

Review

# The Impact of Biases on Health Disinformation Research

Carmen Peñafiel-Saiz <sup>1,\*</sup> , Lázaro Echegaray-Eizaguirre <sup>2</sup> and Amaia Perez-de-Arriluzea-Madariaga <sup>3</sup> 

<sup>1</sup> Journalism Department, Campus of Leioa, University of the Basque Country, 48940 Leioa, Spain

<sup>2</sup> Sociology and Market Research Area, Camarabilbao University Business School, 48011 Bilbao, Spain; lazaro.echegaray@camarabilbaoubs.com

<sup>3</sup> LAMS Programme, Faculty of Arts of the Araba Campus, University of the Basque Country, 01006 Vitoria-Gasteiz, Spain; aperezdearri001@ikasle.ehu.es

\* Correspondence: carmen.penafiel@ehu.es

**Abstract:** This work analyses the treatment of elements such as biases and their relationship with disinformation in international academic production. The first step in this process was to carry out a search for papers published in academic journals indexed in the main indexing platforms. This was followed by a bibliometric analysis involving an analysis of the production and impact of the selected publications, using social media techniques and a semantic content analysis based on abstracts. The data obtained from Web of Science, Scopus, and Dimensions, relating to health, biases, and fake news as well as post-truth, show how these works have multiplied in the last decade. The question relating to this research is as follows: How have cognitive biases been treated in national and international academic journals? This question is answered with respect to the scientific or research method. The results, which date from 2000 to 2024, show a considerable academic dedication to exploring the relationship between biases and health disinformation. In all these communities we have observed a relationship between production with the field of medicine as a general theme and social media. Furthermore, this connection is always tied to other subjects, such as an aversion to vaccines in Community 10; disinformation about COVID-19 on social media in Community 5; COVID-19 and conspiracy theories in Community 6; and content for the dissemination of health-related subjects on YouTube and the disinformation spread about them. The community analysis carried out shows a common factor in all the analysed communities—that of cognitive bias.

**Keywords:** health communication; misinformation; disinformation; types of bias; health investigation; communication; social network analysis



**Citation:** Peñafiel-Saiz, C.; Echegaray-Eizaguirre, L.; Perez-de-Arriluzea-Madariaga, A. The Impact of Biases on Health Disinformation Research. *Societies* 2024, 14, 64. <https://doi.org/10.3390/soc14050064>

Academic Editor: Eugène Loos

Received: 31 March 2024

Revised: 21 April 2024

Accepted: 30 April 2024

Published: 8 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction and Theoretical Framework: A Conceptual Vision of Disinformation/Misinformation and the Importance of Biases in Decision Making

The term ‘disinformation’ is understood as the action of inducing confusion in public opinion through the use of false information. The Royal Spanish Academy defines it as the action of giving intentionally manipulated information at the service of certain ends. To give insufficient information or omit it. However, these concepts in English contain important differences: misinformation refers to false or incorrect information, while the term disinformation refers to information deliberately given with the aim of creating confusion or fear among the population. In Spanish, we refer to both concepts as ‘disinformation’ To quote Southwell et al. [1], misinformation is false or inaccurate information regardless of intentional authorship, much discussion about misinformation has focused on malicious acts with the intention of infecting social media platforms with false information. Regarding misinformation and its effects, they pose a risk to international peace; interfere with democratic decision making; endanger the well-being of the planet; and threaten public health [2]. Thus, the concept of misinformation as a problem appears prominently in recent academic literature and public discourse [1].

Thus, from semiotic and psychological perspectives, disinformation occurs when what is communicated does not fit with the current reality of an object. Political science

and public relations approaches see disinformation as an application of manipulation techniques to the masses and public opinion, while a communication and information science perspective considers it a natural characteristic of news media and of a clearly oversaturated communicative ecosystem [3].

The digital ecosystem of social communication, saturated by the rise of social media and the crisis within news media, is a fertile breeding ground for an increase in disinformation in our daily lives. The study of it must begin with a behavioural examination of the symbolic elites that may be able to manipulate collective thinking in favour of their own interests [4]. This structure of restricted information, in part, moves away from being a clear expression of reality and becomes a prudent selection of a public agenda under the triangulation of three fundamental elements of communicative research: discourse, cognition, and society [4], in which a discursive and semiotic analytical focus is required because the majority of disinformative content is spread via text, the spoken word, or image. This takes place in a society that is increasingly incapable of accessing privileged data for itself, which generates an avid demand for information via intermediaries as society strives to understand its own shared realities [3].

Disinformation has developed exponentially with the COVID-19 pandemic, but its effects started to increase exponentially since 2018 [5]. Three studies carried out in Spain during the pandemic and post-pandemic time periods produced important conclusions relating to disinformation: Salaverría et al. [6] developed a study about the fake news generated in Spain during the first month of the State of Alarm, identified by Spanish verification platforms. The research group discovered that these fake news items were generally started on social media, mainly on WhatsApp. In addition according to fake news about science and technology, a great degree of false content there were found relating to political and governmental subjects [7]. Another study, by Noain-Sánchez, centred on the disinformation generated in Latin America and Spain between 1st January and 1st June 2020 [8]. The study was carried out by observing accredited fact-checking platforms. The conclusions indicate that most fake news was spread in text format, and that the most common medium of dissemination was social media platforms: Facebook in Latin America and WhatsApp in Spain. Disinformation in this case not only affected health but also politics. The study by [9] was carried out a year after the start of the pandemic, when the news was making reference to vaccines and the vaccination process. Our team found that in this research too the diffusion of fake news focused on the political debate. In this case, the dissemination was carried out via Twitter and WhatsApp [9].

Disinformation can lead to so-called systematic error, which occurs when citizens select or favour certain responses over others. This is when cognitive biases appear: shortcuts taken by the brain when it processes information. These shortcuts can hinder decision making and generate irrational and incorrect behaviour (Kahneman, 2011) [10]. In his work about the cognitive biases generated in the COVID-19 pandemic, researcher Castro Prieto posed an interesting question: Why do people sometimes show a tendency to take steps that are not beneficial to them or to society? Faced with uncertainty, our brain resorts more to these biases, generating quick and impulsive responses to decisions that require evaluation. Bias analysis is connected to guidelines of social behaviour that are directly related to decision making. Decisions are not always taken in a rational way [11], and, in this sense, decision making can lead to erroneous situations. The minute that a relationship is produced between the news, bias, and behaviour, this relationship should be studied and analysed from scientific and academic perspectives.

