https://doi.org/10.35662/unine-thesis-3015

The content and spread of conspiracy theories

PhD thesis submitted to the Faculty of Science
Institute of Work and Organisational Psychology
University of Neuchâtel
Switzerland

For the degree of PhD in Psychology
by
Alessandro Miani

Dissertation committee:
Prof. **Adrian Bangerter**, thesis director, UniNE
Prof. **Jacques Savoy**, UniNE
Prof. **Lonneke van der Plas**, Idiap, Martigny
Dr. **Pascal Wagner-Egger**, Université de Fribourg
Prof. **Thomas Hills**, University of Warwick, UK

Defense: 27^{th} February, 2023

University of Neuchâtel, 2023

unine
Université de Neuchâtel
Faculté des sciences

Rue Emile-Argand 11
CH-2000 Neuchâtel
doctorat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Alessandro MIANI

Titre :

"The content and spread of conspiracy theories"

sur le rapport des membres du jury composé comme suit:

- Prof. Adrian Bangerter, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Jacques Savoy, Université de Neuchâtel, Suisse
- Prof. Lonneke van der Plas, IDIAP, Suisse
- Dr Pascal Wagner-Egger, Université de Fribourg, Suisse
- Prof. Thomas Hills, University of Warwick, UK

Neuchâtel, le 27 mars 2023

Le Doyen, Prof. R. Bshary

This dissertation is based on the following manuscripts:

1. Miani A., Hills T. & Bangerter A. (2021). LOCO: The 88-million-word Language of Conspiracy Corpus. *Behavior Research Methods*, 54, 1794-1817. https://doi.org/10.3758/s13428-021-01698-z

2. Miani A., Hills T. & Bangerter A. (2022). Interconnectedness and (in)coherence as a signature of conspiracy worldviews. *Science Advances*, 8(43). https://doi.org/10.1126/sciadv.abq3668

3. Miani A., van der Plas L. & Bangerter A. (In preparation). *Overinclusive thinking in conspiracy texts.*

4. Miani A., Schöpfer T., Hills T. & Bangerter A. (In preparation). *Selective exposure drives traffic to conspiracy websites.*

Abstract

Belief in conspiracy theories (CTs) is associated with numerous societal harms, including violence, vaccine refusal, and political extremism. Given the speed and intensity with which information spread through the internet, there is a pressing need to understand what CTs are and how they circulate. Across four studies, This PhD thesis works towards this goal: understanding the content and spread of CTs.

In Study 1, we developed the largest corpus of CTs available today, LOCO, that allows to explore the content and spread of CTs. An analysis of linguistic content showed that conspiracy texts are focused on deception, power, and dominance. Conspiracy webpages that rely on prototypical conspiratorial language are more shared on Facebook.

In Study 2, we found that conspiracy texts are more interconnected, more topically heterogeneous, and more similar to one another. These results provided strong empirical support for an overarching conspiracy worldview in conspiracy narratives.

In Study 3, we developed measures to assess elements of divergent and convergent thinking in texts. We show that conspiracy texts were more original, semantically divergent, and sophisticated, but less appropriate to their context and less variable compared to those in non-conspiracy texts. Results point to an imbalance between divergent and convergent thinking and may explain the accumulation of CTs within people's belief systems.

In Study 4, we devised a field study to compare the impact of social media and individual cognitive biases on online browsing behavior towards websites classified on ideological types and strength. We found that as the websites' conspiratorial ideology increases, the contribution from individual cognitive biases increases at the expense of traffic from social media.

In sum, results obtained from this thesis have practical implications. As for the content of CTs, the presence of conspiracy mentality that emerges from texts represents a foreseeable possibility to develop algorithms for the automatic detection of CTs online. In regard to the spread of CTs, knowing that individual cognitive biases drive access to conspiracy websites suggests that individual-level interventions, such as improving critical thinking, should be prioritized in the fight against the spread of CTs.

Keywords: *conspiracy theories; natural language processing (NLP); content; transmission*

Résumé

La croyance aux théories du complot (TC) est associée à de nombreux préjudices sociétaux, notamment la violence, le refus des vaccins et l'extrémisme politique. Compte tenu de la vitesse et de l'intensité avec lesquelles les informations se propagent sur l'internet, il est urgent de comprendre ce que sont les TC et comment elles circulent. À travers quatre études, cette thèse de doctorat vise à atteindre cet objectif: comprendre le contenu et la diffusion des TC.

Dans l'étude 1, nous avons développé le plus grand corpus de TC disponible aujourd'hui, LOCO, qui permet d'explorer le contenu et la diffusion des TC. Une analyse du contenu linguistique a montré que les textes conspirationnistes sont axés sur la tromperie, le pouvoir et la domination. Les pages web conspirationnistes qui s'appuient sur un langage conspirationniste prototypique sont davantage partagées sur Facebook.

Dans l'étude 2, nous avons constaté que les textes conspirationnistes sont plus interconnectés, plus hétérogènes sur le plan thématique et plus semblables les uns aux autres. Ces résultats ont apporté un soutien empirique solide à l'idée d'une vision globale du monde conspirationniste dans les récits conspirationnistes.

Dans l'étude 3, nous avons développé des mesures pour évaluer les éléments de la pensée divergente et convergente dans les textes. Nous montrons que les textes conspirationnistes sont plus originaux, sémantiquement divergents et sophistiqués, mais moins adaptés à leur contexte et moins variables que les textes non conspirationnistes. Les résultats indiquent un déséquilibre entre la pensée divergente et la pensée convergente et peuvent expliquer l'accumulation de TC dans les systèmes de croyance des gens.

Dans l'étude 4, nous avons conçu une étude de terrain pour comparer l'impact des médias sociaux et des biais cognitifs individuels sur le comportement de navigation en ligne vers des sites web classés en fonction de leur type et de leur force idéologique. Nous avons constaté qu'à mesure que l'idéologie conspirationniste des sites web augmente, la contribution des biais cognitifs individuels s'accroît au détriment du trafic provenant des médias sociaux.

En résumé, les résultats obtenus dans le cadre de cette thèse ont des implications pratiques. En ce qui concerne le contenu des TC, la présence d'une mentalité conspirationniste qui émerge des textes représente une possibilité prévisible de développer des algorithmes pour la détection automatique des TC en ligne. En ce

qui concerne la diffusion des TC, le fait de savoir que les biais cognitifs individuels favorisent l'accès aux sites web conspirationnistes suggère que les interventions au niveau individuel, telles que l'amélioration de la pensée critique, devraient être prioritaires dans la lutte contre la diffusion des TC.

Acknowledgements

Four years have already gone since I arrived in Neuchâtel with my MacBook and an LCD monitor with a VGA cable. I had very high hopes when I started. It turned out, I spent great time exploring and learning in so many respects and being under constant intellectual stimulation. It was a Luna Park. Now, at the end, I can say the PhD ended well above expectations: paid by the elite for my service, I have now updated my MacBook and use an HDMI monitor.

Clearly, not all the merits are mine; there are powerful forces behind this work. First and foremost, my PhD mentor. Thanks Adrian for trusting and taking me as your PhD student, for your guidance, for your constant and invaluable support, and for leading me in this great work. These years have been incredibly formative for me. Thanks, I was a very lucky PhD student!

The four papers that comprise this thesis couldn't be realized without the precious inputs from my co-authors: Thomas, Lonneke, and Théo, thanks!

To my PhD committee, Jacques Savoy, Lonneke van der Plas, Pascal Wagner-Egger, and Thomas Hills: thank you for taking time out of your busy schedules to critique this work. Thanks Thomas for joining us on site, I really appreciate it.

Pascal, it was your article on teleological thinking that revitalized my passion on conspiracy theories; I'm here ***because*** *of it!*

Thanks Maestro Lorenzo for guidance and support. None of these manuscripts will be published in Basel-based journals, I promise.

Thanks Garance for your ever-ready helpfulness at the IPTO.

Lastly, many thanks go to my family and friends for their long-lasting support. Thanks Milli, for providing that kind of emotional support that people write songs about, for your constant encouragement, and for understanding these last days of eremitism.

Neuchâtel, 19 April, 2023

February 2019

February 2023

they want to sound smart. [...]
Instead of using good words like smart, they choose sophisticated or erudite

(Silvia, 2007, p. 60)

Contents

1 General Introduction

In AD 64, the *Great Fire of Rome* burned for nine days destroying ten of Rome's fourteen districts, causing the death of countless people, and making homeless half of the Rome's population. At the time of the fire, the Emperor Nero was 35 miles away, in his villa in the coastal city of Antium. When he returned to Rome, he opened his private and public buildings to provide shelter for the homeless, and promptly started to rebuild the city. Food was brought from neighboring towns and the price of corn was cut. However, to meet the costs inflicted by the fire and to find the funds for the reconstruction, Nero's government increased taxation, causing discontent in the population.

In the aftermath, different accounts started to spread, mostly accusing Nero of starting the fire to clear the land so to rebuild Rome in his image —the *Neropolis*, which included his *Domus Aurea*, a large complex of 200 acres in the heart of Rome. Other rumors described him watching the fire from his palace on the Palatine Hill while singing and playing the lyre. Nero, on his side, was not happy to be the subject of these allegations, and blamed the new Christian community of conspiring, initiating the empire's persecution against the Christians. Following the growing discontent towards Nero's government, the year after the Great Fire, a plot to murder the emperor (the Pisonian conspiracy) was discovered. These rumors were enduring and even exaggerated by Cassius Dio, 165 years later (for a more elaborate discussion, see Brotherton, 2016). Still today, the name of Nero is associated with negative sentiments and with the great fire of Rome (e.g., the software for burning CD ROM, see Figure 1.1).

The example of the Great Fire of Rome was used to illustrate several prominent aspects of conspiracy theorizing. Conspiracy theories are responses to significant social events and proliferate in times of crisis and uncertainty. Conspiracy theories are enduring, lasting for decades (and even centuries). Some conspiracies are true and other are unverified theories. Conspiracy theories can target people in power (e.g., Nero) but also minorities (e.g., the Christian community at that time), generating (but also stemming from) outgroup hostility sentiments.

Figure 1.1: An ancient (circa AD 1997) picture of Nero burning Rom(e)

Conspiracy theories (CTs) are alternative explanations for significant social events with claims of secret plots by powerful actors at the expense of an unwitting population (Douglas et al., 2019). CTs are widely spread across cultures and historical times (Bangerter et al., 2020; Brotherton, 2016; Butter & Knight, 2023; van Prooijen & Douglas, 2017; West, 2003). Many socially relevant past (e.g., the Moon Landing) or contemporary (e.g., COVID-19) events, or scientific facts (e.g., climate change) and physics phenomena (e.g., contrails generated by airplanes), have the potential to generate CTs. CTs are also country-specific (e.g., including national popular figures) and involve any field, from medicine (e.g., vaccine-autism link) to music (e.g., Paul McCartney is dead, and Elvis Presley is alive). Different CTs coexist as explanations for a specific event and can be also in contradiction (e.g., COVID-19 intentionally created in a lab or accidentally leaked). Yet, despite their diversity, almost (if not) all CTs coherently and chronically denounce a hidden malevolent plan.

CTs are unlikely (Wagner-Egger et al., 2019). Simulation studies suggest that even if CTs were true, they would be inclined to fail within short time (Grimes, 2016). This is because the high number of people (claimed to be) involved in keeping the conspiracy secret would increase the odds of leaking, causing the detection of the conspiracy. This suggests that as the number of actors involved in a conspiracy increases, as CTs claim, the time to discover the conspiracy decreases. For example, according to Grimes (2016), the moon-landing conspiracies, if true, would have been discovered in less than 4 years.

