.M.; Menczer, F.; Metzger, M.J.; Nyhan, B.; Pennycook, G.; Rothschild, D.; et al. The science of fake news. *Science* **2018**, *359*, 1094–1096. [CrossRef]
27. Zimdars, M.; McLeod, K. (Eds.) *Fake News: Understanding Media and Misinformation in the Digital Age*; Information Policy Series; MIT Press: Cambridge, MA, USA, 2020.
28. Tandoc, E.C.; Lim, Z.W.; Ling, R. Defining "Fake News". *Digit. Journal.* **2018**, *6*, 137–153. [CrossRef]
29. Case, D.O. *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2007; ISBN 978-0-12-369430-0.
30. Visser, M.; van Eck, N.J.; Waltman, L. Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quant. Sci. Stud.* **2021**, *2*, 20–41. [CrossRef]
31. Furstrand, D.; Pihl, A.; Orbe, E.B.; Kingod, N.; Søndergaard, J. "Ask a doctor about coronavirus": How physicians on social media can provide valid health information during a pandemic. *J. Med. Internet Res.* **2021**, *23*, e24586. [CrossRef]
32. Joy, I.I.; Idowu, A.-I. A Survey of COVID-19 Information Dissemination Behavior of Library and Information Professionals in Nigeria. *Libr. Philos. Pract.* **2021**, *2021*, 1–29.
33. Sharov, K.S. Adaptation to SARS-CoV-2 under stress: Role of distorted information. *Eur. J. Clin. Investig.* **2020**, *50*, e13294. [CrossRef]

34. Vraga, E.K.; Bode, L. Addressing COVID-19 misinformation on social media preemptively and responsively. *Emerg. Infect. Dis.* **2021**, *27*, 396–403. [CrossRef]
35. Hussain, A.; Tahir, A.; Hussain, Z.; Sheikh, Z.; Gogate, M.; Dashtipour, K.; Ali, A.; Sheikh, A. Artificial intelligence-enabled analysis of public attitudes on facebook and twitter toward COVID-19 vaccines in the United Kingdom and the United States: Observational study. *J. Med. Internet Res.* **2021**, *23*, e26627. [CrossRef]
36. Ramasubramanian, S.; Preetha, S.; Premavathy, D.; Prathap, L. Awareness on spread of misinformation and its effect on public with regard to COVID-19. *Int. J. Curr. Res. Rev.* **2020**, *12*, 66–73. [CrossRef]
37. Obiała, J.; Obiała, K.; Mańczak, M.; Owoc, J.; Olszewski, R. COVID-19 misinformation: Accuracy of articles about coronavirus prevention mostly shared on social media. *Health Policy Technol.* **2021**, *10*, 182–186. [CrossRef] [PubMed]
38. Mejova, Y.; Kalimeri, K. COVID-19 on Facebook Ads: Competing agendas around a public health crisis. In Proceedings of the COMPASS'20: 3rd ACM SIGCAS Conference on Computing and Sustainable Societies, Guayaquil, Ecuador, 15–17 June 2020; pp. 22–31.
39. Islam, M.S.; Kamal, A.-H.M.; Kabir, A.; Southern, D.L.; Khan, S.H.; Murshid Hasan, S.M.; Sarkar, T.; Sharmin, S.; Das, S.; Roy, T.; et al. COVID-19 vaccine rumors and conspiracy theories: The need for cognitive inoculation against misinformation to improve vaccine adherence. *PLoS ONE* **2021**, *16*, e0251605. [CrossRef] [PubMed]
40. Preston, S.; Anderson, A.; Robertson, D.J.; Shephard, M.P.; Huhe, N. Detecting fake news on Facebook: The role of emotional intelligence. *PLoS ONE* **2021**, *16*, e0246757. [CrossRef] [PubMed]
41. Sánchez, A.N. Disinformation and COVID-19: Quantitative analysis through the hoaxes debunked in Latin America and Spain. *Estudios sobre el Mensaje Periodístico* **2021**, *27*, 879–892. [CrossRef]
42. Asubiaro, T.; Badmus, O.; Ikenyei, U.; Popoola, B.; Igwe, E. Exploring Sub-Saharan Africa's Communication of COVID-19-Related Health Information on Social Media. *Libri* **2021**, *71*, 123–139. [CrossRef]
43. Galhardi, C.P.; Freire, N.P.; Minayo, M.C.S.; Fagundes, M.C.M. Fact or fake? An analysis of disinformation regarding the COVID-19 pandemic in Brazil. *Ciência & Saúde Coletiva* **2020**, *25*, 4201–4210. [CrossRef]

44. Neely, S.; Eldredge, C.; Sanders, R. Health information seeking behaviors on social media during the COVID-19 pandemic among american social networking site users: Survey study. *J. Med. Internet Res.* **2021**, *23*, e29802. [CrossRef]
45. Kothari, A.; Foisey, L.; Donelle, L.; Bauer, M. How do Canadian public health agencies respond to the COVID-19 emergency using social media: A protocol for a case study using content and sentiment analysis. *BMJ Open* **2021**, *11*, e041818. [CrossRef]
46. Keller, S.N.; Honea, J.C.; Ollivant, R. How social media comments inform the promotion of mask-wearing and other COVID-19 prevention strategies. *Int. J. Environ. Res. Public. Health* **2021**, *18*, 5624. [CrossRef] [PubMed]
47. Teixeira, M.M.; Junior, D.R.C. In the midst of the Coronavirus pandemic in Brazil, watch out for snowmen in the north and northeast regions! Post-truth under discussion. *Praksis* **2021**, *2*, 128–146. [CrossRef]
48. Pool, J.; Fatehi, F.; Akhlaghpour, S. Infodemic, Misinformation and Disinformation in Pandemics: Scientific Landscape and the Road Ahead for Public Health Informatics Research. *Stud. Health Technol. Inform.* **2021**, *281*, 764–768. [CrossRef] [PubMed]
49. Guimarães, V.H.A.; de Oliveira-Leandro, M.; Cassiano, C.; Marques, A.L.P.; Motta, C.; Freitas-Silva, A.L.; de Sousa, M.A.D.; Silveira, L.A.M.; Pardi, T.C.; Gazotto, F.C.; et al. Knowledge about COVID-19 in Brazil: Cross-sectional web-based study. *JMIR Public Health Surveill.* **2021**, *7*, e24756. [CrossRef] [PubMed]
50. Lupton, D.; Lewis, S. Learning about COVID-19: A qualitative interview study of Australians' use of information sources. *BMC Public Health* **2021**, *21*, 662. [CrossRef]
51. Biancovilli, P.; Makszin, L.; Jurberg, C. Misinformation on social networks during the novel coronavirus pandemic: A quali-quantitative case study of Brazil. *BMC Public Health* **2021**, *21*, 1200. [CrossRef] [PubMed]
52. Armitage, L.; Lawson, B.K.; Whelan, M.E.; Newhouse, N. Paying SPECIAL consideration to the digital sharing of information during the COVID-19 pandemic and beyond. *BJGP Open* **2020**, *4*. [CrossRef]
53. Ebardo, R.; de la Cuesta, J.; Catedrilla, J.; Wibowo, S. Peer influence, risk propensity and fear of missing out in sharing misinformation on social media during the COVID-19 pandemic. In Proceedings of the

28th International Conference on Computers in Education, Virtual, 23–27 November 2020; Volume 1, pp. 351–359.
54. Landi, S.; Costantini, A.; Fasan, M.; Bonazzi, M. Public engagement and dialogic accounting through social media during COVID-19 crisis: A missed opportunity? *Account. Audit. Account. J.* **2021**, *35*, 35–47. [CrossRef]
55. Wahidie, D.; Yılmaz-Aslan, Y.; Ölcer, S.; Aksakal, T.; Brzoska, P. Reasons for rejecting official recommendations and measures concerning protection against SARS-CoV-2—A qualitative study of social media posts. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz* **2021**, *64*, 616–624. [CrossRef]
56. Tripodi, F.B. ReOpen demands as public health threat: A sociotechnical framework for understanding the stickiness of misinformation. *Comput. Math. Organ. Theory* **2021**, *28*, 321–334. [CrossRef]
57. Obi-Ani, N.A.; Anikwenze, C.; Isiani, M.C. Social media and the COVID-19 pandemic: Observations from Nigeria. *Cogent Arts Humanit.* **2020**, *7*, 1799483. [CrossRef]
58. Puri, N.; Coomes, E.A.; Haghbayan, H.; Gunaratne, K. Social media and vaccine hesitancy: New updates for the era of COVID-19 and globalized infectious diseases. *Hum. Vaccines Immunother.* **2020**, *16*, 2586–2593. [CrossRef]
59. Cato, S.; Iida, T.; Ishida, K.; Ito, A.; Katsumata, H.; McElwain, K.M.; Shoji, M. The bright and dark sides of social media usage during the COVID-19 pandemic: Survey evidence from Japan. *Int. J. Disaster Risk Reduct.* **2021**, *54*, 102034. [CrossRef] [PubMed]
60. Sakya, S.M.; Scoy, L.J.V.; Garman, J.C.; Miller, E.L.; Snyder, B.; Wasserman, E.; Chinchilli, V.M.; Lennon, R.P. The impact of COVID-19-related changes in media consumption on public knowledge: Results of a cross-sectional survey of Pennsylvania adults. *Curr. Med. Res. Opin.* **2021**, *37*, 911–915. [CrossRef] [PubMed]
61. Ahmad, A.R.; Murad, H.R. The impact of social media on panic during the COVID-19 pandemic in iraqi kurdistan: Online questionnaire study. *J. Med. Internet Res.* **2020**, *22*, e19556. [CrossRef]
62. Superio, D.L.; Anderson, K.L.; Oducado, R.M.F.; Luceño, M.T.; Palcullo, V.E.V.; Bendalian, M.V.T. The information-seeking behavior and levels of knowledge, precaution, and fear of college students in Iloilo,