The research draws on international literature sources to establish a list of biases to be taken into account in the process. Focussing on the work of Kahneman and Tverski, various research teams have centred their investigations on a series of biases that are worth considering. It comes as no surprise that these authors conclude that human rationality is affected by behavioural or cognitive heuristic biases. Similarly, the author Cerezo, aiming to detect biases produced during the pandemic, gives a brief summary of these types of cognitive bias: one of them is loss aversion, which forms part of prospective theory [12] and

explains how humans give more importance to a loss than a gain; in the case of a pandemic, a population will think more about what it is losing with the crisis than what it stands to gain by following proposed health measures. According to the author, when this happens it is those responsible for public health policy who should mitigate this overvaluation of losses. Another suggested bias is the so-called carry-over effect, also known as the bandwagon effect: this is doing what the majority are doing with no concern for the correction or suitability of the action with regard to the circumstances. One example is how, in the first stages of the pandemic, people bought more food than they needed. The way to deal with this bias is by appealing to rationality. Another bias may come from the little value given to long-term consequences. This bias violates the motivation to carry out certain actions when the threat they pose is not observed in the short term. This is also known as present bias, and shows that we prefer to enjoy the present than think about future consequences. The authorities should promote hopeful messages to counteract this bias, showing the short-term results of proposed measures. Optimism bias is based on our belief that it is unlikely that we will experience a negative event. In terms of illness, this bias makes us see ourselves as less vulnerable than other collectives. The solution to this is thought to be through clear messages from health authorities and news media—emotional messages based on sentiments (i.e., empathy, fear, and individual responsibility). Availability bias, which caught Freudenburg’s attention (1993) following the research of [12], overestimates the most accessible or closest information and can be mitigated by reinforcing the idea that citizens are responsible and obey rules. Good image bias translates to the continuity of habitual dynamics, paying no attention to advice involving a change of attitude or behaviour. In some way, those who develop this bias think that not doing what is recommended shows them in a more favourable light since it makes them more attractive than the rest. It can be combatted by appealing to and involving leaders in the promotion of recommended behaviour. Confirmation bias evaluates information from a position of existing ideas and beliefs, which means we always find reassurance for our own arguments. Thus, the search for information centres on looking for news that supports our own points of view; only supporting information that is of interest. This bias is related to post-truth theory and fake news or conspiracy theories. In this case, it is necessary that both health authorities and news media, along with opinion leaders, neutralise them (Cerezo, 2020) [11]. Authors such as Helena Matute (2019) [13] suggest other types of bias, such as that of familiarity, which leads the brain to trust what it knows and mistrust what is new, or that of the illusion of causality. Table 1 summarises these biases, quoting those authors who have had the strongest relationships with them, highlighting Kahneman and Tverski [12], among others like Leibenstein [14], Freudenburg [15], Lewicki [16] and Nickerson [17], the pioneer being Leibenstein in 1950 with the carry over bias.

**Table 1.** The most Notable Cognitive Biases and its authors. Source: own production from authors’ literature.

Most Notable Cognitive Biases	Definition
Loss aversion	More importance is given to a gain than to a loss [12].
Carry-over effect (bandwagon)	Actions are taken in line with the majority, with no reflection about the action carried out [14].
Present bias	The preference for enjoying the present without worrying about future consequences [11].
Optimism bias	The tendency for a person to think that it is unlikely that negative things will happen to them [12].
Availability bias	Underestimates the nearest and most available information [12,15].
Good image bias	Habitual behaviour and conduct are reinforced without considering the need for change [16].
Confirmation bias	Information is accepted or rejected based on the existence of preconceived ideas [17].
Familiarity	Makes people trust what they know and mistrust what they donot [11].

Disinformation campaigns do not usually limit themselves to the dissemination of fake news but aim to create a malicious narrative, largely based on the idea of post-truth. It has become the scourge of our times, infringing deontological codes and undermining democracies. Therefore, it presents a challenge for the information professionals fighting against it, but also an opportunity since it is necessary to confront it with the veracity of facts, rigour, and ethics.

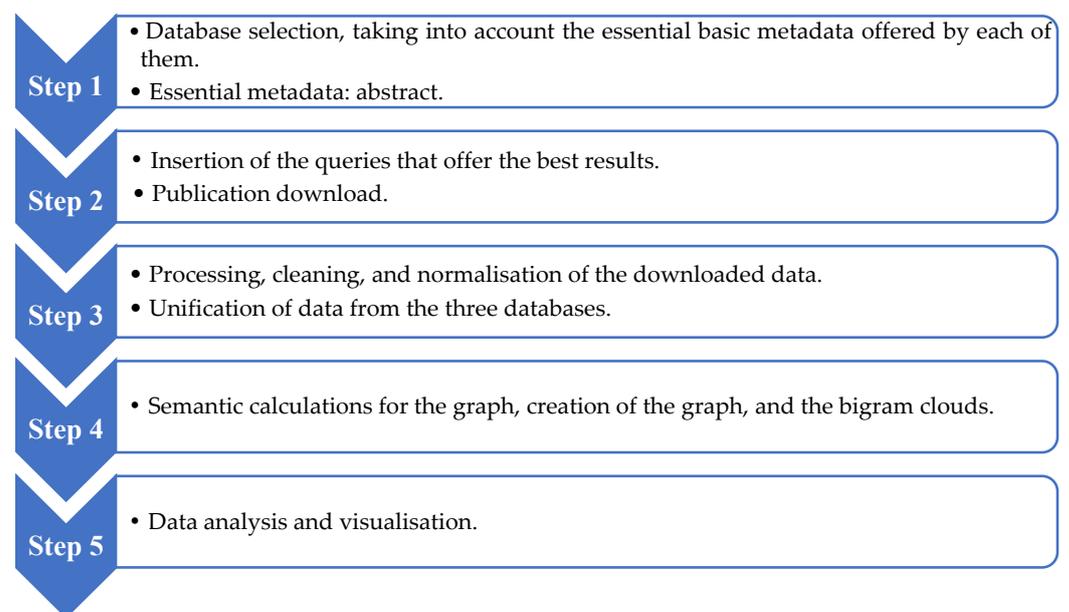
The Oxford Dictionary defines post-truth as the consequence of a political climate in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief [18]. Thus, post-truth arises when what is important is not the truth but perception, and so the conflict between your facts and mine is resolved without worrying about which of these is based on truth [19,20]. The problem is that inaccurate information, rumours, and conspiracy theories can seriously impair people's ability to make health, environmental, political, social, and economic decisions that are crucial to their lives [21].