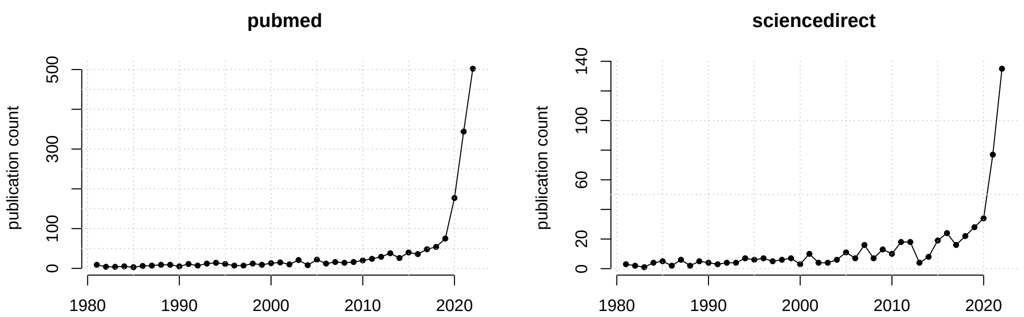
Although a conspiracy research agenda started in early 2000s, with some works in the '90s and before,[1] it is only recently, with the COVID-19 outbreak in 2020, that CTs suddenly attracted the attention of lay people and researchers. Following the pandemic, people recognized the dangerous effects associated

[1] see e.g., Goertzel (1994); Pigden (1995); Hofstadter (1964)

with the spread of CTs such as reduced intention to adopt health preventive measures and vaccination[2]. In a race to develop a COVID-19 vaccine and attempting to contain the spread of the virus via social distancing, researchers from all over the world started to realize (and focus on) how the spread of CTs has a negative impact on the social world.

Figure 1.2 shows the interest of academia in CTs over time, by visualizing the total publication count of papers related to CTs (or conspiracies) from 1980 to 2022 gathered from two databases of published scientific articles: PubMed (left) and ScienceDirect (right).[3] As the figure shows, the year 2020 was characterized by a remarkable increase in publication count related to conspiracies. Moreover, only in the last year, we saw a proliferation of reviews (e.g., Douglas & Sutton (2023) in *Annual Review of Psychology* as well as Biddlestone, Green, et al. (2022); Turska-Kawa et al. (in press)), meta-analyses (Biddlestone, Green, et al., 2022; Bierwiaczonek et al., 2022; Stasielowicz, 2022; Yelbuz et al., 2022), and themed issue (e.g., in *Current Opinion in Psychology*, Volume 47, 2022, edited by Jan-Willem van Prooijen and Roland Imhoff; and in *Journal of Experimental Social Psychology*, Volumes 103, 2022, edited by Kai Sassenberg, Paul Bertin, Karen Douglas, and Matthew Hornsey).

Figure 1.2: Academic interest over time on conspiracy theories



[2]Although vaccine-related CTs were not a novel phenomenon, see e.g., Gangarosa et al. (1998), P. Davies (2002)

[3]Searches were done by querying: "`conspiracy OR conspiratorial OR conspiracism OR conspiracies`" in titles, abstracts, and keywords of articles indexed in the two databases. Although numbers are different, with PubMed being more inclusive than ScienceDirect, the two time-series (log-transformed) correlate at $r(40) = .87$, p $< .001$ ($r = .98$ without log transformation).

1.1 Conspiracy mentality

While some of the conspiracies may be true (see e.g., Nixon's Watergate, MK-Ultra, Tuskegee Syphilis experiments), CTs are false for the majority, or at best unverified, and in any case very unlikely (Wagner-Egger et al., 2019). It is therefore important to define the boundaries of CTs research, distinguishing *conspiracy theories* from *actual conspiracies* and describing the difference between *conspiracy belief* and *conspiracy mentality*.

A **conspiracy**, according to the Oxford Dictionary of English, refers to "*a secret plan by a group to do something unlawful or harmful*". Researchers suggest that such plots are carried on by powerful actors (Keeley, 1999; Pigden, 1995). There are cases of real conspiracies such as the Nixon's Watergate, the MK-Ultra project, the Tuskegee Syphilis experiments, or the Pisonian conspiracy against Nero. In many cases, to be defined as such, a conspiracy should be societally relevant, malevolent and secret. Therefore, it follows that Santa Claus and the Tooth Fairy, for example, while they work in secret, they do not act malevolently, so they cannot be considered a conspiracy.[4] Similarly, a corner-shop robbery, although criminal, cannot be considered a conspiracy because it is not societally relevant.

Different from actual conspiracies, **conspiracy theories** are *attempts to explain* significant social events with claims of conspiracy, which may or may not be true. It is important to highlight that while a conspiracy refers to an act, a conspiracy theory refers to a perception (Pipes, 1999). Also, while actual conspiracies are grounded in reality, conspiracy theories often involve omniscent and omnipotent agents (Franks et al., 2013). Actual and theorized conspiracies also differ at the methodological (see e.g., Wagner-Egger, 2022) and narrative (Tangherlini et al., 2020) levels.

Conspiracy belief refers to a *belief in a specific conspiracy theory* (e.g., the 9/11 terrorist attack is an inside job; Lady Diana princess of Wales faked her death). Although popular, conspiracy theories are not equally believed. It is likely that the plausibility of a theory determines its popularity. For example, while a large proportion of Americans, still nowadays, believe that President Kennedy was not assassinated by Lee Harvey Oswald acting alone (see section 1.4.1), I suspect that only few people believe that lizard aliens rule the world.

[4] but see the *conspiracy of angels* (Walker, 2013), where secret forces work in secret to improve people's lives

A **Conspiracy mentality** (or *mindset*, *thinking*, *worldview*) is an underlying tendency to prefer conspiratorial explanations instead of official narratives. Conspiracy mentality is different from belief in a single CT. While a CT arises in response to a specific event attempting to provide an explanation, the conspiracy mentality is a disposition to interpret the world according to a conspiratorial worldview (Imhoff, Bertlich, et al., 2022). As such, different conspiracies can be endorsed within a conspiracy mentality (Goertzel, 1994; Lewandowsky et al., 2018; Lukić et al., 2019; Swami et al., 2011; Wal et al., 2018; Wood et al., 2012).

In Psychology research, different tools (i.e., scales) have been developed to measure conspiracy belief and conspiracy mentality (see list below). These scales vary in terms of specificity. They can focus on single events, such as the 9/11 terrorist attack, measuring belief in different versions of a single event (1). There are scales that measure belief in different CTs (2). There are also scales that measure, in more abstract terms, generic conspiracy belief (3). Items in these scales do not refer to specific events but are broadly formulated so to tap to a variety of specific cases (e.g., using *virus* instead of *Ebola* or *HIV*). There are scales that measure conspiracy mentality as a tendency to endorse conspiracy ideation (4). As such, items in these scales are formulated in broad terms, mostly tapping into core elements of CTs such as politicians, governments, and organizations that have influence, monitor citizens, and/or hide their decisions. Lastly, there are very general scales such as the single-item scale that measure the core element of conspiratorial mentality, i.e., authorities are deceptive (5).

1. **Specific conspiracy theory about a single event** (Swami et al., 2010): *"The Pentagon was not hit by a passenger aircraft but something smaller, possibly a missile"*; *"The World Trade Centre towers were brought down by a controlled demolition"*.
2. **Specific conspiracy theories** (Bruder & Manstead, 2009): *"There are clandestine religious groups that carry out human sacrifice"*; *"Governments have deliberately spread HIV amongst ethnic minorities"*.
3. **Generic conspiracy** (Brotherton et al., 2013): *"The spread of certain viruses and/or diseases is the result of the deliberate, concealed efforts of some organization"*; *"A small, secret group of people is responsible for making all major world decisions, such as going to war"*.
4. **Conspiracy mentality** (Bruder et al., 2013): *"Government agencies*

closely monitor all citizens"; *"Politicians usually do not tell us the true motives for their decisions"*.

5. **Single-Item conspiracy Scale** (Lantian et al., 2016): *"The official version of the events given by the authorities very often hides the truth"*.

One of the most recurrent findings in CTs research, since the seminal work of Goertzel (1994), is that people who believe in one specific CT tend to believe in other CTs. For example, believing that the AIDS virus was deliberately engineered in a government laboratory is associated with being more likely to believe that the FBI was involved in Martin Luther King's assassination. This phenomenon refers to the *conspiracy mentality*, the tendency to adopt CTs as explanation for important societal phenomena (Swami et al., 2010). The *conspiracy mentality* is such a strong worldview that can accommodate even logically incompatible and fictitious CTs (Lewandowsky et al., 2018; Lukić et al., 2019; Swami et al., 2011; Wood et al., 2012). For example, Swami et al. (2011) reported that the more participants believed in popular CTs (e.g., about 9/11 terrorist attack), the more they perceived as real a set of fictitious CTs created *ad hoc* for their study. In another work, individuals who believe that Lady Diana was killed by government agencies (e.g., MI6) were also more likely to believe that she faked her death (Wood et al., 2012).[5]

The conspiracy mentality is a self-sustaining worldview comprised of a network of mutually supportive beliefs (Wood et al., 2012). Each theory serves as justification for new theories. For example, the unethical behavior of the tobacco industry in the past can explain the wrongdoings of the pharmaceutical industry today (e.g., delivering harmful vaccines, Albarracín, 2022). With the accumulation of CTs within a system of belief, then, also implausible CTs can be endorsed, e.g., lizard aliens rule the world and the earth is flat. In doing so, people with high conspiracy mentality have developed a worldview that explain a world dominated by deceitful authorities that conceal information. As such, events, regardless of their nature, can be turned into a conspiracy. The underlying psychological mechanism is distrust to authorities: if authorities deceived us with the 9/11, then they can deceive us with vaccines. Such an overarching worldview is capable to support incoherent, incompatible, and

[5]Note that the study of Wood et al. (2012) has been criticized due to the correlation coefficient being driven by few participants believing in both statements (see e.g., https://climateaudit.org/2013/11/07/more-false-claims-from-lewandowsky). Pascal Wagner-Egger also observed that only around 1% of people (in a sample of N > 700) endorsed incompatible CT items (personal communication).

fictitious theories as long as these theories are coherent with an overarching worldview that sees authorities as deceitful (Goertzel, 1994; Lewandowsky et al., 2018; Lukić et al., 2019; Swami et al., 2011; Wal et al., 2018; Wood et al., 2012).

It must be noted that the conspiracy mentality is a more stable construct compared to belief in specific CTs (Imhoff, Bertlich, et al., 2022). In psychological research, conspiracy mentality scales allow to assess conspiracism free from cultural and temporal differences (Imhoff, 2022). For example, CTs surrounding the death of Lady Diana might not be relevant outside the UK; CTs surrounding the death of Slobodan Milošević might not be relevant outside ex-Yugoslavia; CTs surrounding the death of president Kennedy might be old and not anymore relevant in 2020s. Differently, the idea that authorities (e.g., the pharmaceutical industry or governments) are deceptive is not related to specific time and space.

1.2 Who believes in conspiracy theories

The largest part of psychology research on CTs has been focused on understanding *who* is most likely to believe in CTs. These studies provide a psychological profile of the conspiracy believer. Administering scales that measures belief in specific or generic CTs or conspiracy mentality (see section 1.1), scholars have generated a large corpus of literature describing the nomological network of conspiracy belief/mentality. In the following, I provide an overview of the predictors of conspiracy thinking mostly related to this work. Note that, although not reviewed here, a large corpus of literature has focused on, for example, personality traits (Biddlestone, Green, et al., 2022; see e.g., Goreis & Voracek, 2019; Turska-Kawa et al., in press) and more recently on COVID-19-related behavior and attitudes (see e.g., Turska-Kawa et al., in press). These studies will not be reviewed here.