Philippines amidst the COVID-19 pandemic. *Int. J. Disaster Risk Reduct.* **2021**, *62*, 102414. [CrossRef]
63. Radwan, E.; Radwan, A.; Radwan, W. The role of social media in spreading panic among primary and secondary school students during the COVID-19 pandemic: An online questionnaire study from the Gaza Strip, Palestine. *Heliyon* **2020**, *6*, e05807. [CrossRef] [PubMed]
64. Apuke, O.D.; Omar, B. User motivation in fake news sharing during the COVID-19 pandemic: An application of the uses and gratification theory. *Online Inf. Rev.* **2021**, *45*, 220–239. [CrossRef]
65. Bruns, A.; Harrington, S.; Hurcombe, E. 'Corona? 5G? or both?': The dynamics of COVID-19/5G conspiracy theories on Facebook. *Media Int. Aust.* **2020**, *177*, 12–29. [CrossRef]
66. Mansour, A. Affordances supporting mothers' engagement in information-related activities through Facebook groups. *J. Librariansh. Inf. Sci.* **2021**, *53*, 211–224. [CrossRef]
67. WHO Launches a Chatbot on Facebook Messenger to Combat COVID-19 Misinformation. Available online: https://www.who.int/news-room/feature-stories/detail/who-launches-a-chatbot-powered-facebook-messenger-to-combat-COVID-19-misinformation (accessed on 15 September 2021).
68. Pennycook, G.; McPhetres, J.; Zhang, Y.; Lu, J.G.; Rand, D.G. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychol. Sci.* **2020**, *31*, 770–780. [CrossRef]
69. Revez, J.; Corujo, L. Librarians against fake news: A systematic literature review of library practices (Jan. 2018–Sept. 2020). *J. Acad. Librariansh.* **2021**, *47*, 102304. [CrossRef]
70. Igbinovia, M.O.; Okuonghae, O.; Adebayo, J.O. Information literacy competence in curtailing fake news about the COVID-19 pandemic among undergraduates in Nigeria. *Ref. Serv. Rev.* **2021**, *49*, 3–18. [CrossRef]
71. Dib, F.; Mayaud, P.; Chauvin, P.; Launay, O. Online mis/disinformation and vaccine hesitancy in the era of COVID-19: Why we need an eHealth literacy revolution. *Hum. Vaccines Immunother.* **2021**, *18*, 1–3. [CrossRef] [PubMed]
72. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. *Science* **2018**, *359*, 1146–1151. [CrossRef] [PubMed]

Disinformation Echo Chambers on Facebook

Mathias-Felipe de-Lima-Santos and Wilson Ceron

Abstract: Abstract: The landscape of information has experienced significant transformations with the rapid expansion of the internet and the emergence of online social networks. Initially, there was optimism that these platforms would encourage a culture of active participation and diverse communication. However, recent events have brought to light the negative effects of social media platforms, leading to the creation of echo chambers, where users are exposed only to content that aligns with their existing beliefs. Furthermore, malicious individuals exploit these platforms to deceive people and undermine democratic processes. To gain a deeper understanding of these phenomena, this chapter introduces a computational method designed to identify coordinated inauthentic behavior within Facebook groups. The method focuses on analyzing posts, URLs, and images, revealing that certain Facebook groups engage in orchestrated campaigns. These groups simultaneously share identical content, which may expose users to repeated encounters with false or misleading narratives, effectively forming "disinformation echo chambers." This chapter concludes by discussing the theoretical and empirical implications of these findings.

1. Introduction

The information landscape has undergone significant transformations with the widespread adoption of the internet and online social networks. This has led to both positive and negative consequences. On the positive side, information can now spread quickly and reach a vast audience. Social media platforms have played a crucial role in fostering a culture of participation by motivating people to actively create and share content. However, there were also drawbacks. Social media platforms employ algorithms that restrict the diversity of content users are exposed to, leading to the

reinforcement of pre-existing beliefs, commonly referred to as "echo chambers" [1]. These occur when individuals are exposed only to opinions that align with their own viewpoints, a phenomenon known as "confirmation bias" [2]. Furthermore, like-minded individuals often form homogeneous clusters where they reinforce and polarize their opinions through content diffusion [3]. The functionalities and features of social media platforms, including ranking algorithms, selective exposure, and confirmation bias, have played a significant role in the development of online echo chambers [4,5].

Online social media platforms have also provided a platform for individuals with malicious intent to disseminate false or misleading narratives with the intention of deceiving the public and undermining democratic processes [6]. The concern over false and misleading information is not a novel one. However, after the 2016 US presidential elections [7] and the Brexit referendum in the UK [8], the propagation of extremism, hate speech, violence, and false news on platforms has significantly accentuated their societal impact [9]. Such content often falls into the epistemological rabbit hole of "fake news" [10]. In this chapter, we use the term "disinformation" to encompass the concept of information disorder, encapsulating so-called fake news, which includes false or misleading information created and disseminated for economic gain or intentionally deceiving the public [11].

The issue of disinformation becomes even more concerning in the context of health crises. Previous public health emergencies, such as the 2014 Ebola outbreak [12–14], showcased the widespread dissemination of inaccurate information on social media. Similarly, during the H1N1 epidemic in 2009, an array of erroneous or deceptive content was propagated, ranging from conspiracy theories to unfounded rumors intended to cause harm [15,16].

Over the past decades, anti-vaccination movements have gained momentum globally [17], coinciding with a resurgence of previously controlled infectious diseases [18]. Debates on vaccines have been fueled by misleading, incorrect, and taken-out-of-context information, contributing to the perception that vaccinations are unsafe and unnecessary [18,19]. This influx of disinformation is jeopardizing

the progress made against vaccine-preventable diseases, as it fuels vaccine hesitancy [20].

The COVID-19 pandemic has vividly demonstrated the disruptive potential of information disorder, shifting the focus from the health crisis toward political disinformation, which can erode the outcomes of public health policies [21]. In this context, particularly concerning are the myriad myths surrounding the safety and efficacy of COVID-19 vaccines [22–24].

Hence, online disinformation permeates all strata of society, necessitating multidisciplinary approaches for comprehension and the implementation of countermeasures. Since there is no one-size-fits-all solution to combating disinformation, experts propose a combination of interventions, such as rectifying false information and enhancing media literacy skills, to mitigate its impact [25].

Despite concerted efforts, curbing disinformation has proven to be more challenging than initially anticipated. For instance, collusive users, as outlined in the literature, purposefully promote false narratives in others' minds, amassing high counts of retweets, followers, or likes, thereby influencing public discourse. They are often funded or composed of individuals who exchange followers among themselves to amplify their visibility [26]. Such users corrode the trustworthiness and credibility of online platforms in a manner akin to spam accounts.

In response, online social media companies have adopted diverse strategies to combat information disorders. Twitter, for example, has been actively removing accounts engaged in spam and platform manipulation.

Facebook has been actively combatting problematic content on its platform since 2018 by employing the concept of "coordinated inauthentic behavior" [27]. While Facebook's efforts have been subject to criticism for their enforcement of policies, the company has substantiated the link between coordinated behavior and the dissemination of problematic information as a measure to counter manipulation attempts on its platform. This term encompasses not only bots and trolls that propagate false content, but also unwitting

citizens and polarized groups recruited to play orchestrated roles in influencing society.

Thus, instead of establishing a distinct demarcation between problematic and non-problematic information, the company has adopted what can be described as an "ill-defined concept of coordinated inauthentic behavior" (CIB). This strategic decision is aimed at effectively tackling and curbing the spread of disorderly information throughout its platform, all the while avoiding the complexity of unequivocally labeling it as false content [9]. Academic literature suggests that these coordinated efforts have become fertile ground for the proliferation of political disinformation [28–31], a phenomenon observed across various social media platforms [32].