News media, along with politics, are also responsible to some extent for the construction of post-truth via the propagation of fake news.

The endorsement of disinformation has negative consequences for public health since there is a tendency to believe that the official information exaggerates the risks of COVID-19 [22]. The conclusion of this work highlights the effect that the disinformation from conservative news media had on public health; an example is the tendency of consumers of these media to think that politicians were exaggerating the importance of COVID-19 [22], a misconception that had serious consequences. At the present time, it seems that post-modernity and post-truth proclaim an end to the desire for the search for truth; what is sought nowadays is the legitimisation of ideology. The truth or lie behind what is alleged is no longer important; what matters is the willingness to give it the category of truth [23]. The concept of fake news is prior to that of post-truth, but the two go hand-in-hand.

## 2. Methodology

The methodology for the bibliometric analysis based on the semantic content of abstracts is articulated around several fundamental steps (i.e., Figure 1), aimed at exploring and visualising research trends as well as thematic patterns in the academic literature:



**Figure 1.** Workflow of the production and impact bibliometric analysis model.

For this purpose, data collection was carried out using three sources: Scopus, Web of Science (WoS), and Dimensions, to ensure a broad and diversified dataset, in addition to creating quantitative and qualitative variables that would then be used in the subsequent analyses. The queries performed in the search were 'health AND disinformation AND bias' and 'health AND misinformation AND bias'. The first query resulted in a total of 102 publications, while the second resulted in 680 publications. No filters of any kind were used in the queries. The research areas of the publications were acquired. Therefore, the initial corpus consists of 782 publications from the three data sources.

The next step was the standardisation of categories across databases to unify the analysis. Standards in bibliometric research understand that the ideal practice is to use a single type of database where indicators are already standardised [24]. In our case, this was not possible due to the high atomisation and specialisation of the databases; however, no relevant metadata gaps were found, and differences in metadata were mitigated by selecting higher-quality metadata, such as the number of citations of a publication (cited or cited count).

Data cleaning was carried out in different phases: Firstly, we eliminated those publications that did not contain an abstract, those publications whose typology corresponded to preprint, letter, note, or survey, and duplicate publications, considering the title, abstract, and DOI, after standardising spelling and punctuation marks. In the event that the same publication was found in two or more data sources, priority was given to maintaining the highest number of citations. On the other hand, the research areas were standardised in their nomenclature, assuming the Scopus classification to be valid (i.e., the Dimensions and WoS areas were translated into the Scopus ones), as it is the most generic and inclusive. The final dataset was reduced to 374 publications.

Next, the semantic analysis of the abstracts was carried out using the LaBse model (i.e., language-agnostic BERT sentence embedder [25]) for sentence embedding, a deep learning-based natural language processing (NLP) technique that converts sentences into vectors based on prior training. These vectors represent the semantic features of sentences in multidimensional space (i.e., matrices). After applying this process of vectorising the sentences in the abstracts, those abstracts with similar meanings are placed close to each other in this multidimensional vector space, while abstracts with different meanings were placed further away.

In order to minimise the loss of semantic information in the vectors, the distances between them were calculated from the cosine similarity, a calculation commonly used in the context of sentence embeddings because of its ability to deal with multidimensional data [26]. Then, a graph of semantically similar abstracts was constructed based on the highest values obtained in the previous calculation: those with a cosine similarity equal to or greater than 0.9. The threshold was calculated in conjunction with modularity (see below) to obtain a graph partition that would allow for the observing of relevant differences between communities without decomposing the graph into isolated subgraphs. With this process, a mathematical object was created to analyse the different narratives in the corpus of abstracts. Additionally, cosine similarities greater than 0.99 were used to identify and eliminate previously undetected duplicate publications from the network. Finally, the abstract network consisted of 366 nodes and 4394 edges.

The graph was transformed into an undirected network with the Python library NetworkX [27]. The modular structure of the network was then identified with the Louvain algorithm [28]. This is a procedure that identifies communities in networks based on modularity optimisation, which is a measure that quantifies the structure of modules, clusters, or communities within a graph [29]. In this case, the modularity value achieved was 0.203: a low value that places us before a set of poorly differentiated narratives (i.e., which is congruent with a methodology such as the one used to study a very specific and delimited scientific field), but which derives from the methodological decision of the 0.9 threshold. At all times we will bear in mind that the communities we are analysing have many more elements in common than distinctive ones, which also accounts for the

particularities of the empirical object we are dealing with. Next, the weighted degree of the nodes of the graph was calculated (i.e., the number of neighbours of each node multiplied by the power of the links). All of these data were exported and added to the initial dataset.

With the aim of deepening the semantic analysis of the communities created with the previously calculated modularity, a bigram calculation was carried out with the abstracts of the publications to capture contextual relationships between words. These are pairs of consecutive elements taken from a sequence of elements, such as the co-occurrence of two words in a text, which are calculated after having removed the stop words (i.e., function words) from the texts. Two different bigram calculations were carried out: one on the abstracts—on the one hand, a bigram for each community, and, on the other hand, two bigrams for all communities based on pre-and post-2020 time spans—and the other focusing on specific terms of interest, specifically on the word bias. Thus, as we will see below, each community has a bigram cloud related to the contexts of the abstracts of the publications of that community, and then two general bigrams specific to the periods before and after 2020 are analysed. Finally, two separate bigram analyses were carried out to examine the bigrams created based on the word bias.

To visualise the data, Gephi was used to synthesise the graph using Force Atlas 2, a brute force algorithm that moves linked nodes closer together and away from unlinked ones [30]. The rest of the visualisations were made with Python in the case of the bigram clouds for each cluster—both those describing the co-occurrences in the abstracts of the publications of each cluster and the bigram clouds containing the word bias in one of the two n-grams—and PowerBI for the rest of the visualisations.

The exploration of results was carried out in two stages: Firstly, bibliographic production and impact were analysed on the basis of general indicators. Secondly, a detailed analysis of each community in the graph was carried out. The methodology includes a detailed examination of topics within specific research communities based on the identified communities, as well as a collective analysis of topics outside these communities (i.e., communities isolated from the graph; not exceeding the threshold of 1% of nodes).