1.2.1 Demographics

Belief in CTs is associated with a range of demographic factors (Douglas et al., 2017, 2019; Douglas & Sutton, 2023). Being a male,[6] unmarried, unemployed, and with low income is associated with belief in CTs (Freeman & Bentall, 2017,

[6]But note that results on gender differences are inconsistent, see e.g., van Prooijen et al. (2022).

2017; Uscinski & Parent, 2014). Also, being part of an ethnic minority (Freeman & Bentall, 2017; Goertzel, 1994), such as Black and Hispanic individuals in the US, is associated with belief in CTs. Similarly, members of low-status social groups are more likely to endorse CTs than members of high-status groups (Abalakina-Paap et al., 1999; Goertzel, 1994; Uscinski & Parent, 2014). Minority and low-status groups are often stigmatized and discriminated, causing distress in people. For example, respondent from minority groups in the sample of Goertzel (1994) were in fact more insecure about their job opportunities.

People from disadvantaged or stigmatized groups tend to believe that dominant groups are conspiring against them (Davis et al., 2018; Thomas & Quinn, 1991). Then, it is more likely that people from a stigmatized group believe in CTs directed at their own group if they personally have experienced discrimination (Sharon Parsons, William Simmons, Fr, 1999; Simmons & Parsons, 2005). Thus, belief in outgroup conspiracies can be fueled by past experiences such as when one's social group has been treated unfairly.

Education is another important predictor (Freeman & Bentall, 2017), although the effect of education seems to be mediated by analytic thinking and belief in simple solutions (van Prooijen, 2017). People believing in CTs tend to have also lower verbal (Fiagbenu, 2022), crystallized (Stieger et al., 2013), and self-assessed (Swami et al., 2011) intelligence. Furthermore, news media literacy has been found to decrease CTs endorsement (Craft et al., 2017) and CTs spread (e.g., Twitter, Mosleh et al., 2021).

These demographic factors can also speak to the psychological needs that motivate people to pursue CTs. For instance, low education can explain the need for knowledge, and low income is associated with existential threats to security. The latter is important to understand the importance of the social context. For example, cross-national studies (Alper, 2022; Hornsey & Pearson, 2022; Imhoff, Zimmer, et al., 2022; Salvador Casara et al., 2022) show that conspiracy mentality is positively associated with economic inequality, ethnic fragmentation, and threat to peace, while it is negatively associated with countries' GDP (Gross domestic product), democracy and development indexes, level of corruption, and political stability.

All these factors point to situational stress and life history. People with a history of stress and experiencing stress are more likely to believe in CTs. For example, childhood family experiences, low physical and psychological well-being, and a range of psychiatric criteria (e.g., alcohol and drug abuse and dependence, depression, panic disorder, attention disorder, and social phobia) have been found to be associated with belief in CTs (Freeman & Bentall, 2017), confirming that stress is an important factor of belief in CTs (Swami et al., 2016).

Overall, results on demographic predictors of belief in CTs are inconclusive and no meta-analyses nor reviews have been published in this regard (except Freeman & Bentall, 2017). I suspect that searching for demographic predictors of belief in CTs would be inconclusive. Older people, for example, tend to have lower digital literacy (Guess & Munger, 2022) hence are more likely to go down the *rabbit hole* (Sutton & Douglas, 2022) and spread CTs (Guess et al., 2019), eventually believing them (Pennycook et al., 2018). However, age correlates with education, simply because youngest people are likely to still be on education, which would act as protective factors against CTs belief. Nevertheless, older people are more likely to endorse conservative political views (Geys et al., 2022; Tilley & Evans, 2014; but see also Peterson et al., 2020), which often correlates with belief in (Imhoff, Zimmer, et al., 2022) and spread of (Guess et al., 2019) CTs. Therefore, a crude correlation between age and belief in CTs would be uninformative *per se* —yet useful to be included as covariate for statistical control.

1.2.2 Motives

People subscribe to CTs prompted by *epistemic* (to achieve knowledge and certainty), *existential* (to feel safe and in control), and *social* (to maintain a positive image of the self or group) needs (Douglas et al., 2017, 2019; Douglas & Sutton, 2023). In the following, I provide an overview of the motives that lead people to endorse CTs.

1.2.2.1 ***Epistemic***

Complex events (e.g., a pandemic) require complex and elaborate explanations. CTs allow to simplify the real-world complexity, gaining knowledge and reducing uncertainty with the associated distress caused by loss of feeling in control. Conspiracy beliefs have been linked to the need for cognitive closure (Leman & Cinnirella, 2013; Marchlewska et al., 2018), especially when events are important and lack a simple clear-cut official explanation. The need for closure is a desire for "*an answer on a given topic, any answer*" (Webster & Kruglanski, 1994). People with high need for closure tend to jump to conclusion quicker (Colbert & Peters, 2002), which is a bias characterized by excessive intuition, and reduced analysis in information processing and may lead to sub-optimal decision-making (Zander-Schellenberg et al., 2021)

CT believers, with high need for closure, attempt to make sense of events (Bangerter et al., 2020; Franks et al., 2013). They are moved to search for hidden motives in search for truth (which, according to believers, is concealed by official accounts and mainstream media), quickly accepting alternative scenarios. In doing so, believers tend to identify meaningful relationships among randomly co-occurring events (Wal et al., 2018), confuse aspects of reality such as believing that prayers have the capacity to heal (Lobato et al., 2014), and believe in the paranormal (Darwin et al., 2011).

Cognitive biases might lead to endorse CTs over official accounts in order to satisfy an epistemic need. These are: conjunction fallacy (Brotherton & French, 2014; Dagnall et al., 2017), which is an error of probabilistic reasoning whereby people overestimate the likelihood of co-occurring events, hypersensitive agency detection (Brotherton & French, 2015; Tempel & Alcock, 2015), the tendency to attribute agency and intentionality where it does not or it is unlikely to exist, and teleological thinking (Wagner-Egger et al., 2018), the attribution of purpose and a final cause to natural events and entities.

1.2.2.2 *Existential*

Besides needs for knowledge, causal explanations for complex events, such as those provided by CTs, allow people to feel safe and secure in their environment which, thanks to CTs, appears to be more controllable. Therefore, people subscribe to CTs to manage the existential threat stemming from an unpredictable and complex world (van Prooijen, 2020). In fact, belief in CTs increases particularly following stressing societal events that elicit existential threat such as terrorist attack, natural disasters, economic crises, wars, and rapid societal change (Pipes, 1999; van Prooijen & Douglas, 2017).

Studies show that when people's sense of control is lost they tend to believe in CTs (Bruder et al., 2013; Nyhan, 2017; Uscinski & Parent, 2014; van Prooijen & Acker, 2015). Belief in CTs is also associated with feelings of powerlessness (Abalakina-Paap et al., 1999; Pratt, 2003; Zarefsky, 2014), state and trait anxiety (Grzesiak-Feldman, 2007, 2013; Radnitz & Underwood, 2017), anxious attachment style (Green & Douglas, 2018), existential anxiety (Newheiser et al., 2011), and stress (Freeman & Bentall, 2017; Swami et al., 2016). It is however debated whether lack of control increases belief in CTs and other types of illusory pattern perception. Two out of seven experiments conducted by van Elk & Lodder (2018) showed a decreased tendency to engage in magical thinking and CTs belief following a control threat manipulation. In another work (Dieguez et al., 2015), contrary to van Prooijen et al. (2018), conspiracy believers did not differ from non-believers in the rejection of randomness heuristic (i.e., the *nothing-happens-by-accident* bias). As the meta-analysis of van Elk & Lodder (2018) suggests, there might be evidence for publication bias in the literature.

Empirical research suggests that threats to control can increase belief in conspiracy theories about the government (van Prooijen & Acker, 2015), yet at the same time, threats to control may also increase people's support for that same government (A. C. Kay et al., 2008). This observation led van Prooijen (2020) to suggest a theoretical model in which existential threat increases people's motivation to make sense of their social and physical environment, which leads to search for CTs. However, this sense-making process leads to CTs only when the outgroup is salient. Hostility towards the outgroup is in fact an important characteristic of CTs. Uscinski et al. (2011) coded 100,000 letters to the editor of the New York Times, from 1897 to 2010, finding that CTs are often directed to outgroup enemies: when threat from external enemies was high, CTs focused more often on foreign actors (e.g., Stalin, the United Nations). These effects suggest that spreading conspiracy theories is a way to manage threat.

It is important to note that CTs can be source of existential threat themselves, stimulating further conspiracy theorizing and contributing to a generalized conspiracy mentality (van Prooijen, 2020). Longitudinal studies show that prolonged exposition to CTs, instead of satisfying epistemic and existential needs, frustrate people even more reinforcing the negative experience of anxiety, uncertainty, and existential threat, which in turn increases conspiratorial belief and subsequent search for CTs (Liekefett et al., 2021).

1.2.2.3 ***Social***

According to some scholars, the social component in CTs belief is paramount: "*if conspiracy beliefs were an individual process, no conspiracy theory would be alike*" (Albarracín, 2022, p. 1). In general, people desire to be part of a social group and to maintain a positive image of the self or the group which they are part of. CTs allow to satisfy needs for social inclusion, but, at the same time, offer the possibility to distinguish oneself and the own group from other individuals and groups. Endorsement and dissemination of CTs may be used to pursue social goals such as recruitment of newcomers, coordination of allies for cooperative actions, and signaling devotion to other group members (Marie & Petersen, 2022).

People who believe in CTs consistently score high in the desire to be unique (Imhoff & Lamberty, 2017; Lantian et al., 2017). Need for uniqueness might explain people's consumption and transmission of CTs to enhance personal and social identity (Sternisko et al., 2020) by feeling special (Tian et al., 2001). People share CTs on social media to obtain likes and reactions expecting to generate social engagement (Ren et al., 2023). CTs are in fact attention-grabbing narratives that elicit reactions of interest and excitement (van Prooijen et al., 2022).

People who believe in CTs are often stigmatized (Lantian et al., 2018), hence at risk of being socially excluded. Because of this, being part of a group is crucial to satisfy a need of belongingness. Online, CT believers join homogeneous clusters of communities that share the same interest, named *echo chambers* (Bessi, 2016; see e.g., Cinelli et al., 2021).

Group belonging can deviate into narcissism, which is an exaggerated view of the self or of the own social group. Studies show that the endorsement of CTs is associated with individual, collective, and nationalistic narcissism (Cichocka et al., 2016; Golec de Zavala & Cichocka, 2012; C. S. Kay, 2021; Marchlewska et al., 2019; see Hornsey et al., 2022 for an overview). High levels of narcissism allow to maintain positive esteem of the self and of the own social group, a superiority that needs to be recognized and respected by others (Golec de Zavala et al., 2009, 2019). Within social conflicts, CTs derive from the need to validate the own group image by denigrating the outgroup (Cichocka et al., 2016). Such a nationalistic, or group, attachment is further amplified when people experience situational threats or crisis situation, which lead to increase belief in CTs (Mashuri & Zaduqisti, 2014; van Prooijen & Douglas, 2017).