This chapter undertakes an examination of disinformation narratives concerning COVID-19 vaccines that have been propagated by users on Facebook. Through the lens of the echo chamber concept, this study delves into the role of user-generated content (UGC) exhibiting signs of "coordinated inauthentic behavior" within Facebook groups. To this end, the study is guided by the following research questions:

(RQ1) To what extent is problematic content shared on Facebook?

(RQ2) How are these groups interconnected?

(RQ3) How do these coordinated networks possess characteristics that contribute to the formation of echo chambers?

To answer these RQs, our approach involved sourcing fact-checked stories related to COVID-19 vaccines from two major Brazilian fact-checking initiatives, namely Agência Lupa and Aos Fatos. These stories were published during the period spanning January 2020 to June 2021, and they provided the foundation for generating keywords that were then utilized in our queries on CrowdTangle. This process was instrumental in identifying false narratives that were actively circulating on Facebook. In total, our study made use of 276 instances of debunked content to uncover and analyze disinformation narratives that were being disseminated across this online social media platform. Our analysis takes the form of a computational strategy aimed at predicting instances of coordinated behavior within Facebook groups. These groups engage in inauthentic tactics with the intent

of boosting the visibility and reach of particular content, ultimately contributing to the amplification of problematic information on the platform [9].

Our computational approach involves an analysis of content frequency and similarity, which enables the detection of potential traces of "coordinated inauthentic behavior." This can manifest through the replication of widely available narratives within specific Facebook groups or the sharing of common links in a condensed timeframe, often leading to external websites. Additionally, we extended our analysis to encompass the coordinated dissemination of visual content, commonly referred to as memes. These images are particularly susceptible to manipulation, rendering them more challenging to identify using conventional computational methods [33]. To address this, we leveraged a computer vision (CV) algorithm provided by Facebook to extract and analyze the textual content embedded within these images. This allowed our method to ascertain whether multiple images shared the same message over a brief period.

Our findings reveal a concerted endeavor to manipulate public discourse with the strategic objective of establishing "disinformation echo chambers". This is achieved by fostering a high level of engagement with false narratives across various groups, a substantial number of which are characterized by political affiliations. These fabricated information pieces possess the potential to reinforce existing biases, erode public health efforts, and trigger adverse societal consequences in relation to COVID-19 vaccines. Furthermore, the content propagated within these diverse groups can be construed as beliefs that gain potency through repeated exposure within these tightly-knit communities, effectively shielding them from counterarguments and perpetuating echo chambers [34].

In addition to these implications, the coordinated efforts to manipulate discussions within Facebook groups pose specific societal risks. This manipulation can deceive users into replicating these fabricated narratives in offline scenarios, where the tendency to resist vaccination might be exacerbated. Ultimately, our study concludes by highlighting the overarching dangers posed by these coordinated inauthentic efforts, including the propagation of confusion and mistrust

among individuals, all while hindering the effectiveness of public health responses.

This chapter seeks to expand the expanding literature on disinformation and digital platforms by illustrating how coordinated inauthentic information can potentially give rise to echo chambers by effectively amplifying specific false or misleading narratives. Moreover, the scrutiny of the structural attributes of these Facebook groups, which exhibit well-defined coordinated networks, offers insights into the potential hazards and challenges posed by disinformation narratives in influencing individuals' decision-making processes regarding vaccines. This influence can result in ignorance and misperceptions that jeopardize the formulation and execution of crucial public health policies, such as vaccination campaigns [35]. The subsequent subsections delve into an exploration of the current landscape of research concerning echo chambers and disinformation.

1.1. Transitioning from Open Channels of Communication to Echo Chambers

In their initial stages, online social networks were hailed for their potential to influence democracy and the public sphere by facilitating the exchange of information, ideas, and discussions in an unrestricted manner [36]. Online social media platforms embodied an optimistic perspective, driven by the disruption of traditional communication patterns in shaping public opinion, such as the gatekeeping role of newspapers in other forms of expert and non-expert communications [37]. These hopeful viewpoints championed the expansion of freedom, the transformation of democratic discourse, and the creation of a communal online knowledge hub [38]. However, these positive outlooks have given way to a more pessimistic stance, characterized by the recognition of homophily structures within these networks. This suggests that users tend to interact more frequently with individuals who share similar viewpoints, resulting in a limited range of perspectives that could foster social division and stimulate polarized outlooks [39].

Within this framework, the metaphor of echo chambers has gained prominence as a way to elucidate these behaviors, amplified by the

algorithms of social media platforms. It illustrates a scenario where existing beliefs are echoed and reinforced, resembling reverberations within an acoustic echo chamber [40]. Alongside the homogeneity inherent to online social networks and exacerbated by their algorithms, the concepts of selective exposure and confirmation bias have also played pivotal roles in the formation of these echo chambers within digital platforms [4,5]. Previous research has indicated that online social networks and search engines contribute to the widening ideological gap between users. Similarly, studies have identified instances of echo chambers on online social media, particularly among groups divided along ideological lines [41,42] and on controversial issues [43].

Although some studies have suggested that these effects are relatively modest [44], others argue that the term "echo chambers" might oversimplify the issue, as it is not solely a consequence of platform mechanisms but also a result of existing social and political polarizations [45]. Scholars have also put forth the argument that the extent of ideological segregation in online social media usage has been overstated, challenging the assertion that echo chambers are universally present [46].

Conversely, Facebook employs various mechanisms that could potentially exacerbate exposure to like-minded content, including the social network structure, the feed population algorithm, and users' content selection. Thus, the combination of these mechanisms might increase exposure to ideologically diverse news and opinions. However, these mechanisms still leave individuals' choices to play a "stronger role in limiting exposure to cross-cutting content" [47].

On Twitter, researchers have examined both political and nonpolitical matters to comprehend the presence of echo chambers. According to their outcomes, political topics tend to foster more interactions among individuals with similar ideological leanings compared to nonpolitical subjects [4,40,48]. In other words, their findings suggest that homophilic clusters of users dominate online interactions on Twitter, particularly concerning political subjects [49].

In the context of studying echo chambers on online social media, it is apparent that conceptual and methodological choices significantly

impact research findings [40]. For instance, studies relying on interactions or digital traces tend to indicate a higher prevalence of echo chambers and polarization compared to those focusing on content exposure or self-reported data [40]. These amplifications of pre-existing beliefs can also be shaped by the technological features of online social media platforms. In essence, the interplay between online social media interfaces and the user-technology relationship can influence the emergence of echo chambers [50].

Hence, it is crucial not only to analyze the nature of social media interactions but also to comprehend the content that users encounter in their news feeds or the groups they engage with. If the content within online groups promotes the limitation of exposure to diverse perspectives in favor of reinforcing like-minded groups that deliberately disseminate messages to larger audiences, consequently reinforcing a shared narrative, we argue that the network of groups resulting from these coordinated communication dynamics indeed resembles "echo chambers" [51]. In this chapter, we employ the term "echo chamber" to describe Facebook groups where the online media ecosystem is characterized by selective exposure, ideological segregation, and political polarization, with specific users assuming central roles in discussions.

1.2. The Never-Ending Challenge of "Fake News"

Online social networks exist in a paradoxical realm, characterized by the coexistence of homophilous behavior and the potential for information dissemination. This duality has given rise to an environment where conflicting facts and contradictory expert opinions flourish, allowing false news to proliferate and conspiracies to take root [10]. Since 2016, the term "fake news" has gained global recognition as a descriptor for this false or misleading information spread in online spaces. This content can either be fabricated or intentionally manipulated to deceive individuals [11].

However, the term "fake news" has been wielded by politicians to undermine the media [10,52], leading to the emergence of alternative synonyms such as "information disorder," "fake facts," and "disinformation" [11]. Scholars engage in debates about differentiating

between “disinformation” and “misinformation” [53]. Some argue that the distinction lies in intent, with misinformation lacking the deliberate intent to deceive. Yet, establishing intent can be challenging [54].

Despite these nuances, the term “disinformation” appears to be the most suitable to encompass this intricate landscape, as it covers both fabricated and intentionally manipulated content [11]. In the current complex information ecosystem, it is crucial to shift our focus from intention to the influence of the narratives that these posts align with. This is because people are not solely influenced by individual posts, but rather by the broader narratives they fit into [35]. The harmful consequences of information disorder arise from the human tendency to default to assuming the truth of a statement in the absence of compelling evidence to the contrary [5].

The rapid surge of disinformation from 2017 onward has fueled an extensive field of study, generating numerous publications approaching this multifaceted issue from diverse angles [35]. Some researchers aim to categorize various types of information disorders that emerge, while others scrutinize the social and individual dimensions of disinformation’s effects on the public and political spheres [11,55]. Computational methodologies have also been employed to detect so-called “fake news” [35].

Over the years, automated accounts, or bots, have attracted significant attention from researchers for their potential to influence conversations, shape content distribution, and manipulate public opinion. Although terms such as “bots”, “automated accounts”, “fake accounts”, and “spam accounts” have often been employed interchangeably, they do not always denote the same type of activity. Bots are accounts controlled by software to automate posting or interactions, while spammers generate unsolicited mass content. Fake accounts, in turn, impersonate real individuals on online platforms [54,56].