### 3. Results

#### 3.1. Production and Impact per Community

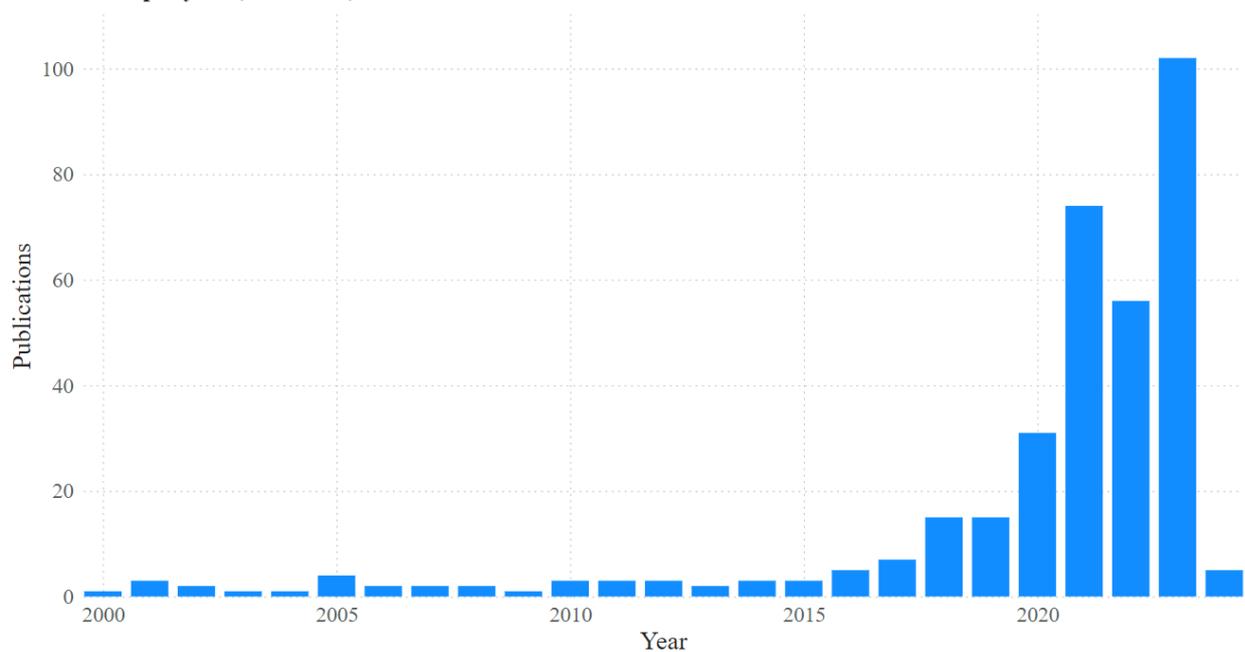
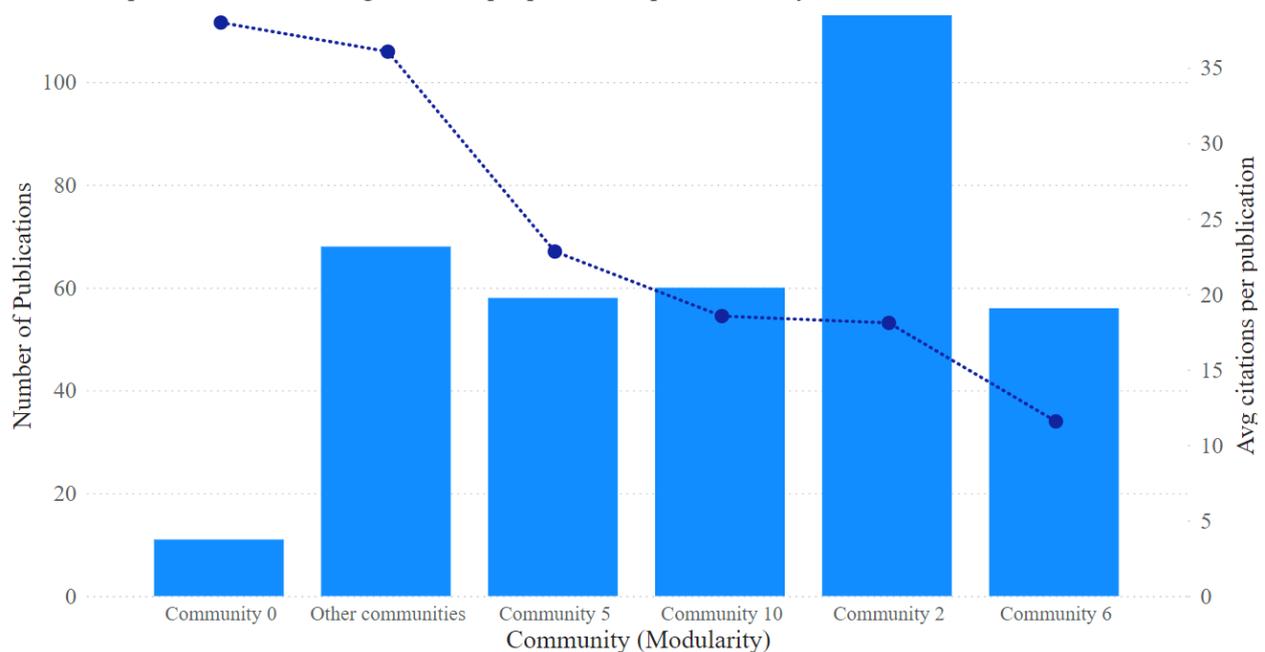
##### 3.1.1. Number of Publications per Year

The production of the topics analysed (i.e., the queries ‘health AND disinformation AND bias’ and ‘health AND misinformation AND bias’) starts to grow from 2016, before the COVID-19 pandemic, and an exponential trend can already be observed between 2018 and 2020 (i.e., Figure 2); despite the drop in production in 2022, in 2023 it grows again, surpassing that of 2021. This suggests a great scientific interest in the topic, which also concerns several research areas, as we will see in the cluster analysis.

Although there are publications before 2000, the following graph has been filtered to show publications from 2000 onwards, which is when most of the activity starts to accumulate.