1.2.3 Political orientation

Several studies highlight that belief in CTs tends to be higher among people holding strong political ideology (van Prooijen et al., 2015), and, in general, more pronounced on the conservative side of the political spectrum (Imhoff, Zimmer, et al., 2022; Linden et al., 2021) —although others did not find strong evidence for partisanship (Uscinski, Enders, Diekman, et al., 2022).

van Prooijen et al. (2015) and Imhoff, Zimmer, et al. (2022) investigated the relationship between conspiracy mentality and political orientation testing both linear (tendency towards a specific political party) and quadratic (tendency towards an extreme ideology regardless of political position) functions. Across three studies, van Prooijen et al. (2015) found strong evidence for a quadratic effect of belief in CTs on political orientation yet evidence for a linear effect was weak. Imhoff, Zimmer, et al. (2022) tested this relationship across 26 countries for a combined sample size of more than one hundred thousand participants, finding both linear and quadratic effects. Overall, this suggests that conspiracy mentality increases as a function of ideology and this relationship tends to be stronger on the right.

1.2.3.1 *Extreme ideologies*

Studies consistently show a u-shaped relationship, although right-skewed, between political partisanship and belief in CTs (Imhoff, Zimmer, et al., 2022; see e.g., van Prooijen et al., 2015). Although substantial differences between left- and right-party ideologies, both ideologies are grounded in a similar underlying psychology at the extremities (Greenberg & Jonas, 2003). Both left and right political extremists rely on a relatively simplified mental processing style characterized by black-and-white dichotomous thinking, in which reality is rigidly classified either as good or bad (Greenberg & Jonas, 2003). People holding extreme political ideologies tend to display more ingroup favoritism than moderates, to blame other people rather than circumstances for negative events, to show intolerance to ambiguity, and to embrace a fundamentalist mindset characterized by paranoid suspiciousness and violence (Basit, 2021; Borum, 2014; Greenberg & Jonas, 2003; van Prooijen et al., 2015).

Important precursors of extreme ideologies are deception and injustice, which are the core elements of conspiratorial worldview (Midlarsky, 2011). In fact, there is an important conspiratorial component in political extremes (Allington

& Joshi, 2020; Bartlett & Miller, 2010; Prooijen & Kuijper, 2020; van Prooijen et al., 2015). Both conspiracy and extreme ideologies stem from feelings of uncertainty and fear (McCauley & Moskalenko, 2008; Prooijen et al., 2015; van Prooijen et al., 2015). The two ideologies attempt to make sense of the world so to regulate the uncertainties that people encounter in their life (van Prooijen & Krouwel, 2015). People at the extreme ideologies tend to display more ingroup favoritism than moderates (Greenberg & Jonas, 2003), which might explain the endorsement of CTs in response to ingroup threat and high scores in collective narcissism.

Parallels with radicalized groups has been suggested (Kruglanski et al., 2022). For example, members of different extremist groups frequently spread CTs (Bartlett & Miller, 2010), while both far-right and far-left violent groups use CTs in their propaganda (Basit, 2021). Individuals joining radical groups are moved by need for significance and identity with a group (Borum, 2014). Feelings of insignificance can be caused by situational factors that pose existential threat. Radical groups promise to restore significance by providing a sense of belongingness and heroism, promising to fight for a holy cause (Doosje et al., 2016).

It is important to note that such a notion of heroism echoes a model of conspiracy radicalization: from step to step, individuals move through an epistemic and spiritual journey moving towards a process of awakening/conversion, where few heroic figures act in proselytizing the majority of the "sheep" people (Franks et al., 2017). In the process of radicalization, alike a religious conversion (Franks et al., 2013), both extremists (Strozier et al., 2010) and conspiracy believers (Franks et al., 2017) experience a change that is not simply an endorsement of specific ideas, but it is transformative and comprehensive (Harambam & Aupers, 2017).

1.2.3.2 *On the right*

White conservative males are significantly more likely than other Americans to endorse and spread CTs (Guess et al., 2019; McCright & Dunlap, 2011; Min, 2021). In the US context, Linden et al. (2021) show that conservatives were more likely than liberals and extreme Liberals to endorse specific CTs and displaying a conspiratorial worldviews. Still in the US, Uscinski, Enders, Diekman, et al. (2022) found that the strongest positive predictors of both specific CTs and conspiracy mentality were populism, Manicheanism, Trump support, and political violence, while the strongest negative predictors were support for Biden and trust in government. As reviewed in Linden et al. (2021), people holding conservative (vs liberal) views are more likely to mistake political opinions for facts, are more vulnerable to fake news and pseudo-profound bullshit, and are less interested in scientific knowledge. A survey of more than five thousand participants from 24 countries showed that politically conservative respondents were more likely to hold anti-vaccination attitudes, which is positively related to belief in CTs (Hornsey & Pearson, 2022). Higher number of conservatives are skeptic towards childhood vaccination (56% of Donald Trump supporter vs 11% of Bernie Sanders, Hoffman et al., 2019). On April 2021, Republicans were five times more likely, compared to Democrats, to refute COVID vaccinations (Piltch-Loeb et al., 2021).

Such a right- (vs left-) skewed relationship between political attitude and belief in CTs could be explained by two factors at play: cognitive characteristics and worldview. Conservatives and liberals have different cognitive styles that are rooted on genetic basis, emerge early in life, and result in brain structural differences (Bavel & Pereira, 2018). There is in fact a *cognition of the right* (Zmigrod et al., 2018, 2021). People on the right have higher epistemic needs to attain order, certainty, structure, closure, and existential needs to attain safety, security, and a sense of reassurance (Jost et al., 2003a, 2003b). These are also known predictors for belief in CTs (Douglas et al., 2017, 2019; Douglas & Sutton, 2023). Cognitive biases such as confirmation bias (W. Hart et al., 2009; Nickerson, 1998; Stroud, 2014) and motivated reasoning (Kunda, 1990) seem to be at the basis of such relationship.

Lewandowsky & Oberauer (2021) show that worldviews play an important role in endorsing CTs. Specifically, people on the conservative political spectrum tend to endorse CTs that reject scientific evidence. Science is based on communism (scientific research should be the common property of the scientific community) and universalism (knowledge should transcend racial, class, national, or political barriers) and is in contrast with conservative values (based on individualism, nationalism, and patriotism). Therefore, science-target CTs such as those directed to climate change and health-related issues (e.g., vaccination) tend to be endorsed by people holding conservative views (Lewandowsky & Oberauer, 2021; Lu et al., 2021; van Prooijen et al., 2015). But how about CTs on the liberal side? One might speculate that the extreme leftists believe in CTs about capitalism (van Prooijen et al., 2015), about the environmental contribution of intelligence (Lewandowsky & Oberauer, 2021), about the claims that President George W. Bush possessed advance knowledge of the terrorist attacks of 9/11 and chose not to intervene (Linden et al., 2021), or about Big Pharmaceutical Companies that pushed childhood vaccinations (Linden et al., 2021). There is little evidence for the Left supporting these theories (Lewandowsky & Oberauer, 2021; Linden et al., 2021).

1.2.3.3 *Power asymmetry*

CTs find a fertile soil in power asymmetries, that is, when political parties are deprived of political control. In the US context, Democrats tend to believe that Republicans are committing electoral fraud, and Republicans accuse Democrats of the same wrongdoing (Edelson et al., 2017; Wang & van Prooijen, 2022). Uscinski et al. (2011) found that when Democrats were in power, CTs accusing Democrats of wrongdoing were more prevalent and a similar figure emerged when Republicans were in power (high prevalence of CTs about Republicans). Similarly, across 26 countries, Imhoff, Zimmer, et al. (2022) found stronger conspiracy mentality among voters of opposition parties, supporting the notion that CTs *are for losers* (Uscinski et al., 2011). CTs are in fact powerful tools for vilifying opponent. CTs can be transmitted strategically for persuasive purposes, such as when they are intentionally shared on social media to affect outcomes like voter behavior (Bangerter et al., 2020; Douglas et al., 2019) and for propaganda and lure recruits by extremist groups (Bartlett & Miller, 2010; Basit, 2021).

1.2.4 Psychopathology

A multitude of studies view conspiracy beliefs as a symptom of an underlying psychological disorder such as paranoia (Bruder et al., 2013; Darwin et al., 2011; Freeman et al., 2022; Furnham & Grover, 2022; Kuhn et al., 2021; Larsen et al., 2021), schizotypy (D. Barron et al., 2014, 2018; Darwin et al., 2011; Denovan et al., 2020; Dyrendal et al., 2021; Furnham & Grover, 2022; Georgiou et al., 2019; J. Hart & Graether, 2018), proneness to delusion (Georgiou et al., 2019; Larsen et al., 2021), psychoticism (Bowes et al., 2021; Teličák & Halama, 2021), and dissociative tendencies (Pisl et al., 2021).

Paranoid ideation and schizotypy share similar traits, including suspicion, magical thinking, and endorsement of odd and unusual beliefs (Barlow et al., 2017). In paranoid ideation, people believe that others have hostile intentions toward them, involving physical or verbal threats, deception, exploitation, and disloyalty (Darwin et al., 2011; Freeman et al., 2005). Paranoia, however, differs from belief in CTs in one important respect: while people with paranoia find a suspected hostility against the self, people believing CTs usually perceive hostility towards their (in)group, assuming that a powerful and hostile outgroups conspire against them (van Prooijen & Douglas, 2018).

Schizotypy is particularly relevant to our work. Schizotypy is the milder and non-clinically significant form of schizophrenia. The two overlap substantially and are characterized, yet on different levels, by an impairment in thought and perception that lead to psychotic symptoms (Ettinger et al., 2014). Individual with schizotypal personality tend to jump to conclusions quicker and make decisions based on less evidence compared to people without schizotypy (Juárez-Ramos et al., 2014) and are vulnerable to misinformation (Bronstein et al., 2019).

1.3 The content of conspiracy theories

Little is known so far about the linguistic content of CTs. *What* do CTs talk about and *how* they do it? Do they have specific language style? Can we distinguish conspiracy narratives from non-conspiracy narratives? How much do they converge (and diverge) from intuitively similar phenomena such as rumors, urban legends, and fake news? Are there any specific linguistic markers (e.g., lexical, syntactical, semantic) that allow to identify CTs?

These questions pertain to the study of the *content* of CTs. This field of investigation allows to advance research on CTs in both theoretical and practical terms. Knowing the content of CTs allows to first define what is a CT and how it diverges from other types of similar narratives. Once CTs are distinguished from other similar types of narratives, then they can be identified and tracked. Algorithms for automatic detection of CTs can be developed allowing to study their endorsement, spread, and potential deleterious effects with more precision. Therefore, the study of the content of CTs, although overlooked so far, would theoretically and practically contribute to the ultimate goal of the whole research on CTs agenda, that is, fighting their spread.

The content of CTs has been investigated in a handful of studies, mostly focused on user-generated texts such as comments and posts gathered from social media e.g., from Twitter (Fong et al., 2021; Mitra et al., 2021; Wood, 2018), Facebook (Bessi, Zollo, et al., 2015; Bessi, 2016; Brugnoli et al., 2019; Smith & Graham, 2019), Reddit (Klein et al., 2018, 2019; Samory & Mitra, 2018a, 2018b), Gab (Zannettou et al., 2018), or from comment sections of news websites (Wood & Douglas, 2013, 2015). From this text material, researchers have performed analyses relying on lexical/sentiment analyses (Del Vicario, Vivaldo, et al., 2016; Faasse et al., 2016; Fong et al., 2021; Klein et al., 2019; Samory & Mitra, 2018a; Wood & Douglas, 2015) to explore *how* CTs use language, on topic distributions (Bessi, Zollo, et al., 2015; Klein et al., 2018; Mitra et al., 2021; Samory & Mitra, 2018b) to identify *what* CTs talk about, and on narrative patterns (Samory & Mitra, 2018b; Tangherlini et al., 2020) to identify the interplay of *whats* and *hows*.