In this respect, studies demonstrate that external events and major global incidents trigger increased manipulation attempts on platforms, particularly during elections and health crises [21]. In these occurrences, traces of coordinated bot behavior could be detected [9,57–59]. For example, on Twitter, estimates vary regarding the prevalence of bots,

with some analyses suggesting 9% to 15% of profiles are automated accounts [60]. However, contrasting views also exist, asserting that bot accounts constitute more than 50% of Twitter users [61]. Interestingly, the platform itself provided an official statement in a public filing, indicating that fewer than 5% of its 229 million daily active users are categorized as "false" or "spam" accounts, as determined by an internal review of a sample [62].

To effectively address this issue, computational methods such as textual or social network analysis (SNA) play a crucial role in identifying and suspending harmful bots from platforms. These methods enable scholars to not only detect the detrimental effects of bots but also to mitigate their impact successfully. By understanding the nuanced differences between various types of automated accounts and their behaviors, researchers can develop more targeted strategies for preserving the authenticity and integrity of online conversations and content distribution [61].

Researchers have also identified the role of bots in amplifying the spread of disinformation and hoaxes by analyzing common interactions and network integrations. Hashtags used by these users have also been relevant for detecting automated accounts, as human users tend to use more generic ones and maintain a diverse range of social connections. Botometer, formerly known as BotOrNot, has been a widely used tool for bot detection on Twitter. It evaluates the extent to which a Twitter account exhibits characteristics similar to those of social bots, aiding in the study of inauthentic accounts and manipulation on online social media for over a decade [63]. However, scholars have also pointed out that bots are becoming more sophisticated around human behavior, which presents limitations for these tools [64]. Additionally, Botometer is exclusive to Twitter, making it challenging to detect malicious actors on other platforms.

Other techniques have been employed to detect manipulation attempts on online platforms, including disinformation and conspiracy narratives. These methodologies encompass statistical approaches such as linear regression [65] as well as social network analysis (SNA) that considers the diverse relationships users form within networks. Additionally, artificial intelligence (AI) methods, such as

naive Bayes models and convolutional neural networks (CNN) [66,67], have been utilized. These different techniques have been employed both individually and in combination. Despite their utility, some of these methods come with certain limitations. While AI holds potential for enhanced detection, it necessitates a wide range of input data and exhibits higher accuracy with more recent datasets. Ensuring datasets are consistently up-to-date is challenging. Additionally, the strategies employed by malicious bots have undergone substantial evolution in recent years, hampering these methods.

Other methods have been employed to detect manipulation attempts on platforms, including disinformation and conspiracy narratives. These techniques involve statistical methods, such as linear regression [65], social network analysis (SNA), which considers different types of relationships among users that form these networks, and artificial intelligence (AI) methods (e.g., naive Bayes models and convolutional neural networks—CNN) [66,67]. These methods have also been employed singularly or in combination. Despite this, some of them present caveats. While AI solutions hold promise for improved detection, they require diverse input data and are more accurate with recent datasets. However, datasets are not always up-to-date, and the strategies of malicious bots have evolved considerably in recent years.

Given these factors, there is a clear need for more sophisticated bot detection models or a greater reliance on methodologies that scrutinize the scope of activity within coordinated campaigns. When multiple entities collaborate within a network to achieve a common goal, the presence of coordination becomes evident [61. In this vein, CIB strives to monitor the manipulation of information across online social networks, leveraging content dissemination through automated means to amplify its reach. This shift in focus from content and automated accounts to information dynamics within social networks aligns with Facebook's policies, which link coordinated behavior with the sharing of problematic information [9,68].

Some scholars advocate for the advancement of techniques targeting bot coordination over mere bot detection, as orchestrated bot activities can prove significantly more detrimental [61]. This aligns with Facebook's approach to its policies, employing the term CIB

to underline the association between coordinated behavior and the propagation of problematic information [27].

Similarly, researchers have examined group-level features using graphs to identify orchestrated activities through users' shared relationships, such as friends, hashtags, URLs, or identical messages [68]. In this respect, previous studies have explored CIB through shared links on Facebook pages, groups, and verified public profiles [9].

Coordinated behaviors in online networks have been associated with the creation of echo chambers, as users intentionally orchestrate communication dynamics to disseminate messages to large audiences [69,70]. Another study has revealed a connection between the rapid dissemination of false information and the existence of echo chambers, primarily due to the existence of polarized clusters of opinions and networks that contribute to the spread of such information [71].

While researchers have recognized collective behavior among malicious actors driven by economic and ideological motives, the academic literature has not extensively explored coordinated mechanisms for spreading false or misleading content through messages and memes. Notably, the COVID-19 pandemic has highlighted the prevalence of visual content sharing for disseminating disinformation on online social networks [70]. In this context, Facebook groups could serve as pivotal conduits for the propagation of intricate contagions of viral disinformation.

This chapter seeks to address this knowledge gap by delving into this subject, specifically focusing on COVID-19 vaccine disinformation within public Facebook groups. In the subsequent section, we provide an in-depth overview of our methodology for pinpointing echo chambers of disinformation on the Facebook platform.

2. Data and Methods

2.1. Data Collection and Preparation

Recognizing that isolated bots might not represent the most critical issues on the platform, we opted to focus on coordinated activities to investigate disinformation campaigns within Facebook

groups. We contend that these communities inherently function as echo chambers, where users intentionally join these Facebook groups to be exposed selectively to information that aligns with their beliefs and values. Hence, these communities offer an ideal context to delve into information dissemination dynamics. To explore this avenue, our study follows a three-step approach.

Initially, we identified disinformation narratives circulating on Facebook by analyzing debunked content from two prominent fact-checking agencies in Brazil: Agência Lupa and Aos Fatos. Both organizations adhere to the transparency standards set by the International Fact-Checking Network (IFCN), a coalition dedicated to upholding excellence in the fact-checking industry [72]. Our data collection spanned from January 2020 to June 2021, yielding a total of 2860 items. We employed an algorithm to filter out debunks that did not include the term "vaccine" or related variations in their titles. This process yielded 250 debunks specifically addressing COVID-19 vaccines. Subsequently, we subjected these debunks to qualitative analysis, confirming that they were all false or misleading, and eliminating any that did not meet this criterion.

Moving on to our second step, we extracted relevant data to locate these debunked posts within Facebook. We utilized academic access to CrowdTangle, an insights tool owned and operated by Meta since 2016. It is important to note that prior research has highlighted certain limitations of this tool, such as incomplete metrics and restricted access to fully public spaces on the broader Facebook or Instagram platforms. CrowdTangle only encompasses public groups with a certain user threshold, as opposed to the entire spectrum of groups [9,73,74].

In our search, we aimed to pinpoint sentences that could be readily identified and would not yield unrelated results. For instance, we refrained from using phrases such as "COVID-19 vaccines" or similar constructs that could encompass both disinformation and credible information. Our search criteria aligned with the timeframe of the debunks, spanning from January 2020 to June 2021. Through this process, we retrieved a total of 21,614 posts containing disinformation across 3912 groups. Importantly, this data extraction was performed after Facebook's public announcement that it had removed false content

from its platform [75]. This announcement holds particular significance, as these posts should have been eradicated from the platform by that time, which could have hindered our study. Nevertheless, our findings reveal that this announcement was not fully realized, as many debunked posts persisted on the platform. This discrepancy suggests that the volume of such posts within Facebook public groups could be even more substantial.

In our third phase, we proceeded to download all the identified posts. Due to data extraction limitations within the tool, we segmented the process into timeframes, later amalgamating the data into a unified dataset. This database underwent a process of duplicate removal based on post IDs, resulting in the elimination of 1707 duplicated entries from our initial dataset. Consequently, our final dataset encompassed 19,457 distinct entries.

2.2. Data Analysis and Visualization

In prior investigations of coordinated inauthentic behavior, researchers utilized estimated time thresholds to identify items shared in near-simultaneity over a short period. Similarly, a statistical metric was proposed to identify concurrent link sharing by assessing the interarrival time—the interval difference in seconds between successive shares of URLs [28]. However, we chose not to adopt these thresholds in our study for several reasons.

Unlike previous studies that centered on URLs [9,28], our analysis seeks to identify CIB within textual and visual content. Moreover, our study focuses solely on coordinated activities among non-human accounts, necessitating a more stringent approach. This threshold determination was guided by similar studies that calculated this value based on a subset of the 10% of URLs with the shortest time intervals between the first and second shares [9,28,57–59]. Our empirical tests demonstrated that the timeframe calculated from the shortest intervals of 10% of URLs could range from 30 s to a minute. Consequently, depending on the dataset in use, this threshold might extend to around one minute, a timeframe that could feasibly be performed by humans.