##### 3.1.2. Number of Publications and Average Number of Citations per Publication and Community

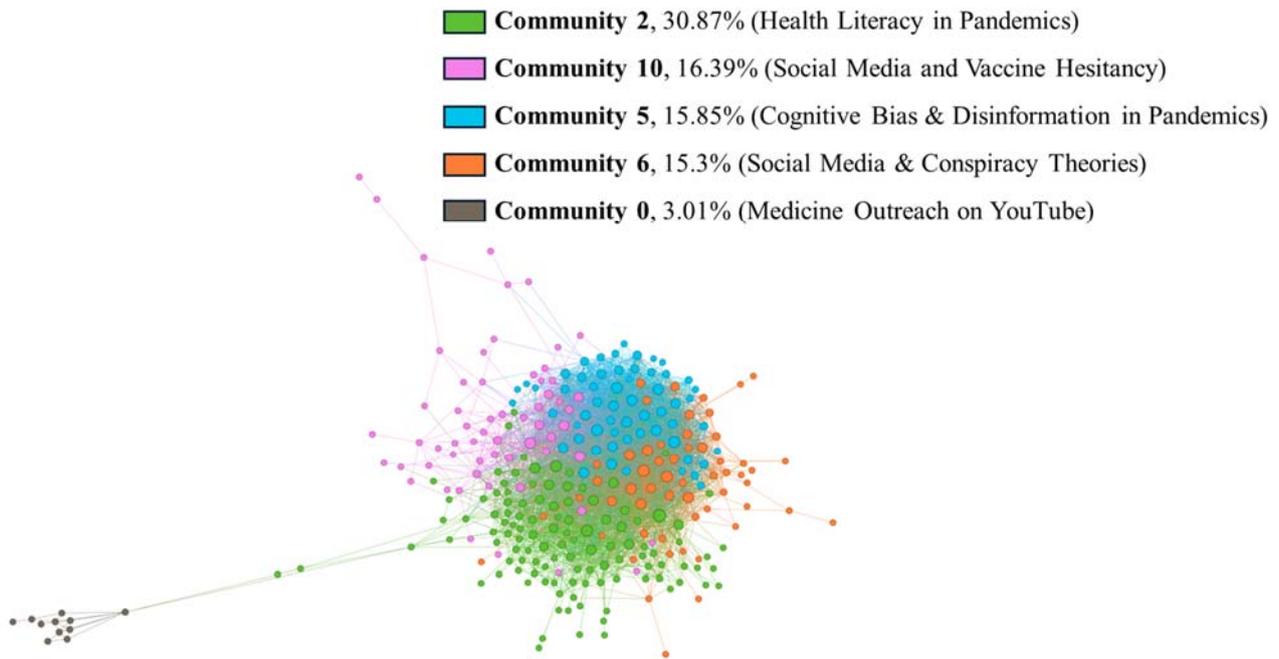
The following bar chart (i.e., Figure 3) shows, on the left vertical axis, the number of publications per community identified in the graph (i.e., 0, 10, 2, 5, 6, and the rest of the communities, as we will see in the next section), and, on the right axis, the average number of citations per publication in each community. To analyse the success of each community based on the number of publications and the average number of citations per publication, it is understood that the lower the number of publications and the higher the average number of citations per publication, the more successful the community will be.

**Publications per year (since 2000)****Figure 2.** Publications per year since 2000. Source: own elaboration with PowerBi.**Number of publications and average citations per publication per community****Figure 3.** Number of publications and average number of citations per publication and per community. Source: own elaboration with PowerBi.

Therefore, the community that stands out the most is Community 0 (i.e., the fifth according to the number of nodes), as it has the lowest number of publications (11) but the highest average number of citations per publication (38), as well as the highest impact. On the other hand, Community 2 (i.e., the first according to the number of nodes in the graph) has the highest number of publications (113), but its average number of citations per publication (18.10) is not as high as that in the previous community. This indicates that it is a community with a lot of production, but with a relatively low impact in comparison to this production.

### 3.2. Community Analysis

The following undirected graph (i.e., Figure 4) is made up of 366 nodes and 4394 weighted edges (i.e., taking weights between 0.9 and 0.99 according to the cosine similarity value between abstracts). After the application of the Louvain algorithm for the identification of communities, the result was 71 communities and a modularity of 0.203. Of the 71 communities, the most relevant ones were selected for the analysis, which are those that exceed 1% of the nodes: 5 communities in total.



**Figure 4.** Graph of abstracts with featured communities. Source: own elaboration with Gephi.

The above graph has a relatively low modularity value, suggesting little differentiation between communities. Its relatively high density of 0.066 reinforces this notion, as the volume of links between nodes—denoting similarity relationships—is relatively abundant. On a visual level, we can highlight that Communities 2, 5, and 6 are part of the same node set, while Communities 10 and, above all, 0 remain structurally separate. This, in the interpretative context of the Force Atlas 2 algorithm, means that a greater differentiation from the rest of the communities can be expected from these two.

To analyse each community in the graph, a dashboard was developed with PowerBi and a series of visualisations with the Wordcloud and Matplotlib libraries for Python [31,32] (i.e., Figure 5). The indicators shown for each community are as follows: (1) the total sum of publications and citations for each community, (2) the top five publications with five citations or more ordered by prototypicality (i.e., the sum of the edge weights with regard to the other nodes of the graph, i.e., the weighted degree), (3) a timeline of the number of publications per year, (4) a word cloud with the most important bigrams of the abstracts (i.e., the higher the co-occurrence of the words that make up the bigram, the larger they appear in the visualization), and (5) the top five research areas by number of publications, considering that there are publications with more than one research area.

#### 3.2.1. Community Analysis—Com. 2—30.87% of Nodes (Health Literacy in Pandemics)

As we saw earlier in the bar chart analysis (i.e., Figure 3), Community 2 (i.e., Figure 5) is the least successful community in terms of the number of publications and the average number of citations per publication. Most of the publications are in the fields of medicine and computer science. The most frequent bigrams are social media, public health, health information, and health misinformation, which indicates that the central theme of this



with the role of cognitive biases in the search for and acceptance of (dis)information in social media when (dis)informing oneself about public health issues and COVID-19 in the pandemic.

This community is specifically concerned with the flow of public-health-related information on social media, as well as how disinformation can influence public perception and social behaviour with respect to public health recommendations in pandemics. The volume and speed of the information shared on these channels can lead to both the widespread dissemination of truthful data and the circulation of fake news as well as conspiracy theories that undermine the preventive efforts of institutions regarding public health issues.

With respect to the role of biases, these may influence people to favour information that confirms their prior beliefs (i.e., confirmation bias) or to have the perception that nothing negative is likely to happen to them (i.e., optimism bias), in addition to affecting the adoption of health practices.