Overall, studies converge in showing that users who believe in CTs use a language loaded of negative emotions with a focus on topics such as deception, death, religion, and power (Fong et al., 2021; see e.g., Klein et al., 2019). In a case-study, Tangherlini et al. (2020) described and contrasted the narrative structures of the *Pizzagate*[7] CT with an actual conspiracy, the *Bridgegate*.[8] The authors show that, compared to the actual conspiracy, the CT is initially composed of a relatively small number of actors, but, with time, other theories conflate rendering the network based on multiple domains. In another work, Samory & Mitra (2018b) identified the key narrative elements of CTs by identifying *agents*, *performance*, and *targets* (by extracting Subject, Verb, and Object) from (the parsed texts of) comments and posts on the social media Reddit. In doing so, the authors were capable of identifying common narrative motif such as *"governmental agency controls communications"* or *"science uncovers health threat"*. This study provided insights into both *what* and *how* users online use language when referring to CTs.

Although an effort has been made to understand the content of CTs, the main barrier to understand the language of CTs, in previous studies, is represented by the material researchers used in these studies, namely user-generated texts on social media. Focusing on social media has the advantage of exploring

[7]A theory circulating in 2016 claiming that Hillary Clinton and Democratic elites were running child sex trafficking

[8]A political scandal involving appointees of New Jersey Governor Chris Christie colluding to intentionally cause a massive traffic problem at the George Washington Bridge for political reasons

ecologically valid samples of text as a complement to psychological investigations of CT beliefs. Specifically, analysis of these texts provides information about what and how users endorsing CTs talk about CTs. Note, however, that it is difficult to reliably extract measures of individual belief from comments embedded within the noisy and heterogeneous discussion threads of conspiracy believers and debunkers (Wood & Douglas, 2013, 2015), but see how Klein et al. (2019) carefully matched CT-believers and non-believers.

The main limitation in studying the content of CTs via user-generated texts on social media is that discussion threads are not conspiracy narratives per se. While comments and posts on social media might instill curiosity in other users, reinforce existing beliefs, or support conversion (see e.g., Franks et al., 2017; Olshansky et al., 2020), they often do not constitute the actual source through which CTs are transmitted. Discussion threads, practically speaking, limit the utility of extracted text to perform text analysis because comments and posts are brief and are contextualized in the discussion in which they are embedded. It also follows that comments are incapable of spreading independently from the whole thread hence of low utility for studying the interplay of content and spread of CTs.

1.4 Why the spread of conspiracy theories matters

CTs propagate on social networks rapidly in a way that parallels infectious diseases (Kauk et al., 2021). People might reach CTs inadvertently (Sutton & Douglas, 2022), via a link posted on a social network or via a Google search about a particular issue (e.g., COVID-19). Once a conspiratorial worldview has started to build up, it can grow, accelerating in a non-linear fashion due to recursive dynamics (Sutton & Douglas, 2022). CTs are persuasive because, accumulating evidence (see e.g., Oswald, 2016), they give an impression of an overall truth (Pennycook et al., 2018) and because, freed from veridicality, they can exploit cognitive biases for appealing information (Acerbi, 2019; Hills, 2019; van Prooijen et al., 2022). It follows that exposure to CTs reinforces conspiracy belief which in turn motivate people to search for more CTs and join conspiracy groups online (Uscinski, Enders, Klofstad, & Stoler, 2022). Prolonged exposition to CTs, instead of satisfying epistemic and existential needs (Douglas et al., 2017), increases frustration (Liekefett et al., 2021) and the likelihood to search for more CTs.

1.4.1 They are popular

CTs are widely endorsed. Studies in the 2000s show that around half of surveyed populations believe in at least one CT surrounding the assassination of president Kennedy (Enders et al., 2018), pharmaceutical industry (Oliver & Wood, 2014), and the 9/11 terrorist attack (Allen, 2008). In 2004, 49% of New York City residents believed the US government to be complicit in the 9/11 terrorist attacks (Sunstein & Vermeule, 2009). At the beginning of the COVID-19 pandemic, in April 2020, about 85% of about 3,000 respondents believed in more than one CT about COVID-19 and 60% believing in three or more (Miller, 2020). Although these numbers seem not to have changed in years (Uscinski, Enders, Klofstad, Seelig, et al., 2022), they are alarming because CTs spread and cause detrimental consequences.

1.4.2 They spread

1.4.2.1 *Content affecting spread*

False information exploits cognitive biases for appealing information (Acerbi, 2019; Hills, 2019; van Prooijen et al., 2022) and is more likely to become viral, eliciting responses of surprise (Friggeri et al., 2014; Vosoughi et al., 2018). Such a persuasive advantage might stem from specific content features that make narratives more believable (Oswald, 2016; Vosoughi et al., 2017; Vosoughi et al., 2018). For example, emotional content is a successful feature of narrative stickiness and transmission (Franks et al., 2013; Heath et al., 2001). In the domain of urban legends, Heath et al. (2001) showed that stories rated as more disgusting, interesting, and eliciting surprise were also rated more likely to be transmitted. Similarly, but on Twitter, false news that elicited surprise and disgust were more re-twitted (Vosoughi et al., 2018).

Plausibility is another narrative feature that impacts positively transmission (Heath et al., 2001). In the study of Heath et al. (2001), rumors that were rated as more plausible were more likely associated with intention of transmission. In the context of CTs, this reminds the notion of *minimal counterintuitiveness* developed in the context of cognitive science of religion (Boyer, 2001; Norenzayan et al., 2006). Like religious narratives, conspiracy narratives combine intuitive with counterintuitive concepts (Franks et al., 2013).

Actors in CTs are represented as humans, but with extraordinary omniscience and omnipotence: they exert their power on the whole population while keeping it in secret (Bangerter et al., 2020). These features are likely to make CTs more memorable and more easily transmitted (Norenzayan et al., 2006).

CTs are based on a rhetoric of persuasiveness that emulates formal features of academic discourse (Oswald, 2016). On Twitter, higher levels of linguistic sophistication have been recorded in malicious (vs non-malicious) rumors, perhaps to make the rumors look more legitimate and believable (Vosoughi et al., 2017). Some scholars have described the rhetoric of CTs as characterized by an accumulation of "evidence" (or *mille-feuille* argumentative style, or *Gish gallop*) that overwhelms the reader by presenting a number of arguments with no attention to their accuracy or strength (Břízová et al., 2018; Goodwin, 2019; Oswald, 2016; Wagner-Egger et al., 2019).

Rumors and false news that appear novel, are more likely to be transmitted (Brooks et al., 2013; Vosoughi et al., 2018). Novelty could be also a predictive factor of CTs spread. For example, narratives that tap into the believers' need for uniqueness (see Section 1.2.2.3) might be more likely transmitted. Believers might spread CTs to enhance personal and social identity (Sternisko et al., 2020) by feeling special (Tian et al., 2001). For example, people find CTs interesting, exciting, and attention-grabbing, especially those high in sensation-seeking (van Prooijen et al., 2022).

1.4.2.2 *Contextual factors*

As seen in Section 1.2.2.2, CTs are particularly appealing for providing causal explanation (hence satisfying epistemic needs, see Section 1.2.2.1) during stressing societal events that elicit existential threat (Pipes, 1999; van Prooijen & Douglas, 2017). Time of crisis accentuate the salience of existential and epistemic needs: people need to be in control of —and know about— their environment. It is not surprising, then, that at the beginning of the first wave of the COVID-19 pandemic, in 2020, health-related misinformation attracted four times as much traffic as official health sources on social media (AVAAZ, 2020). COVID-19 misinformation not only attracted interest in CTs from a large number of researchers, but also provided new challenges and stimulated collaborative efforts to fight the spread of misinformation (Briand et al., 2021; Zarocostas, 2020): "*We're not just fighting an epidemic; we're fighting an infodemic*" (Zarocostas, 2020, p. 676).

1.4.2.3 *Individual differences*

Scholars have suggested that cognitive biases, such as confirmation bias, might be responsible for shaping individuals' news environment hence providing affordances for the spread of misinformation and CTs (Hills, 2019). Confirmation bias is a defensive behavior enacted to preserve an ideology that is crucial for making sense of the world (Koltko-Rivera, 2004) and to reduce the unpleasant state of conflict between old and new information (Festinger, 1957; W. Hart et al., 2009). Two primary mechanisms drive confirmation bias: selective exposure to congenial information and avoidance of challenging information (W. Hart et al., 2009; Nickerson, 1998; Stroud, 2014).

Confirmation bias is ideologically transversal but is accentuated in people highly committed to their beliefs (Knobloch-Westerwick & Meng, 2009; Taber & Lodge, 2006) and with a rigid thinking style (W. Hart et al., 2009). Confirmation bias increases as function of political (Pearson & Knobloch-Westerwick, 2019; Westerwick et al., 2017) and conspiratorial (Georgiou et al., 2021; Kuhn et al., 2021; Zollo et al., 2017) ideological strength. Studies show that CTs are used as tools to denigrate their opponents (Smallpage et al., 2017), and this is accentuated when supporters of political parties are not in power (Imhoff, Zimmer, et al., 2022; Uscinski et al., 2011). As seen in Section 1.2.3.3, CTs against political opponents proliferate when they are in power.

Another individual characteristic associated with spread of CTs is schizotypal personality, which is linked to vulnerability to misinformation (Bronstein et al., 2019). Individual with schizotypal personality tend to jump to conclusions quicker and make decisions based on less evidence compared to people without schizotypy (Juárez-Ramos et al., 2014). This might facilitate the spread of CTs via users with low reflective thinking who shares CTs on social media without rationally assess the content of the narrative.

Analytic thinking, consistently low in people believing in CTs (Swami et al., 2014; van Prooijen, 2017; Wagner-Egger et al., 2018), is a robust protector against sharing fake news and CTs (Mosleh et al., 2021). People with high analytic thinking are less likely to believe in fake news (Bago et al., 2020; Pennycook et al., 2020; Pennycook & Rand, 2019b), they report lower likelihood of sharing them on social media (Pennycook et al., 2020; Pennycook & Rand, 2019b), and they rate fake news or hyper-partisan news sources as unreliable (Pennycook & Rand, 2019a).

1.4.2.4 *The media*

The internet constitutes a system of information proliferation by which many people form opinions in regard to political parties, social issues, and health-related information (Betsch et al., 2011). Some scholars claim that on the internet, misinformation spreads faster, farther, and deeper within echo chambers, i.e., groups of like-minded individuals (Del Vicario, Bessi, et al., 2016; Del Vicario, Vivaldo, et al., 2016; Vosoughi et al., 2018), although others suggest that the internet did not facilitate the spread of CTs (Clarke, 2007; Uscinski et al., 2018; Uscinski, Enders, Klofstad, Seelig, et al., 2022).

In the Web 2.0 version of the internet, information is produced and consumed in a horizontal fashion, allowing anyone to create and share content, with few editorial filters (Aupers, 2012; Bessi, Coletto, et al., 2015; Bessi, Zollo, et al., 2015). CT texts may thus have the same epistemological weight for many users as mainstream texts do, and compete with them for attention (Bessi et al., 2014; Eicher & Bangerter, 2015; Hills, 2019). The perceived credibility of epistemic sources is also a function of belief in CTs (Imhoff et al., 2018). This makes epistemic authority difficult to evaluate, especially when conspiratorial narratives are promoted by political leaders (e.g., Donald Trump Barkun, 2017), by scholars in prestigious journals (e.g., the infamous case of autism-vaccine link, see Wakefield et al., 1998), and by Nobel Prize winners in predatory journals, see e.g., Luc Montagnier publishing in the *"International Journal of Research -GRANTHAALAYAH"* (Perez & Montagnier, 2020).