Given these constraints, we undertook manual testing to ascertain a timeframe unlikely for consecutive human posting. Our tests

indicated that this interval should be less than 30 s. We acknowledge that factors such as internet speed and computing power might impact this performance. Nevertheless, we opted to adopt a threshold of 30s between two posts, as it represented the minimum time required for consecutive postings. Our approach also considered a recursive 30-s timeframe, accounting for the possibility of repeated new posts within short intervals—a scenario unlikely to occur frequently. This approach allowed us to identify coordinated posts that were disseminated over an extended period.

Considering these temporal criteria, our computational model assessed four elements to determine coordination between posts. First, the method analyzed the "message" field, encompassing the textual content of a Facebook post. Second, it scrutinized the "description" field, which provides textual information accompanying external URLs or images shared on Facebook thumbnails. For example, the description for the post in Figure 1 was "Uma catastrófica análise sobre as vacinas contra o vírus chinês: 'Interferem diretamente no material genético'," identical to the content in the thumbnail. Third, our methodology leveraged CrowdTangle's computer vision algorithm to detect text within images and ascertain if these visual contents were disseminated through automated means. It is worth noting that prior research has highlighted that CrowdTangle's computer vision capabilities for text recognition have been a recent development and are not without limitations [74]. Lastly, our process examined whether multiple entities rapidly and consistently shared the same URL, which serves as another indicator of coordinated activity [9,28].

Figure 1. The illustration depicts a post referring to an external URL, showcased in a thumbnail. The description field is located below the link jornaldacidadeonline.com.br." This mechanism is also observed with internal URLs, which redirect to a Facebook post. Source: Reprinted from Facebook.

To visualize the coordinated behaviors among different Facebook groups more effectively, we constructed a graph $G = (V, E)$, where each vertex $V = \{v_1, v_2, v_3, \ldots, v_n\}$ represents a Facebook group, and the edges $E = \{e_1, e_2, e_3, \ldots, e_m\}$ indicate the sharing of posts with signals of coordinated activity across these groups. This process was applied to the entire dataset, resulting in the creation of Figure 2. To implement this graph, we utilized the network analysis software Gephi [76], which allowed us to visually demonstrate the stronger connections between certain groups and the presence of structures that resemble "echo chambers." The Louvain method was employed to identify network communities within this graph [77]. This community detection algorithm relies on modularity optimization, resulting in a fast process to generate clusters [78]. Through this technique, we could pinpoint closely linked Facebook groups that formed more significant echo chambers.

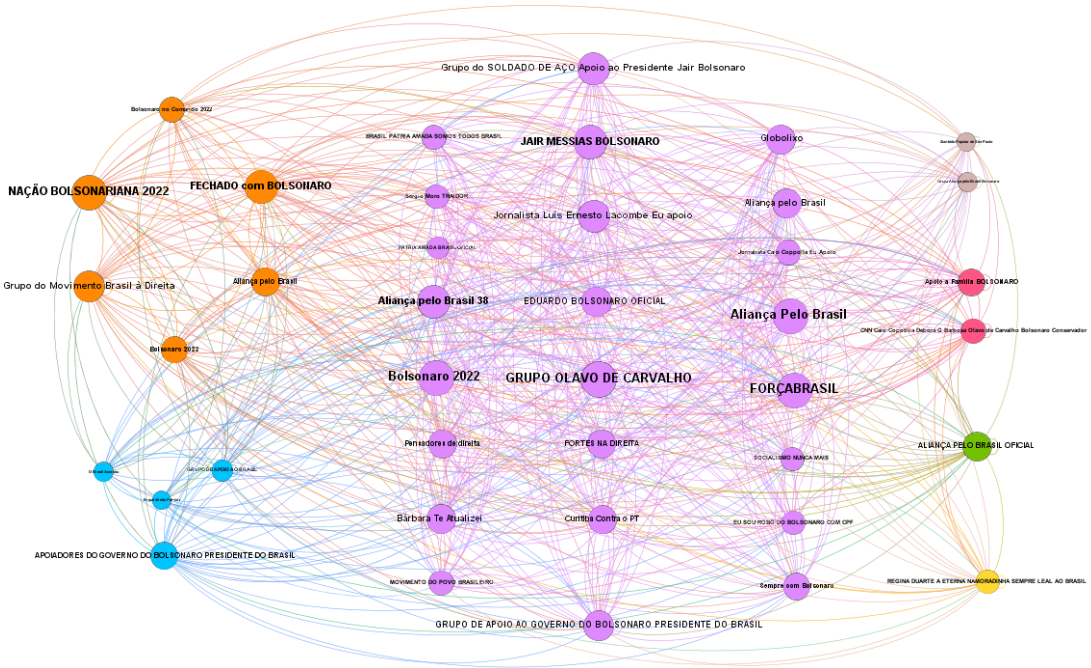

Figure 2. This graph exclusively features Facebook groups possessing degrees exceeding 100. In this context, these groups have shared a minimum of 100 coordinated posts. Remarkably, a significant portion of these groups have adopted political titles. Source: Figure by authors.

Furthermore, we generated a second graph (see Figure 3) illustrating the five most shared instances of disinformation content. This graph, denoted as $G = (D, F)$, consisted of nodes of different types, where the set of disinformation content $D = \{d_1, d_2, d_3, \ldots, d_n\}$ was connected to the set of Facebook groups $F = \{f_1, f_2, f_3, \ldots, f_n\}$ through an edge set $E = \{e_1, e_2, e_3, \ldots, e_m\}$, signifying the coordinated activity signals within the dataset. This graph vividly demonstrates the robust correlation between echo chambers and the widespread dissemination of disinformation. In the subsequent section, we delve into our findings and present these visualizations.

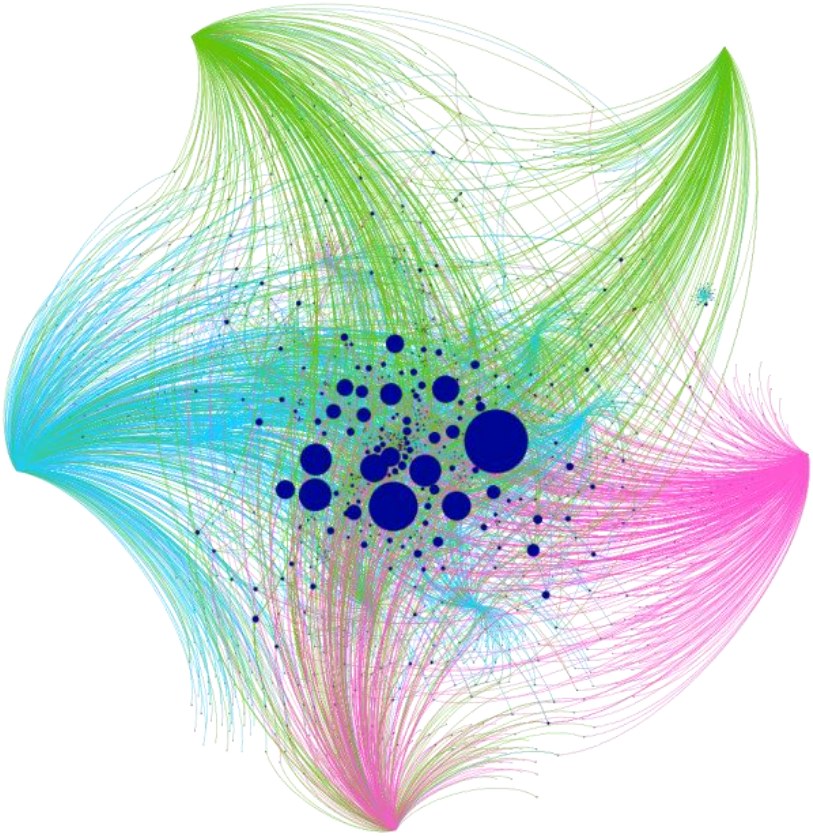

Figure 3. This graph illustrates the five most widely shared instances of disinformation content, highlighting their interconnectedness across various groups (dark blue nodes at the center). The edges portrayed in pink signify the shared videos within these groups, while the blue edges represent memes/photos, and the green edges signify URLs. Source: Figure by authors

3. Results

Within our dataset, we were able to identify that approximately 1504 out of the 3912 Facebook groups displayed indications of coordinated activity. In other words, nearly 38.5% of these groups engaged in the near-simultaneous sharing of identical content. The results also underscore that these orchestrated endeavors to manipulate public discourse span across various groups with political designations. The concern is heightened considering the nature of these posts, which contain false or misleading information.

The correlation between political Facebook groups and specific political behaviors introduces challenges to community cohesion and trust dynamics. A substantial body of literature addressing politics

and social media explores the potential impact of echo chambers on individuals' behaviors and how these might undermine efforts to uphold democratic values [79,80]. These online groups, in particular, exhibit indications of selective exposure, ideological segmentation, and political polarization. In our sample, they often adopt political labels [51]. This situation compounds existing issues by occupying a privileged position in scientific communication, thereby endangering public health and hindering efforts to manage the coronavirus pandemic. These Facebook groups serve as a tangible example of the intricate and interconnected nature of disinformation rhetoric, making empirical analysis in isolation a complex endeavor. For example, past research has highlighted the penetration of political disinformation narratives in the COVID-19 discourse during the first waves of the pandemic in Brazil [21].