#### 3.2.4. Community Analysis—Com. 6—15.3% of Nodes (Social Media and Conspiracy Theories)

Most of the publications are in the field of medicine with an interdisciplinary approach, including the social sciences and communication. The topic focus of this community is on the interaction between social media and conspiracy theories in the context of the COVID-19 pandemic. This cluster is mainly concerned with the role of cognitive biases and social media (dis)information seeking in (dis)informing about public health issues and COVID-19 in the pandemic. Furthermore, research in this cluster is oriented towards understanding the psychological biases and social mechanisms that underlie the acceptance of false narratives and how to address these in order to counteract disinformation and thus also strengthen societal resistance to disinformation.

#### 3.2.5. Community Analysis—Com. 0—3.01% of Nodes (Medicine Outreach on YouTube)

As we saw earlier in the bar chart analysis, this community is the most successful in terms of the number of publications and the average number of citations per publication. Most of the publications are in the field of medicine. The central topics of this community are health content disseminated via YouTube and the issue of the misinformation as well as disinformation associated with it. Studies in this community critically evaluate the quality, bias, and accuracy of videos available on YouTube, covering topics as diverse as treatments for common diseases and management options for specific health conditions such as prostate cancer.

This community is also dedicated to understanding the impact that health content on YouTube has on patients' knowledge and decisions. Recognising that many users turn to this social media platform as a primary source of information, the researchers seek to identify how the variable and often biased quality of videos can lead to misunderstandings and the adoption of potentially harmful health practices.

#### 3.2.6. Community Analysis—Other Communities—18.58% of Nodes

Most of the publications belong to the field of medicine. The central topic of this community is the relationship between social media and the proliferation of disinformation in the COVID-19 pandemic. This cluster is mainly about the role of news media and social media in creating and disseminating misinformation on health issues of different types and in different contexts, not only in the COVID-19 pandemic, and how disinformation can affect people's willingness to get vaccinated, as well as other critical health-related decisions.

The most impactful research from this community—the publications in this cluster could not be ordered by prototypicality, so they have been ordered from highest to lowest impact, i.e., number of citations—also focuses on strategies with which to combat misinformation in social media, exploring how to develop methods to identify and rectify disinformation.



- Community 2 focuses on confirmation bias and other cognitive biases that affect judgements and decisions; publication bias, which influences scientific literature; and commercial bias, which can distort content for economic interests. These factors, especially the cognitive ones, are key to understanding and dealing with disinformation.
- In Community 10, risk bias, linked to how subjective perceptions can alter the assessment of real risks, and publication bias, which is based on the non-publication of research results depending on the direction of the research findings, stand out.
- Community 5 focuses on optimism bias, the tendency of people to believe that they are less likely to experience negative events. Media bias (i.e., social media bias, for the contexts of the communities analysed so far) also appears prominently, reflecting concerns about how the media can influence public perception.
- Community 6 focuses on cognitive bias and unconscious bias, highlighting how unconscious prejudices and reasoning errors affect the interpretation of information.
- Community 0 focuses on commercial bias; commercial interests can bias content and information, affecting the objectivity and usefulness of the data presented, especially regarding the dissemination of biased information and its impact on public perception.
- Finally, the other communities present a richer collection of types of biases, although they focus on cognitive biases, those related to commercial bias, confirmation bias, and disinformation bias.



Figure 8. Bigram clouds containing the biases for each community. Source: own elaboration with Python.

Not all biases found in the analysis coincide with those that appear in the literature. The analysis shows new biases that have not been previously mentioned and that should be taken into account: publication, commercial, risk, media, unawareness, and disinformation.

### 3.5. Bigrams Containing Bias: Before and after 2020

In addition to the previous analysis of bigrams for all communities before and after 2020 period, another one was carried out focusing on a pair of bigram clouds containing bias before and after 2020. The first cloud (i.e., Figure 9) shows bigrams containing bias for all communities before 2020, the most prominent bigrams include confirmation bias, misinformation bias, weight bias, risk bias, and cognitive biases. This suggests a focus on discussions about certain types of cognitive biases and the quality of information in this period. After 2020 (i.e., Figure 10), on the other hand, the terms that stand out are cognitive biases, commercial bias, confirmation bias, media bias, and risk bias. In both bigram clouds the prominence of cognitive and confirmation biases is prevalent, although in the case of the post-2020 period media bias and risk bias stand out.



Figure 9. Bigram cloud containing bias before 2020. Source: own elaboration with Python.



Figure 10. Bigram cloud containing bias after 2020. Source: own elaboration with Python.

In both figures, the recurrent presence of terms related to bias and misinformation may indicate an ongoing concern with how cognitive biases and biased information can affect understanding and decision making in various contexts, especially those with important implications for society, such as public health and media consumption. These core terms reflect relevant topics of discussion in academia, the media—probably social media—and public conversations before and after 2020.

## 4. Discussion and Conclusions

Having observed the results obtained in the search and analysis process, this research shows the existence of academic articles or papers that have dealt with biases and their

relationship with how the world of academia has treated disinformation. Thus, in terms of community analysis, the results show a considerable academic dedication towards the existence of biases and disinformation. In all these communities we have observed a relationship between production, with the field of medicine as a general theme, and social media, with this connection always being tied to other subject matter, such as an aversion to vaccines in Community 10; disinformation about COVID-19 on social media in Community 5; COVID-19 and conspiracy theories in Community 6; and material for the dissemination of health-related subjects on YouTube, as well as the disinformation that exists about them.

This community analysis reveals a common factor in all of the communities analysed—that of cognitive bias; however, it must be taken into account that, according to various authors, there are significant relationships between different types of bias, and this helps people to protect the image they have of themselves, leading them to analyse information in a particular and partial way [38].

Precision about the type of cognitive biases treated in academic publications can be found in the second synopsis of the Results section—the bigrams containing bias per each community. Thus, Community 2 deals with confirmation bias, along with other cognitive biases, such as publication and commercial. The first analyses and interprets information from existing ideas and beliefs, giving preference to one's own arguments. In this same analysis, Community 10 refers to risk bias and also publication bias, focussing in this case on those papers referring to the COVID-19 pandemic and which were accepted or rejected depending on the results produced.

Other biases were taken into account in the preparation of this work, such as the bias of optimistic self-perception, which appears in Community 5 and refers to the tendency of one to think that bad things will not happen to them. This bias appears in the lifting of information as optimism bias. Additionally, prejudices and errors of reasoning appear in the analysis—they are cognitive and unconscious biases, which appear in Community 6. Similarly, commercial bias appears in Community 0, found in a relationship between commercial interests and information bias.