Some scholars propose that the spread of CTs and misinformation is facilitated by the structure and algorithmic ranking (Pariser, 2011) of online social networks (e.g., Facebook, YouTube) that promote the formation of echo chambers (Cinelli et al., 2021) and foster radicalization towards extreme ideologies (Bessi, Coletto, et al., 2015; Mocanu et al., 2015; Zollo et al., 2017). Despite overwhelming evidence for echo chambers on social media (Bessi, 2016; Brugnoli et al., 2019; Cinelli et al., 2021; Del Vicario, Vivaldo, et al., 2016; Raemdonck, 2019), some scholars suggest that echo chambers do not represent an urgent concern because information diet, at least in the US, is mostly centrist and diverse (Guess, Nyhan, Lyons, et al., 2018), yet *"the potential for a balkanized future remains"* (Guess, Nyhan, Lyons, et al., 2018, p. 16) and misinformation continues to spread (Lasser et al., 2022).

Not all social media are equal in terms of spreading CTs. Some social media, more than others, are particularly suited for allowing people to discuss and spread extremist content and CTs. Among those investigated by scholars there are 4chan.org (Tuters et al., 2018), Gab (Cinelli et al., 2021; Mathew et al., 2019; Zannettou et al., 2018)], Stormer (Törnberg & Törnberg, 2022), Parler (Aliapoulios et al., 2021), and Truth Social, the social media owned by the Trump Media & Technology Group launched on February 21, 2022. CTs spread also via messaging platforms such as Telegram (Hoseini et al., 2021) and Whatsapp (Theocharis et al., 2021).

There are also differences among the most popular social media. In one study, Theocharis et al. (2021) tested which social media use predicted belief in CTs finding that Facebook, YouTube, and Whatsapp were positively linked to the spread of CTs while the use of Twitter was negatively related to CTs. In another study, Cinelli et al. (2021) found that platforms organized around social networks and news feed algorithms based on filter bubbles —the algorithmic ranking on which social media are based (Pariser, 2011)— such as Facebook and Twitter promote the emergence of echo chambers. The authors also found higher segregation on Facebook than Reddit. Paralleling these findings, Raemdonck (2019) suggests that Facebook, compared to Reddit, facilitates the process of echo chambers formation allowing individuals to be selective in their beliefs and groups of peers, facilitating proselytism and radicalization. Echo chambers appear mainly on platforms where people choose to join groups of like-minded peers avoiding dissenting ideas. Moreover, Facebook, compared to Reddit, fosters the emergence of echo chambers because it relies on "cookies" that provide a more personal experience (e.g., by sorting posts via algorithms that factor in a number of different values, including friend relationships, explicit user interests and prior user engagement, DeVito, 2017).

Overall, it seems that the algorithmic ranking on which social media are based, i.e., filter bubbles (Pariser, 2011), are particularly suited to satisfy people's need to confirm their beliefs (see also section 1.4.2.3). The YouTube recommendation algorithm, for example, works by proposing conspiratorial videos if users watch conspiratorial videos (Faddoul et al., 2020; Hussein et al., 2020) hence amplifying and facilitating access to CTs (Alfano et al., 2021).

1.4.2.5 *The right*

As already discussed in Section 1.2.3.2, the relationship between belief in CTs and extreme political spectrum is more leaned towards the right side of the political spectrum. Such a skew is evident also in terms of CTs transmission. Republicans are far more likely to share misinformation than Democrats (Guess et al., 2019; Lawson & Kakkar, 2021; Osmundsen et al., 2021) and are more resistant to debunking (Rathje et al., 2022). Donald Trump is a fervent conspiracy disseminator (Barkun, 2017). Between 2011 and 2015, he spread at least 120 tweets on global warming (Matthews, 2017), while he also seems to believe that vaccinations cause autism and that climate change is a hoax (Lewandowsky & Oberauer, 2021). During the 2016 US presidential campaign, supporters of Trump (vs supporters of Hillary Clinton) were more likely to visit fake news websites, which were overwhelmingly pro-Trump (Guess, Nyhan, & Reifler, 2018). An analysis of seven online extremist forums revealed that spreading extremist ideologies about COVID-19 on violent right-wing extremist and incel[9] online forums increased significantly following the declaration of the pandemic, while the same was not true of left-wing or jihadist forums (G. Davies et al., 2021).

1.4.3 They have consequences

The high number of people subscribing to CTs (see Section 1.4.1) is alarming because CTs are capable of mobilizing people (Marie & Petersen, 2022) causing detrimental societal effects (Pummerer, 2022). CT belief is related to general distrust and political alienation along with endorsement of non-normative (vs. normative) political intentions (Einstein & Glick, 2015; Imhoff et al., 2021; Jolley & Douglas, 2014a). Exposure to CTs reduces pro-social attitudes (Linden, 2015) and trust towards institutions (Einstein & Glick, 2015) and can lead to the disruption of the social order, as the Capitol Hill storm on January 6$^{\text{th}}$, 2021, has revealed (Kaplan, 2021). CTs provide justification for engaging in everyday crime and violence (Jolley et al., 2019; Jolley & Paterson, 2020) and are associated with anti-Semitic and Islamophobic attitudes (Golec de Zavala & Cichocka, 2012; Swami et al., 2018). As such, they are often used by extremist groups for (online) propaganda and lure recruits (Bartlett & Miller, 2010; Basit, 2021), in some instances culminating in cases of murdering and criminal incidents perpetuated by hand of individuals who believe in CTs (Allington & Joshi, 2020; Bartlett & Miller, 2010; Basit, 2021; Dave et al., 2021; Garry et al., 2021).

[9] A community of men who consider themselves unable to attract women, typically associated with misogyny

CTs represent an obstacle to the adoption of evidence-based policies such as in medical and environmental sciences (Ball, 2020; Linden, 2015; Raab, Auer, et al., 2013; Raab, Ortlieb, et al., 2013). For example, belief in climate CTs is linked to lower support for pro-climate policies and reduced pro-environmental intentions (Biddlestone, Azevedo, et al., 2022). In medical settings, the acceptance of anti-scientific claims such as CTs can drive patients to reject mainstream medicine in favor of alternative medicine (Cassileth, 1984) or refute vaccines (Jolley & Douglas, 2014b; Lazarus et al., 2020; Salmon et al., 2005). At the population level, endorsement of health-related CTs causes loss of human lives (Gangarosa et al., 1998), and public funds (P. Davies, 2002). In Canada, anti-vaccine misinformation caused an estimated loss of —at least— $300,000,000 and nearly 3,000 unnecessary deaths (Council of Canadian Academies, 2023). HIV-related CTs with decreased condom use and lower antiretroviral medication adherence (Bogart et al., 2010; Bogart & Bird, 2003). Endorsement of COVID-19 CTs resulted in resistance to COVID-19 containment measures and vaccination (Biddlestone et al., 2020; Bierwiaczonek et al., 2020; Lazarus et al., 2020; Pummerer et al., 2022),

2 This thesis

In the previous chapter, I have introduced the topic of CTs. I defined the boundaries of CT research and described the conspiracy mentality (Section 1.1). It followed a brief literature overview of the psychological and social precursors of conspiracy mentality and belief in CTs (Section 1.2). In two sections, then, I described the state of the art of research on the content (section 1.3) and spread (Section 1.4) of CTs. Now, in the following section, I present our work, starting with a brief overview of the four studies that comprise this thesis (section 2.1), followed by a brief methodological note (Section 2.2).

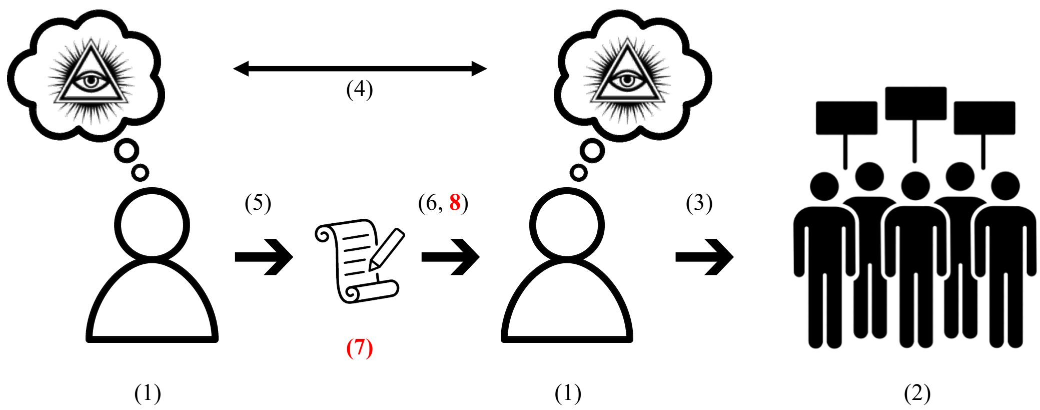
In Figure 2.1, I summarize the main fields of research on CTs described in the previous chapter, providing also a graphical overview and contextualization of my thesis (highlighted in red).

1. Individual predispositions and contextual factors that explain vulnerability to CTs (see Section 1.2)
2. Consequences associated with endorsing CTs (see Section 1.4.3)
3. How beliefs affect behavior (see Section 1.4.3)
4. How like-minded conspiracy believers form groups and influence each other, e.g., via social media, in echo chambers (see Section 1.4.2.4)
5. Factors that affect the spread of CTs (see Section 1.4.2)
6. How exposure to CTs affects belief (see Section 1.4.2)
7. The content of CTs (see Section 1.3)
8. How CTs spread or are reached (see Section 1.4.2)

2.1 Overview of the thesis

The goal of this thesis is to provide a better understanding of the content and spread of CTs. To this goal, four manuscripts, which are attached in the Appendix (see Section 5), have been produced.

Figure 2.1: Research fields in conspiracy theories



1. In Section 2.1.1, I describe LOCO, the corpus of CTs that, due to its rich set of data and metadata, allows to explore both the content and spread of CTs.
2. In Section 2.1.2, I describe our work on conspiracy interconnectedness. We used LOCO to perform a large-scale text analysis testing whether conspiracy narratives form a hyper-connected network of different topics. We found that conspiracy narratives show in fact strong patterns of interconnectedness, texts are characterized by low cohesion, and are, overall, more similar to each other compared to non-conspiracy narratives.
3. In Section 2.1.3, I describe our work on overinclusive thinking. We extracted from LOCO's texts measures of creativity that have been previously associated with divergent and convergent thinking. We found an imbalance of the two, which suggests the presence of overinclusive thinking in CTs narratives.
4. In Section 2.1.4, I describe our work on selective exposure, that is a specific way of accessing information (in our case, online). In this study, we analyzed the incoming traffic towards websites leaned on different ideologies. We found that users of conspiracy and political websites search for information that is consistent with confirmation bias.