Our method successfully identified certain groups that exhibited stronger associations in disseminating these disinformation campaigns compared to others. As depicted in Figure 2, the Facebook groups highlighted in pink (a total of 117 nodes) form a particularly robust "disinformation echo chamber." Within it, inauthentic actors appear to be swiftly and repeatedly amplifying inappropriate content. This occurrence transpires at a notably higher frequency than in other groups, as evidenced by a clustering coefficient of 0.85. This shows the propensity of nodes within this network to cluster together, resulting in the formation of triangles and the manifestation of robust community structures within this network [81,82].

Additionally, the magenta nodes consist of 257 Facebook groups that showcase coordinated behavior. These groups also exhibit a high clustering coefficient (0.83), indicating the presence of a strong community structure. Lastly, the blue nodes represent 150 Facebook groups wherein multiple actors appear to make concerted efforts to enhance the visibility of specific content by employing coordinated activities. This community boasts a more robust structure than the magenta one (with a clustering coefficient of 0.84), albeit with a smaller number of nodes. Our analysis further revealed the existence of smaller communities that also display traces of activities aimed at artificially boosting the popularity of certain online content. Consequently,

these Facebook groups, which likely emerge from orchestrated communication dynamics intending to disseminate messages to wide audiences, can be likened to "disinformation echo chambers."

Figure 3 underscores how the five most frequently shared narratives are extensively propagated among these Facebook groups. Housing potentially inauthentic actors, these online communities appear to amplify these problematic contents in an endeavor to elevate their visibility. This creates a causal connection that potentially links the spread of disinformation with the presence of online echo chambers [71].

In essence, when a network of groups within an online media environment engages in nearly simultaneous and recurrent sharing of disinformation narratives, the emergence of "disinformation echo chambers" becomes apparent.

4. Discussion and Conclusions

This chapter delves into the concerning prevalence of digital disinformation within online social networks, specifically highlighting how political Facebook groups have become conduits for amplifying the reach of such narratives. Our approach successfully identified instances of disinformation narratives being shared in close proximity by various entities within a short timeframe. This encompassed URLs, posts, and memes, all of which contributed to the proliferation of echo chambers on online social media platforms.

In fact, Facebook groups inherently function as echo chambers, as users deliberately join these groups to expose themselves selectively to information that aligns with their pre-existing beliefs and values [34]. However, these groups reinforce confirmation biases and contribute to the polarization of views by limiting exposure to diverse perspectives. These Facebook groups are the spaces where one gets their daily dose of confirmation bias, exacerbating their problematic behavior [4, 5]. Although some of these habits are influenced by both social and political polarization as well as platforms' algorithms, Facebook groups have emerged as fertile ground for disseminating false or misleading information [83]. This is especially evident during periods of uncertainty, such as the COVID-19 pandemic [22].

Our research outcomes highlight the significant purposeful interconnection among particular groups, driven by coordinated endeavors to propagate disinformation narratives. Specifically, these posts, especially those linking COVID-19 vaccines with inaccurate or deceptive information, have played a role in fostering the expansion of anti-vaccination sentiments. This network of interrelated groups, united by the circulation of shared content, underscores the echo chamber phenomenon, wherein they reinforce their confirmation biases. Consequently, it is plausible to view these Facebook groups as "disinformation echo chambers."

We assert that these "disinformation echo chambers" emerge from orchestrated actions aimed at intentionally spreading false or deceptive narratives to wide audiences. In our context, this poses threats to strategies aimed at curbing the impact of the COVID-19 pandemic, including vaccination efforts [34]. Furthermore, our study underscores that, despite efforts to eliminate false or misleading content related to COVID-19 vaccines, such material remained accessible to users, even when it had been debunked by fact-checking organizations collaborating with Meta/Facebook. This situation is concerning, as it indicates that the effectiveness of these measures is questionable.

It is crucial to note that our analysis primarily focused on coordinated activities driven by automated accounts. However, real users can also contribute to coordinated inauthentic behavior [84], as recently highlighted by Facebook's expanded policies against such actions. The company announced a crackdown on coordinated campaigns of actual users that cause harm on and off its platforms, expanding its measure against coordinated activities [85].

Our study's fixed threshold approach might not capture all instances of near-simultaneous sharing, considering the evolving strategies of malicious actors. Addressing such complex scenarios requires combining various methods and approaches to effectively combat information disorder in rapidly changing online environments.

Similarly, our analysis was limited to large public groups. Similar dynamics might be at play in smaller and private groups, potentially exacerbating exposure to false narratives for these individuals.

Exploring the interplay between false content dissemination in private and public groups could be a fruitful avenue for future research.

In conclusion, the ongoing pandemic has underscored the critical importance of comprehending and countering the propagation of problematic information online. This study presents an innovative computational method that uncovers the existence of "disinformation echo chambers" within public Facebook groups using different ways to manipulate the public discourse (e.g., memes, URLs, etc.). By disseminating deceptive narratives, these groups can undermine COVID-19 vaccination efforts and erode public trust in health measures. Our findings not only shed light on these inauthentic tactics but also suggest novel approaches for detection and mitigation to combat the visibility and impact of misleading content.

Author Contributions: M.-F.d.-L.-S.: Mathias-Felipe de-Lima Santos. W.C.: Wilson Ceron. Conceptualization: M.-F.d.-L.-S. and W.C.; Methodology: M.-F.d.-L.-S. and W.C.; Validation: W.C..; Formal Analysis: M.-F.d.-L.-S.. and W.C.; Data Curation: W.C..; Writing Original Draft Preparation: M.-F.d.-L.-S.; Writing Review and Editing: M.-F.d.-L.-S. and W.C.; Visualization: M.-F.d.-L.-S. and W.C.; Supervision: M.-F.d.-L.-S.; Project Administration: M.-F.d.-L.-S..; Funding Acquisition: M.-F.d.-L.-S.. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the University of Amsterdam's RPA Human(e) AI and by the European Union's Horizon 2020 research and innovation program No. 951911 (AI4Media).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sunstein, C.R. *#Republic: Divided Democracy in the Age of Social Media*; Princeton University Press: Princeton, NJ, USA, 2017.
2. Garimella, K.; De Francisci Morales, G.; Gionis, A.; Mathioudakis, M. Political Discourse on Social Media. In Proceedings of the 2018 World Wide Web Conference on World Wide Web—WWW '18; ACM Press: New York, NY, USA, 2018; Volume 2, pp. 913–922. [CrossRef]
3. Del Vicario, M.; Bessi, A.; Zollo, F.; Petroni, F.; Scala, A.; Caldarelli, G.; Stanley, H.E.; Quattrociocchi, W. The Spreading of Misinformation Online. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 554–559. [CrossRef] [PubMed]

4. Garrett, R.K. Echo Chambers Online?: Politically Motivated Selective Exposure among Internet News Users. *J. Comput.-Mediat. Commun.* **2009**, *14*, 265–285. [CrossRef]
5. Nickerson, R.S. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Rev. Gen. Psychol.* **1998**, *2*, 175–220. [CrossRef]
6. Biehl, J. A Political Economy of Pharmaceuticals. *Will Live* **2021**, *18*, 10–13. [CrossRef]
7. Allcott, H.; Gentzkow, M. Social Media and Fake News in the 2016 Election. *J. Econ. Perspect.* **2017**, *31*, 211–236. [CrossRef]
8. Bennett, W.L.; Livingston, S. The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions. *Eur. J. Commun.* **2018**, *33*, 122–139. [CrossRef]
9. Giglietto, F.; Righetti, N.; Rossi, L.; Marino, G. It Takes a Village to Manipulate the Media: Coordinated Link Sharing Behavior during 2018 and 2019 Italian Elections. *Inf. Commun. Soc.* **2020**, *23*, 867–891. [CrossRef]
10. Farkas, J.; Schou, J. Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood. *Javnost* **2018**, *25*, 298–314. [CrossRef]
11. Wardle, C. The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder. *Digital Journalism* **2018**, *6*, 951–963. [CrossRef]
12. Abramowitz, S.; McKune, S.L.; Fallah, M.; Monger, J.; Tehoungue, K.; Omidian, P.A. The Opposite of Denial: Social Learning at the Onset of the Ebola Emergency in Liberia. *J. Health Commun.* **2017**, *22* (Suppl. 1), 59–65. [CrossRef]
13. Guidry, J.P.D.; Jin, Y.; Orr, C.A.; Messner, M.; Meganck, S. Ebola on Instagram and Twitter: How Health Organizations Address the Health Crisis in Their Social Media Engagement. *Public Relat. Rev.* **2017**, *43*, 477–486. [CrossRef]
14. Guidry, J.P.D.; Meganck, S.L.; Perrin, P.B.; Messner, M.; Lovari, A.; Carlyle, K.E. #Ebola: Tweeting and Pinning an Epidemic. *Atl. J. Commun.* **2021**, *29*, 79–92. [CrossRef]
15. Smallman, S. Whom Do You Trust? Doubt and Conspiracy Theories in the 2009 Influenza Pandemic. *J. Int. Glob. Stud.* **2015**, *6*, 1–24.
16. Wagner-Egger, P.; Bangerter, A.; Gilles, I.; Green, E.; Rigaud, D.; Krings, F.; Staerklé, C.; Clémence, A. Lay Perceptions of Collectives at the Outbreak