Therefore, it can be seen how academic literature has treated the analysis of cognitive biases present in the relevant published information. It is the analysis of information generated at one particular time, in one particular context, and in one particular circumstance that finds common elements in the search for biases. A large number of those that had been taken into account in this work appeared in the analysis. Others that had not been taken into account refer to two elements that should generate special interest: commercial bias and publication bias. The first favours an informative focus not based on objectivity but on certain economic market interests. Publication bias is also particularly important since it focuses on information—in this case academic—depending on the results of research. This is highly damaging for the practice of academic research, since it causes it to lose objectivity the minute publishing companies have a preference for some results over others, which clearly shows a lack of objectivity in a sector that is based precisely on objectivity. Special attention should also be paid to this bias and its existence since it tends to create a loop, a spiral in the world of science, a sector in which every new academic production is based on previous studies. Are we facing a spiral of disinformation generated by the publishers of certain scientific journals? It is important to be aware of this subject in order to detect and avoid the so-called 'scientific fraud' that has been reported in the press for years. What is more, they are biases that have not been generated by the psychology of citizens in the form of shortcuts to enable decision making—over and above the particular consequences that these might have—but by communication companies that have an interest in publication depending on what type of results are produced about health information.

It is also important to highlight the intentions of the research teams whose papers have been analysed in the defence of communication for health education, centred on the predominance of social media in the analysis that has revolutionised the way of accessing information about medicine and health [39]. In short, the importance of information and knowledge in the development of processes of political—or even economic—democratisation [40] indi-

cates that communication is the central axis of this development. But this axis should not be limited; health education plays an important role in the development of healthy, capable, free, and empowered societies.

Beyond the comments made on academic production in the mentioned fields, our research shows that there is a relationship between biases and disinformation when it comes to people's decision-making processes, including for health-related decisions. It also confirms that the existence of cognitive biases generates effects such as those mentioned in this article: interference in decision making, danger to the well-being of the planet, and a threat to public health. At the same time, it underlines the WHO's position by referring to an infodemic and the impossibility of addressing health problems—among others—without first solving the issue of disinformation.

This study has been conducted to identify the biases used in the era of misinformation related to health topics. Misinformation is not simply bad or false information, but rather selective information that circulates even among isolated and disconnected groups. Disinformation is not simply bad or false information directed at communities of users, but also selective information circulating among isolated and disconnected groups. One of the important contributions of this study is that it has worked with a novel and innovative methodology, through a methodology with a bibliometric approach that also combines social network and discourse analyses of abstracts of international academic literature. The great contribution is that it sheds light on the relationship between cognitive biases and health, exploring topics of interest to society such as vaccine hesitancy and the COVID-19 pandemic. Future research will focus on understanding and analyzing what makes communities and individuals less susceptible to manipulation. We will rely on data that can help mitigate mass manipulation.

Beyond the search for this relationship between biases, it is clear that the academic literature has dedicated space to articles or papers that have observed how these subjects have developed during times of pandemic.

**Author Contributions:** Conceptualization: C.P.-S.; Methodology: A.P.-d.-A.-M.; Software: C.P.-S., L.E.-E. and A.P.-d.-A.-M.; Validation: C.P.-S. and A.P.-d.-A.-M.; Formal analysis: A.P.-d.-A.-M.; Data curation: A.P.-d.-A.-M., C.P.-S. and L.E.-E.; Writing—Preparation of the original draft: C.P.-S., L.E.-E. and A.P.-d.-A.-M.; Writing, review & edition: C.P.-S., L.E.-E. and A.P.-d.-A.-M.; Abstract, title and keywords: C.P.-S.; Introduction and question research: C.P.-S. and L.E.-E.; Methodology: A.P.-d.-A.-M.; Results: A.P.-d.-A.-M.; Discussion and conclusions: L.E.-E. and C.P.-S.; Display: C.P.-S., L.E.-E. and A.P.-d.-A.-M.; Supervision: C.P.-S. and A.P.-d.-A.-M.; Project management: C.P.-S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is part of the R+D+i research project “Pseudoscience, conspiracy theories, fake news and media literacy in health communication”—COMSALUD, Reference PID2022-142755OB-I00, funded by the [State Research Agency (AEI in Spanish)]. It is also part of the scientific production of the Consolidated Research Group Gureiker (IT1496-22) from the System of the University of the Basque Country.

**Data Availability Statement:** The data can be obtained from Perez-de-Arriluzea-Madariaga, A. (2024). The Impact of Biases on Health Disinformation Research [Data set]. Zenodo. Available online: <https://doi.org/10.5281/zenodo.11120540>.

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

## References

1. Southwell, B.G.; Niederdeppe, J.; Cappella, J.N.; Oh, A.; Peterson, E.B.; Wen-Ying, S.C. Misinformation as a misunderstood Challenge to public health. *AJPM Am. J. Prev. Med.* **2019**, *57*, 282–285. [CrossRef] [PubMed]
2. West, J.D.; Bergstrom, C.T. Misinformation in and about science. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e1912444117. [CrossRef] [PubMed]
3. Romero Rodríguez, L. Hacia un estado de la cuestión de las investigaciones sobre desinformación/misinformación. *Corresp. Análisis* **2013**, *3*, 319–342. [CrossRef]
4. Van Dijk, T. Discurso y manipulación: Discusión teórica y algunas aplicaciones. *Rev. Signos* **2006**, *39*, 49–74. [CrossRef]
5. Parra, P.; Oliveira, L. Fakenews, una revisión sistemática de la literatura. *Obs. (OBS) J.* **2018**, *12*, 54–78.