2.1.1 Study 1: A corpus of conspiracy theories

Study 1 stemmed from the need to study the language of CTs. In 2019, when we started searching for literature and material, we realized that there were no resources available to study the language of CTs. The material available, at that time, was mostly focused on fake news and rumors (see e.g., Castelo et al., 2019; Kwon et al., 2017; Zubiaga et al., 2016). In Section 1.3, I have described the limitations of investigating the content of CTs via user-generated texts on social media, namely that posts and thread on social media are not CTs per se. A better way to address this limitation is studying CTs as text from webpages. Compared to texts in posts and tweets on social media, webpages provide space for an in-depth and elaborated discourse. Conspiracy websites, specifically, are specialized sources created for the purpose of developing, collecting, and spreading CTs. These websites provide space to discredit official narratives and are trustworthy epistemic sources for CT believers. In addition, webpages constitute standalone, structured texts nested within sources (i.e., websites) and as such provide paratextual information that is important to explore features associated with text content. Finally, as standalone documents, webpages can easily be shared on social media and so can provide measures of spread. Appropriately identified webpages, therefore, would be beneficial for studying the content of CTs and could provide a solid grounding for CT research.

We surveyed the literature assessing weaknesses and strengths of corpora previously built for rumors and fake news, and so we built LOCO (the language of conspiracy corpus, see Section 5.1). LOCO is a multilevel topic-matched corpus composed of standalone documents (N = 96,743) gathered via ready-made lists of conspiracy and non-conspiracy websites (N = 150 in total). LOCO has been built as a freely available text source from which researchers can extract features and/or generate predictive and classification models. The strength of LOCO is that it allows to study both the content and spread of CTs and allows to make comparisons with a topic-matched corpus of non-conspiracy documents. Each document in LOCO is associated with both fine- and coarse-grained semantic indexing (topics and keywords) as well as a rich set of lexical features (e.g., similar to sentiment analysis). Each document is provided with a measure of spread, that is the number of times it was shared on Facebook.

2.1.2 Study 2: Interconnectedness

The conspiracy mentality (see Section 1.1) is a self-sustaining worldview comprised of a network of mutually supportive beliefs (Wood et al., 2012). In Study 2 (see section 5.2), we test this proposition, performing large-scale text and network analyses using texts from LOCO. Results provide strong empirical support for an overarching conspiracy worldview in conspiracy narratives. We show that CTs exhibit a pattern of strong interconnectedness, linking multiple ideas that result in a dense and hyper-interconnected network. Individual CT documents are built from multiple sources and are, on average, less locally (within-document) cohesive than corresponding non-conspiracy documents. They nevertheless exhibit higher global (between-document) coherence, being more lexically similar to each other than non-conspiratorial documents. Noteworthy, hyper-interconnectedness and low textual cohesion suggest the presence of altered semantic processing as in the schizophrenia spectrum (see section 1.2.4), suggesting clinical significance.

2.1.3 Study 3: Overinclusiveness

Study 2 paved the way for Study 3. Evidence for hyper-interconnectedness and low textual cohesion in conspiracy narratives resembles thought alterations as in individuals on the schizophrenia spectrum, which includes the subclinical and milder form schizotypy (that correlates with belief in CTs; see section 1.2.4). Schizotypy and schizophrenia are characterized by an impairment in thought and perception that lead to psychotic symptoms (Ettinger et al., 2014), which can be manifested in language production (Elvevåg et al., 2007; Rezaii et al., 2019). This indicates a certain degree of overlap between the schizophrenia spectrum and belief in CTs in regard to semantic processing, suggesting that manifestation of conspiracy mentality would emerge in texts.

Psychoticism and the schizophrenia spectrum are related to overinclusive thinking (F. Barron, 1993; Eysenck, 2003; Runco, 2010). which is often the product of *"nonconformity, lack of discipline, and blind rejection of what already exists* [...] *regardless of accuracy"* (Cropley, 2006, p. 392). In study 3 (see section 5.3), we explored whether conspiracy texts show elements of overinclusive thinking. We performed a large-scale text analysis using LOCO and developed measures to assess elements of convergent and divergent thinking. Results point to an imbalance between convergent and divergent thinking, suggesting the presence of overinclusive thinking in the process of generating nominal compounds (e.g., *carbon dioxide*) in conspiracy narratives. Overinclusion of ideas without any filter to select pertinent ideas might be at the basis of low cohesion in conspiratorial narratives. Overinclusive thinking could then help explain the accumulation and interconnectedness of CTs (see section 2.1.2).

2.1.4 Study 4: Accessing conspiracy theories

Studies converge in showing that both social media (Bessi, Coletto, et al., 2015; Cinelli et al., 2021; Lasser et al., 2022; Mocanu et al., 2015) and cognitive biases (Georgiou et al., 2021; Guess et al., 2019, 2020; Kuhn et al., 2021; Pearson & Knobloch-Westerwick, 2019; Westerwick et al., 2017; Zollo et al., 2017) affect information foraging in people committed to strong ideologies such as CTs and political extremes. However, due to heterogeneity of methods and a lack of comprehensive and comparative data, the magnitude of their respective contribution cannot be conclusively drawn.

In LOCO, we have assessed the impact of both social media and cognitive biases on online traffic towards CTs and politically biased websites, showing that both factors drive people to search for information in highly ideological websites. However, also LOCO's analyses presented some weaknesses such the limited sample size of websites used for analyses (N = 91), the broad categorization of website categories (conspiracy, biased, least-biased, and pro-science websites), and the lack of important confound variables.

In study 4 (see section 5.4), we devised a field study to assess and compare the magnitude of the impact of social media and cognitive biases in online information foraging. We replicated LOCO's analyses with a more robust methodology. Results show that cognitive biases are the main drive for accessing conspiracy and political websites and increase as a function of ideological strength. This behavior is particularly marked in conspiracy websites followed by conservative websites and reduced in liberal websites. Conspiracy and conservative websites' users show similar patterns of information foraging.

2.2 Methods

The goal of this thesis is to provide a better understanding of the content and spread of CTs. Across four studies, we formulated questions within the domain of psychology and developed/borrowed methods from computational sciences to answer these questions. In the following, I briefly describe the main methodological implementations that characterize our work.

2.2.1 Automated corpus construction

Traditionally, psychology studies rely on a limited sample of participants (see for example section 1.2), but see collaborative mega studies such as Imhoff, Zimmer, et al. (2022). Even when text analysis is performed, sample sizes remain modest. For example, Fu et al. (2016) used 116 unique webpages, Xu & Guo (2018) analyzed headlines from 615 pro-vaccine and 442 anti-vaccine webpages, and Xu et al. (2019) analyzed 541 pro-vaccine and 382 anti-vaccine webpages. One study performed content analysis on about 100,000 published letters to the editor of The New York Times from 1897 to 2010 (Uscinski et al., 2011). After the data were collected, each document was manually coded as either referring to a conspiracy or not. The authors identified 800 letters (about 0.8% of the corpus) as referring to conspiracies, which is still a rather modest —and unbalanced— sample size.

As reviewed in Section 1.3, the majority of studies aimed at understanding the content of CTs was based on user-generated texts on social media. While working on social media allows to obtain generous sample sizes (see e.g., 1.4M posts from the two online Reddit communities of r/conspiracy and r/science used by Zhang et al., 2021), these works present some limitations, above all short texts (e.g., Twitter's 140- and 280-character limit before and after 2017). Another limitation present in previous works —except a notable exception, see Zubiaga et al. (2016)— was that documents were not matched by topics, which could hinder the extraction of group-specific language (in our case conspiracy).

In our work, we overcome the limitation of sample size by relying on text extraction in an automated way. In LOCO (see section 5.1), for example, we have automated Google searches by creating queries to search keywords within websites, e.g., "`site:bbc.com moon landing`".[1] To our knowledge, this is the first time a corpus was built in such a way. This method allows to build a corpus specifying sources and content while providing a generous sample size.

The possibility to specify sources and content upon which building a corpus represents a crucial innovation. If correctly done, this allows to make comparisons between and within sources and content (i.e., topics). For example, in our case, it is possible to analyze how the language of conspiracy on documents matched by topic, e.g., performing analyses between sources (conspiracy vs non-conspiracy) and within topic (e.g., 9/11). Also, topics can

[1] https://www.google.co.uk/search?q=site%3Abbc.com+moon+landing&hl=en

be aggregated: if one aims at studying diseases, topics such as Ebola, Zika, COVID-19 can be aggregated. The same applies to specific diseases if one is interested in studying differences between the outbreak of coronavirus in China vs in the US.

2.2.2 Representativeness

Because one might be interested in what is the prototypical conspiratorial language, we aimed at extracting a set of the most representative CT documents in LOCO. A representative document should represent as many as possible documents within the conspiracy world and share more words with the conspiracy corpus compared to a less representative document. One way to do this is by extracting the most frequently occurring words in conspiracy documents and compute a measure of between-document similarity. In doing so, we are capable to retrieve documents containing the prototypical language. Recurrent word patterns such as *"they are trying to KILL US!"* or *"know the truth"* might in fact be highly shared across the majority of conspiracy documents. Documents containing these lexical patterns should display high similarity, in theory, with the whole conspiracy corpus, representing the conspiratorial language use.

We did so in LOCO, by computing the cosine similarity between words of each document against all words in the conspiracy corpus. In doing so, each document was associated with a vector of values ranging from 0 to 1 (indicating respectively either no overlap or a perfect overlap of words). By computing the average of each vector, we obtained a value of "how a document is similar to the whole conspiracy corpus". By extracting the documents with the highest values (in our case, one standard deviation above the mean), we obtained a set of documents that share the most words with the whole conspiracy corpus, i.e., cleaned by event-specific CTs (e.g., Lady Diana or 9/11). Below an excerpt from the document with highest cosine similarity value, i.e., the most representative document of conspiracy corpus (document ID: `C01b90`).

> *The reality of what's truly happening inside the United States is far, far more terrifying than just being screwed. Because quite literally, they are trying to KILL US! "Oh, come on now," I can hear you thinking. "Get real!" Well that's exactly what I'm will be doing in this article. I'm going to get very real.*

2.2.3 Network analyses

In Study 2, we combined natural language processing techniques with network analysis to test the assumption that conspiracy documents are more interconnected than non-conspiracy documents. Prior to our study, this conjecture was based on correlation matrices built from participants answering their level of agreement with specific CTs (see e.g., Goertzel, 1994; Wood et al., 2012). To answer this question (i.e., whether conspiracy narratives are more interconnected than non-conspiracy narratives), we have built networks from keywords and topics that describe each document in LOCO. We tested whether the there were differences in the number of edges per each node. Note that a node is either an event associated with CTs (e.g., the death of Lady Diana, 9/11) or a topic (extracted in an unsupervised fashion from the corpus using the latent Dirichlet allocation, LDA, Blei et al., 2003).

3 General discussion

The goal of this thesis was to investigate the content and spread of CTs. When I started, in 2019 prior to the pandemic, resources to investigate the content and spread of CTs were scarce. Therefore, we built the largest corpus of CTs available today: LOCO (Study 1). LOCO is a turnkey resource for researchers: it is freely available, and it provides a rich sets of linguistic features and metadata that link measures of spread with content features associated with texts. Relying on LOCO, we derived two main streams of research that focused on the content (see section 3.1) and spread (see section 3.2) of CTs.

3.1 The content

3.1.1 Language use

In Study 1, we explored the language use in conspiracy narratives. Similar to other works focused on text analysis (see section 1.3), conspiracy documents in LOCO were found to rely on a language loaded of negative emotions with a focus on deception, terrorism, death, religion, and power. We visually inspected the lexical fingerprinting of conspiracy documents with previous lexical analyses performed by other scholars on social media (Fong et al., 2021; Klein et al., 2019) and found presence of lexical overlap. This finding is interesting as it shows that talking (i.e., writing a post on a social media) about CTs relies on a language similar to that of conspiracy narratives. Is this a characteristic of CTs per se or a feature of believers (who use language in a certain way that is different from non-believers)? Future studies might assess in a more systematic way to what extent the language of people delivering CTs overlap with the language of conspiracy narratives embedded in webpages. This can be done via collaborative efforts by merging data sets (e.g., LOCO with data from social media) and testing lexical alignment. Alternatively, experimental studies could provide insight on the causality of such hypothesized alignment, e.g., does exposure to conspiratorial narratives trigger conspiratorial language?