of the H1n1 Epidemic: Heroes, Villains and Victims. *Public Underst. Sci.* **2011**, *20*, 461–476. [CrossRef] [PubMed]
17. Arif, N.; Al-Jefri, M.; Bizzi, I.H.; Perano, G.B.; Goldman, M.; Haq, I.; Chua, K.L.; Mengozzi, M.; Neunez, M.; Smith, H.; et al. Fake News or Weak Science? Visibility and Characterization of Antivaccine Webpages Returned by Google in Different Languages and Countries. *Front. Immunol.* **2018**, *9*, 1215. [CrossRef] [PubMed]
18. Dubé, E.; Vivion, M.; MacDonald, N.E. Vaccine Hesitancy, Vaccine Refusal and the Anti-Vaccine Movement: Influence, Impact and Implications. *Expert Rev. Vaccines* **2014**, *14*, 99–117. [CrossRef] [PubMed]
19. West, J.D.; Bergstrom, C.T. Misinformation in and about Science. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e1912444117. [CrossRef] [PubMed]
20. Hornsey, M.J.; Harris, E.A.; Fielding, K.S. The Psychological Roots of Anti-Vaccination Attitudes: A 24-Nation Investigation. *Health Psychol.* **2018**, *37*, 307–315. [CrossRef]
21. Ceron, W.; De-Lima-Santos, M.-F.; Quiles, M.G. Fake News Agenda in the Era of COVID-19: Identifying Trends through Fact-Checking Content. *Online Soc. Netw. Media* **2021**, *21*, 100116. [CrossRef]
22. Ceron, W.; Sanseverino, G.G.; De-Lima-Santos, M.-F.; Quiles, M.G. COVID-19 Fake News Diffusion across Latin America. *Soc. Netw. Anal. Min.* **2021**, *11*, 47. [CrossRef]
23. Chadwick, A.; Kaiser, J.; Vaccari, C.; Freeman, D.; Lambe, S.; Loe, B.S.; Vanderslott, S.; Lewandowsky, S.; Conroy, M.; Ross, A.R.N.; et al. Online Social Endorsement and Covid-19 Vaccine Hesitancy in the United Kingdom. *Soc. Media Soc.* **2021**, *7*, 205630512110088. [CrossRef]
24. Monari, A.C.P.; Sacramento, I. A "Vacina Chinesa de João Doria": A Influência Da Disputa Política-Ideológica Na Desinformação Sobre a Vacinação Contra a Covid-19. *Revista Mídia e Cotidiano* **2021**, *15*, 125–143. [CrossRef]
25. Bode, L.; Vraga, E. The Swiss Cheese Model for Mitigating Online Misinformation. *Bull. At. Sci.* **2021**, *77*, 129–133. [CrossRef]
26. Gera, S.; Sinha, A. C-ANN: A Deep Leaning Model for Detecting Black-Marketed Colluders in Twitter Social Network. *Neural Comput. Appl.* **2022**, *34*, 15113–15127. [CrossRef]
27. Gleicher, N. *Coordinated Inauthentic Behavior Explained*; Facebook Newsroom. Available online: https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/%250Ahttps://newsroom.fb.

com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/ (accessed on 11 October 2021).
28. Broniatowski, D.A. *Towards Statistical Foundations for Detecting Coordinated Inauthentic Behavior on Facebook*; The George Washington University: Washington, DC, USA, 2021.
29. Freelon, D.; Wells, C. Disinformation as Political Communication. *Political Commun.* **2020**, *37*, 145–156. [CrossRef]
30. Keller, F.B.; Schoch, D.; Stier, S.; Yang, J.H. Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign. *Political Commun.* **2020**, *37*, 256–280. [CrossRef]
31. De-Lima-Santos, M.-F.; Ceron, W. Coordinated Amplification, Coordinated Inauthentic Behaviour, Orchestrated Campaigns: A Systematic Literature Review of Coordinated Inauthentic Content on Online Social Networks. In *Mapping Lies in the Global Media Sphere*, 1st ed.; Filibeli, T.E., Özbek, M.Ö., Eds.; Routledge: London, UK, 2023; pp. 165–184. [CrossRef]
32. Yang, K.C.; Pierri, F.; Hui, P.M.; Axelrod, D.; Torres-Lugo, C.; Bryden, J.; Menczer, F. The COVID-19 Infodemic: Twitter versus Facebook. *Big Data Soc.* **2021**, *8*, 1–16. [CrossRef]
33. Kim, Y.; Song, D.; Lee, Y.J. #Antivaccination on Instagram: A Computational Analysis of Hashtag Activism through Photos and Public Responses. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7550. [CrossRef]
34. Möller, J. Filter Bubbles and Digital Echo Chambers. In *The Routledge Companion to Media Disinformation and Populism*; Tumber, H., Waisbord, S., Eds.; Routledge: London, UK, 2021; pp. 92–100. [CrossRef]
35. Righetti, N. Four Years of Fake News: A Quantitative Analysis of the Scientific Literature. *First Monday* **2021**, *26*, 1–15. [CrossRef]
36. Dahlgren, P. The Internet, Public Spheres, and Political Communication: Dispersion and Deliberation. *Political Commun.* **2005**, *22*, 147–162. [CrossRef]
37. Belair-Gagnon, V.; Nelson, J.L.; Lewis, S.C. Audience Engagement, Reciprocity, and the Pursuit of Community Connectedness in Public Media Journalism. *Journal. Pract.* **2019**, *13*, 558–575. [CrossRef]
38. Benkler, Y. *The Wealth of Networks*; Yale University Press: New York, NY, USA, 2006.

39. McPherson, M.; Smith-Lovin, L.; Cook, J.M. Birds of a Feather: Homophily in Social Networks. *Annu. Rev. Sociol.* **2001**, *27*, 415–444. [CrossRef]
40. Terren, L.; Borge, R. Echo Chambers on Social Media: A Systematic Review of the Literature. *Rev. Commun. Res.* **2021**, *9*, 1–39. [CrossRef]
41. Bright, J. Explaining the Emergence of Political Fragmentation on Social Media: The Role of Ideology and Extremism. *J. Comput. -Mediat. Commun.* **2018**, *23*, 17–33. [CrossRef]
42. Eady, G.; Nagler, J.; Guess, A.; Zilinsky, J.; Tucker, J.A. How Many People Live in Political Bubbles on Social Media? Evidence From Linked Survey and Twitter Data. *SAGE Open* **2019**, *9*, 1–21. [CrossRef]
43. Garimella, K.; Morales, G.D.F.; Gionis, A.; Mathioudakis, M. Quantifying Controversy on Social Media. *ACM Trans. Soc. Comput.* **2018**, *1*, 1–27. [CrossRef]
44. Gt Walker, P.; Whittaker, C.; Watson, O.; Baguelin, M.; Ainslie, K.E.C.; Bhatia, S.; Bhatt, S.; Boonyasiri, A.; Boyd, O.; Cattarino, L.; et al. *The Global Impact of COVID-19 and Strategies for Mitigation and Suppression*; WHO Collaborating Centre for Infectious Disease Modelling, MRC Centre for Global Infectious Disease Analysis, Abdul Latif Jameel Institute for Disease and Emergency Analytics; Imperial College London: London, UK, 2020.
45. Bruns, A. Echo Chambers? Filter Bubbles? The Misleading Metaphors That Obscure the Real Problem. In *Hate Speech and Polarization in Participatory Society*; Pérez-Escolar, M., Noguera-Vivo, J.M., Eds.; Routledge: London, UK, 2021; pp. 33–48. [CrossRef]
46. Flaxman, S.; Goel, S.; Rao, J.M. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opin. Q.* **2016**, *80*, 298–320. [CrossRef]
47. Bakshy, E.; Messing, S.; Adamic, L.A. Exposure to Ideologically Diverse News and Opinion on Facebook. *Science* **2015**, *348*, 1130–1132. [CrossRef]
48. Arguedas, A.R.; Robertson, C.T.; Fletcher, R.; Nielsen, R.K. *Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review*; Reuters Institute for the Study of Journalism: Oxford, UK, 2022.
49. Cinelli, M.; De Francisci Morales, G.; Galeazzi, A.; Quattrociocchi, W.; Starnini, M. The Echo Chamber Effect on Social Media. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2023301118. [CrossRef]