6. Salaverría, R.; Buslón, N.; López-Pan, F.; León, B.; López-Goñi, I.; Erviti, M.C. Desinformación en tiempos de pandemia: Tipología de los bulos sobre la COVID-19. *Prof. Inf. Prof.* **2020**, *29*, 3. [CrossRef]
7. Echeverría, M.; Rodríguez Cano, C.A. La alfabetización digital activa la incredulidad de noticias falsas? Eficacia de las actitudes y estrategias contra la desinformación en México. *Revista de Comunicación* **2023**, *22*, 1–17. [CrossRef]
8. Noain-Sánchez, A. Desinformación y COVID-19: Análisis cuantitativo a través de los bulos desmentidos en Latinoamérica y España. *Estud. Sobre Mensaje Periodístico* **2021**, *27*, 879–892. [CrossRef]
9. Almansa-Martínez, A.; Fernández-Torres, M.J.; Rodríguez-Fernández, L. Desinformación en España un año después de la COVID-19. Análisis de las verificaciones de Newtral y Maldita. *Rev. Lat. Comun. Soc.* **2022**, *80*, 183–200. [CrossRef]
10. Kahneman, D. *Thinking Fast and Slow*; Farrar, Straus and Giroux: New York, NY, USA, 2011.
11. Cerezo Prieto, M. Sesgos cognitivos en la comunicación y prevención de la COVID-19. *Rev. Lat. Comun. Soc.* **2020**, *78*, 419–435. [CrossRef]
12. Kahneman, D.; Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica* **1979**, *47*, 263–291. [CrossRef]
13. Matute, H. Ilusiones y sesgos cognitivos. *Investig. Cienc.* **2019**, *518*, 55–57.
14. Leibenstein, H. Bandwagon, Snob and Veblen effects in the Theory of Consumer's Demand. *Q. J. Econ.* **1950**, *64*, 183–207. [CrossRef]
15. Freudenburg, W.R. Risk and recreancy: Weber, the division of labor, and the rationality of risk perceptions. *Soc. Forces* **1993**, *71*, 909–932. [CrossRef]
16. Lewicki, P. Self image bias in person perception. *J. Personal. Soc. Psychol.* **1983**, *45*, 384. [CrossRef]
17. Nickerson, R.S. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **1998**, *2*, 175–220. [CrossRef]
18. Oxford Learner's Dictionaries. Available online: <https://www.oxfordlearnersdictionaries.com/definition/english/post-truth?q=post+truth> (accessed on 28 March 2024).
19. D'Ancona, M. *Posverdad. La Nueva Guerra Contra La Verdad y Cómo Combatirla*; Alianza Editorial: Madrid, Spain, 2019.
20. Morales-i-Gras, J. Cognitive Biases in Link Sharing Behavior and How to Get Rid of Them: Evidence from the 2019 Spanish General Election Twitter Conversation. *Soc. Media + Soc.* **2020**, *6*, 205630512092845. [CrossRef]
21. Barzilai, S.; Chinn, C.A. A review of educational responses to the "post truth" condition: Four lenses on "post truth" problems. *Educ. Psychol.* **2020**, *55*, 107–119. [CrossRef]
22. Motta, M.; Stecula, D.; Farhart, C.E. How Right-Leaning Media Coverage of COVID-19 Facilitated the Spread of Misinformation in the Early Stages of the Pandemic in the U.S. *Can. J. Political Sci./Rev. Can. Sci. Polit.* **2020**, *53*, 335–342. [CrossRef]
23. Follari, R. La post-verdad contra la ciencia. *Rev. Guillermo Ockham* **2023**, *21*, 5–6. [CrossRef]
24. Thelwall, M.; Sud, P. A comparison of title words for journal articles and Wikipedia pages: Coverage and stylistic differences? *Prof. Inf.* **2018**, *27*, 49–64. [CrossRef]
25. Feng, F.; Yang, Y.; Cer, D.; Arivazhagan, N.; Wang, W. Language-agnostic BERT Sentence Embedding. *arXiv* **2007**, arXiv:2007.01852. [CrossRef]
26. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *arXiv* **2019**, arXiv:1908.10084. [CrossRef]
27. Hagberg, A.; Swart, P.J.; Schult, D.A. Exploring network structure, dynamics, and function using NetworkX. In Proceedings of the Conference: SCIPY 08, Pasadena, CA, USA, 21 August 2008; Varoquaux, G., Vaught, T., Millman, J., Eds.; Los Alamos National Laboratory (LANL): Los Alamos, NM, USA, 2008; pp. 11–15.
28. Blondel, V.D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *10*, P10008. [CrossRef]
29. Newman, M.E.J. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8577–8582. [CrossRef] [PubMed]
30. Jacomy, M.; Venturini, T.; Heymann, S.; Bastian, M. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE* **2014**, *9*, e98679. [CrossRef]
31. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
32. Oesper, L.; Merico, D.; Isserlin, R.; Bader, G.D. WordCloud: A Cytoscape plugin to create a visual semantic summary of networks. *Source Code Biol. Med.* **2011**, *6*, 7. [CrossRef]
33. D'Errico, F.; Cicerelli, P.G.; Corbelli, G.; Paciello, M. Addressing racial misinformation at school: A psycho-social intervention aimed at reducing ethnic moral disengagement in adolescents. *Soc. Psychol. Educ.* **2023**. [CrossRef]
34. Fernandez-Luque, L.; Imran, M. Humanitarian health computing using artificial intelligence and social media: A narrative literature review. *Int. J. Med. Inform.* **2018**, *114*, 136–142. [CrossRef]
35. Peng, W.; Lim, S.; Meng, J. Persuasive strategies in online health misinformation: A systematic review. *Inf. Commun. Soc.* **2023**, *26*, 2131–2148. [CrossRef]
36. Roitero, K.; Soprano, M.; Portelli, B.; Spina, D.; Della Mea, V.; Serra, G.; Mizzaro, S.; Demartini, G. The COVID-19 Infodemic. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Virtual, 19–23 October 2020; pp. 1305–1314. [CrossRef]
37. Soskolne, C.L.; Kramer, S.; Ramos-Bonilla, J.P.; Mandrioli, D.; Sass, J.; Gochfeld, M.; Cranor, C.F.; Advani, S.; Bero, L.A. Toolkit for detecting misused epidemiological methods. *Environ. Health* **2021**, *20*, 90. [CrossRef] [PubMed]

38. Concha, D.; Bilbao-Ramírez, M.A.; Gallardo, I.; Rovira, R.; Fresno, A. Sesgos cognitivos y su relación con el bienestar subjetivo. *Salud Soc.* **2016**, *3*, 115–119. [[CrossRef](#)]
39. Peñafiel, C.; Ronco, M.; Echegaray, L. ¿Cómo se comportan los jóvenes y adolescentes ante la información de salud en Internet? *Rev. Española Comun. Salud* **2016**, *7*, 167–189.
40. Martín-Barbero, J. Retos culturales: De la comunicación a la educación. *Nueva Soc.* **2016**, *169*, 33–43.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.