50. Bucher, T.; Helmond, A. The Affordances of Social Media Platforms. In *The SAGE Handbook of Social Media*; Burgess, J., Marwick, A., Poell, T., Eds.; SAGE Publications Ltd.: London, UK, 2018; pp. 233–253. [CrossRef]
51. Barberá, P. *How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the U.S.*; NYU Center for Data Science, New York University: New York, NY, USA, 2015.
52. Porreca, A.; Scozzari, F.; Di Nicola, M. Using Text Mining and Sentiment Analysis to Analyse YouTube Italian Videos Concerning Vaccination. *BMC Public Health* **2020**, *20*, 259. [CrossRef]
53. van der Linden, S. Misinformation: Susceptibility, Spread, and Interventions to Immunize the Public. *Nat. Med.* **2022**, *28*, 460–467. [CrossRef] [PubMed]
54. Wardle, C. Misunderstanding Misinformation. *Issues Sci. Technol.* **2023**, *29*, 38–40. [CrossRef]
55. Tandoc, E.C.; Lim, Z.W.; Ling, R. Defining "Fake News": A Typology of Scholarly Definitions. *Digit. Journal.* **2018**, *6*, 137–153. [CrossRef]
56. Chang, K.-C.; Menczer, F. *How Many Bots Are on Twitter? The Question Is Tough to Answer—And Misses the Point*; Nieman Journalism Lab. Available online: https://www.niemanlab.org/2022/05/how-many-bots-are-on-twitter-the-question-is-tough-to-answer-and-misses-the-point/ (accessed on 1 September 2022).
57. Giglietto, F.; Righetti, N.; Marino, G. *Understanding Coordinated and Inauthentic Link Sharing Behavior on Facebook in the Run-Up to 2018 General Election and 2019 European Election in Italy*; SocArXiv; LaRiCA—University of Urbino Carlo Bo: Urbino, Italy, 2019. [CrossRef]
58. Giglietto, F.; Iannelli, L.; Rossi, L.; Valeriani, A.; Righetti, N.; Carabini, F.; Marino, G.; Usai, S.; Zurovac, E. *Mapping Italian News Media Political Coverage in the Lead-Up of 2018 General Election*; SSRN Electronic Journal; Università di Urbino Carlo Bo—LaRiCA: Urbino, Italy, 2018. [CrossRef]
59. Giglietto, F.; Righetti, N.; Rossi, L.; Marino, G. Coordinated Link Sharing Behavior as a Signal to Surface Sources of Problematic Information on Facebook. *ACM Int. Conf. Proceeding Ser.* **2020**, *20*, 85–91. [CrossRef]
60. Varol, O.; Ferrara, E.; Davis, C.A.; Menczer, F.; Flammini, A. Online Human-Bot Interactions: Detection, Estimation, and Characterization. In Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017, Montreal, QC, Canada, 15–18 May 2017; pp. 280–289.

61. Khaund, T.; Kirdemir, B.; Agarwal, N.; Liu, H.; Morstatter, F. Social Bots and Their Coordination During Online Campaigns: A Survey. *IEEE Trans. Comput. Soc. Syst.* **2022**, *9*, 530–545. [CrossRef]
62. Dang, S.; Paul, K.; Chmielewski, D. *Do Spam Bots Really Comprise under 5% of Twitter Users? Elon Musk Wants to Know*; Reuters. Available online: https://www.reuters.com/technology/do-spam-bots-really-comprise-under-5-twitter-users-elon-musk-wants-know-2022-05-13/ (accessed on 1 September 2022).
63. Luceri, L.; Deb, A.; Giordano, S.; Ferrara, E. Evolution of Bot and Human Behavior during Elections. *First Monday* **2019**, *24*, 9. [CrossRef]
64. Bugra Torusdag, M.; Kutlu, M.; Selcuk, A.A. Are We Secure from Bots? Investigating Vulnerabilities of Botometer. In Proceedings of the 5th International Conference on Computer Science and Engineering, UBMK 2020, Diyarbakir, Turkey, 9–11 September 2020; Volume 2020, pp. 343–348. [CrossRef]
65. Balestrucci, A. How Many Bots Are You Following? In Proceedings of the Fourth Italian Conference on Cyber Security, Ancona, Italy, 4–7 February 2020; CEUR Workshop Proceedings. pp. 47–59.
66. Bello, B.S.; Heckel, R.; Minku, L. Reverse Engineering the Behaviour of Twitter Bots. In Proceedings of the 2018 5th International Conference on Social Networks Analysis, Management and Security, SNAMS 2018, Valencia, Spain, 15–18 October 2018; pp. 27–34. [CrossRef]
67. Akyon, F.C.; Esat Kalfaoglu, M. Instagram Fake and Automated Account Detection. In Proceedings of the—2019 Innovations in Intelligent Systems and Applications Conference, ASYU 2019, Izmir, Turkey, 31 October–2 November 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–7. [CrossRef]
68. Cresci, S. A Decade of Social Bot Detection. *Commun. ACM* **2020**, *63*, 72–83. [CrossRef]
69. Al-Khateeb, S.; Agarwal, N. Examining Botnet Behaviors for Propaganda Dissemination: A Case Study of ISIL's Beheading Videos-Based Propaganda. In Proceedings of the—15th IEEE International Conference on Data Mining Workshop, ICDMW 2015, Atlantic City, NJ, USA, 14–17 November 2015; IEEE: Piscataway, NJ, USA, 2016; pp. 51–57. [CrossRef]
70. Islam, A.K.M.N.; Laato, S.; Talukder, S.; Sutinen, E. Misinformation Sharing and Social Media Fatigue during COVID-19: An Affordance and Cognitive Load Perspective. *Technol. Forecast. Soc. Chang.* **2020**, *159*, 120201. [CrossRef]

71. Törnberg, P. Echo Chambers and Viral Misinformation: Modeling Fake News as Complex Contagion. *PLoS ONE* **2018**, *13*, 1–21. [CrossRef]
72. IFCN. *International Fact-Checking Network*; Poynter. Available online: https://www.poynter.org/ifcn/ (accessed on 27 July 2020).
73. Bruns, A.; Harrington, S.; Hurcombe, E. 'Corona? 5G? Or Both?': The Dynamics of COVID-19/5G Conspiracy Theories on Facebook. *Media Int. Aust.* **2020**, *177*, 12–29. [CrossRef]
74. de-Lima-Santos, M.-F.; Kooli, A. Instagrammable Data: Using Visuals to Showcase More Than Numbers on AJ Labs Instagram Page. *Int. J. Commun.* **2022**, *16*, 2821–2842. [CrossRef]
75. Reuters. *Facebook Removes Dozens of Vaccine Misinformation "Superspreaders"*; Reuters. Available online: https://www.reuters.com/technology/facebook-removes-dozens-vaccine-misinformation-superspreaders-2021-08-18/ (accessed on 11 October 2021).
76. Bastian, M.; Heymann, S.; Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Icwsm* **2009**, *3*, 361–362. [CrossRef]
77. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast Unfolding of Communities in Large Networks. *J. Stat. Mech.: Theory Exp.* **2008**, *2008*, 10008. [CrossRef]
78. Fortunato, S. Community Detection in Graphs. *Phys. Rep.* **2010**, *486*, 75–174. [CrossRef]
79. Howard, P.N.; Woolley, S.; Calo, R. Algorithms, Bots, and Political Communication in the US 2016 Election: The Challenge of Automated Political Communication for Election Law and Administration. *J. Inf. Technol. Politics* **2018**, *15*, 81–93. [CrossRef]
80. Theocharis, Y.; Jungherr, A. Computational Social Science and the Study of Political Communication. *Political Commun.* **2021**, *38*, 1–22. [CrossRef]
81. Girvan, M.; Newman, M.E.J. Community Structure in Social and Biological Networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [CrossRef]
82. Watts, D.J.; Strogatz, S.H. Collective Dynamics of 'small-World9 Networks. *Nature* **1998**, *393*, 440–442. [CrossRef]
83. Walter, N.; Salovich, N.A. Unchecked vs. Uncheckable: How Opinion-Based Claims Can Impede Corrections of Misinformation. *Mass Commun. Soc.* **2021**, *24*, 500–526. [CrossRef]

84. Mozilla Foundation. *Inside the Shadowy World of Disinformation for Hire in Kenya*; Mozilla Foundation: Nairobi, Kenya, 2021.
85. Culliford, E. *Facebook Cracks Down on German Anti-COVID Restrictions Group Over "Social Harm"*; Reuters. Available online: https://www.reuters.com/technology/facebook-shuts-down-network-linked-german-anti-covid-group-launches-rules-social-2021-09-16/ (accessed on 10 October 2021).

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
www.mdpi.com

MDPI Books Editorial Office
E-mail: books@mdpi.com
www.mdpi.com/books

MDPI Books

Publishing Open Access
Books & Series

mdpi.com/books

ISBN 978-3-0365-1347-8