While the linguistic alignment between conspiracy narratives from websites

(LOCO) and social media posts might suggest that perhaps believers use language differently from non believers, we found another case of alignment that suggests that similarities might be due to the texts themselves —instead of individual characteristics. In LOCO, we reported that non-conspiracy documents that mention either actual conspiracies or CTs (that is, documents that simply included the word `conspir*` in the text) tend to align, in regard to lexical features, to the language of conspiracy documents, as reflected by the focus on crime, terrorism, and deception (among others). While we did not know whether these documents refer to actual conspiracies (e.g., Watergate) or CTs (e.g., Moon landing is a hoax), we nevertheless recorded that simply mentioning a conspiracy has an effect on language use, which converges towards that of CTs. This phenomenon might pose challenges for future endeavors in developing algorithms for automatic detection of CTs. But it also raises questions on the boundaries of CTs, at least in regard to lexical features: what does distinguish a CT from an actual conspiracy in terms of lexical features?

As described in Section 2.2.2, we extracted the prototypical language of CTs, namely a subset of representative conspiracy documents that rely on lexical patterns shared across conspiracy documents, e.g., *"they are trying to KILL US!"*. These documents appear to be an exaggerated version of an average conspiracy document. When compared to other documents in the conspiracy corpus, the representative set is generally more emotionally charged and displays a prototypical language of conspiracy focused on power, dominance, and aggression. At the same time, the representative set of documents is less likely to rely on non-conspiratorial themes such as those related to tourism or vacation. Representative conspiracy documents, more than other conspiratorial documents, are also more likely to use a rhetorical style built upon refutational strategies based on questioning the dubious version of the official story while highlighting the lack of answers from official sources. This provides evidence for the theoretical work of Oswald (2016). Furthermore, in line with the social aspects of CTs (see section 1.2.2.3), we found that prototypical conspiracy narratives are based on process of social identification of the ingroup by exclusion from the outgroup.

In Study 3, we focused on how complex terms are generated and used. We found that conspiracy texts make use of more original, semantically divergent, and sophisticated compound words, but less appropriate to the context and less variable compared to those in non-conspiracy texts. It seems that while the *generation* of compounds shows elements of creativity, the *use* of compound words is rigid. Conspiracy narratives use fewer compounds, repeated over texts, revolving around fewer topics. In other words, sophisticated compound constructions might be borrowed from specialist texts (see e.g., Oswald, 2016) and/or, generated to elicit surprise (perhaps for persuasive purposes, see e.g., Vosoughi et al., 2017; Vosoughi et al., 2018). Once a congenial lexical solution (i.e., a surprisingly complex compound) is found, then it is repeated over across the document, spanning multiple topics. Future studies focusing on narrative production methods (see e.g., Raab, Ortlieb, et al., 2013) could help understand the poietic process in the production of conspiracy texts.

3.1.2 Conspiracy mentality

As revealed by our analyses in Study 2, we provided evidence that conspiracy narratives rely on abundant accumulations of proofs to showcase their conspiratorial arguments (see e.g., Břízová et al., 2018; Goodwin, 2019; Oswald, 2016; Wagner-Egger et al., 2019). Nevertheless, although built from several topics, conspiracy narratives show higher level of similarity, compared to non-conspiracy narratives. The set of representative conspiracy documents (see Study 1) make use of lexical patterns that are shared across the majority of conspiracy texts, e.g., *"they are trying to KILL US!"*. Moreover, as shown in Study 3, the use of language in conspiracy text is rigid, meaning that compound words (or lexical patterns) are repeated over the text regardless of their context. Overall, these findings point to a language use, in CTs, that is rigid and crystallized, suggesting that conspiracy narratives are generated in a systematic way, at least from a linguistic perspective.

I suggest this is a linguistic instance of *conspiracy mentality*. Relevant social events are translated into conspiracies by adding a set of recurrent lexical patterns that can be reused in any conspiratorial context and are shared across conspiracy narratives. These patterns are not specifically related to the event described, but seem to work like boilerplates.[1] Yet, these lexical patterns in CTs can be indeed the linguistic markers of CTs. The example 3.1, shows that any socially relevant event i (e.g., 9/11 or death of Lady Diana), becomes a CT by adding the lexical patterns *"They lied about"*, *"it is a hoax!"*.

[1] Boilerplates are reusable lexical patterns (e.g., openings in mails such as *"hope you're doing well"*, or navigational links in webpages) that are applied to texts without significant changes and without adding information to the main content of a text

$$\textit{They lied about } \{event_i\}\,,\ \textit{I knew it was a hoax}! \tag{3.1}$$

This linguistic behavior parallels a psychological one. More specifically, it resembles an accentuated case of top-down information processing, where inferences about social events are drawn from applying a conspiratorial worldview to those observations (Franks et al., 2017). Ultimately, this cognitive processing style might allow to coerce unconnected or even contradictory observations into support for a global conspiracy (Lewandowsky et al., 2018) and might explain the hyper interconnectedness we found in Study 2. In doing so, observations fit an overarching conspiracy worldview and are translated into a conspiracy by adding a set of recurrent lexical patterns involving language of deception, questioning, social identification, and negative emotions (see Study 1 and Section 3.1.1) that can be reused in any conspiratorial context and are shared across conspiracy narratives.

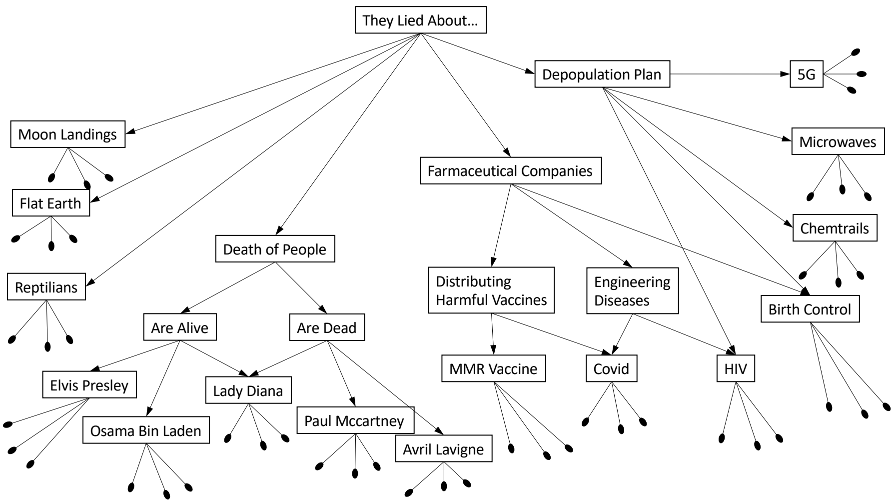
The set of representative conspiracy texts make use of lexical patterns that are shared across conspiracy narratives but, crucially, they are event-free, meaning that they do not refer to a specific CT (see e.g., items from scales that measure specific CTs in Section 1.1). Thinking conspiracy narratives as narratives that vary in degrees of specificity has theoretical and practical implications for advancing research on CTs.

At the theoretical level, correlations between single items and the patterns of strong interconnectedness between CTs (see Study 2) suggests that single specific events are clustered in more generic, or overarching, CTs on several layers. As such, narratives can be nested and classified based on their level of specificity, resembling the classification of scales measuring belief in CTs and conspiracy mentality (see Section 1.1). At the highest level of specificity, different conspiratorial explanations coexist for a single event. For example, Lady Diana is both dead (to prevent her from marrying an Arab or killed by Al-Fayeds' business enemies) and alive (she faked her death, see Wood et al., 2012). The death of Lady Diana is part of a group of CTs that involve socially relevant figures including Osama bin Laden and Elvis Presley. Although unrelated in mainstream worldviews, these figures (or conspiratorial facts, or events) are interconnected in the conspiratorial worldview by the fact that these people faked their death. Other living figures, instead, have been replaced by a look-alike (see e.g., Paul McCartney and Avril Lavigne). These theories form a cluster of CTs on their own. The two clusters of CTs surrounding dead and/or living social figures form themselves a cluster of CTs sharing the view that *"authorities hide information about the death/replacement of socially relevant people"*. The same logic applies to CTs claiming that pharmaceutical companies distribute harmful vaccines and are involved in disease engineering, forming a cluster of CTs revolving around wrongdoings of *Big Pharma*. At the most generic level, these clusters are grouped to an overarching view that *"they lied to us"*. At this level of abstraction, it is not even important to define *who* is lying to us, because *"the official version of the events given by the authorities very often hides the truth"* (Lantian et al., 2016).

In Figure 3.1, I show a tentative taxonomy of CTs based on specificity. Similar to phylogenetic trees, CTs are grouped based on shared features/themes. For example, at the tips of the branches (black dots), there are individual theories such as different explanations for the (faked) death of Lady Diana. The death of popular figures are grouped within a clade. Similarly, there is a clade of people who are in reality alive but according to CTs they are replaced by look-alike. These two clades are nested within one another involving popular figures. At the same level, there are CTs about pharmaceutical companies claiming that they distribute harmful vaccines or engineering diseases. The root of the whole tree represents the core aspect of CTs, that is *"they lied to us"* as seen in the example 3.1 (and from the single-item scale developed by Lantian et al., 2016).

Such a taxonomy has practical implications. Classifying conspiracy narratives on their degree of specificity might help developing algorithms for automatic detection of CTs. From specific theories surrounding an event (e.g., different accounts for the 9/11 terrorist attack) to a generalized distrust towards authorities that hide the truth (see Section 1.1), specificity decreases while generalizability increases. It follows that extracting patterns from less specific documents, namely patterns that are shared across all conspiracy narrative, would help identify also future CTs such as novel viruses, or new theories about old events. This represents a foreseeable possibility to develop an algorithm for detecting CTs.

Figure 3.1: A tentative taxonomy of conspiracy theories



3.2 The spread

3.2.1 Content affecting spread

Overall, we found evidence that conspiracy narratives exploit cognitive biases for appealing information (Acerbi, 2019; Hills, 2019; van Prooijen et al., 2022). In Study 1, we found that the set of representative conspiracy documents had higher shares on Facebook, meaning that those documents had an advantage in transmission. Being rhetorically appealing and emotionally loaded, these documents spread online more successfully than other conspiracy documents. This is in line with the fact that emotional content is a successful feature of narrative stickiness and transmission (Franks et al., 2013; Heath et al., 2001).

In Study 3, we found that conspiracy texts were more original, semantically divergent, and sophisticated compared to those in non-conspiracy texts. Studies show that false information is more likely to be transmitted online, eliciting responses of surprise (Vosoughi et al., 2017; Vosoughi et al., 2018). This transmission advantage might stem from content features such as novelty, semantic divergence, and sophistication that make narratives more attention-grabbing (Oswald, 2016; Vosoughi et al., 2017; Vosoughi et al., 2018). Originality and divergence generate surprise, which attracts attention (Dillard, 2001; Schomaker & Meeter, 2015) and encourages information transmission by conveying social status on people who believe have access to unique information (Berger & Milkman, 2012), tapping into need for uniqueness (Imhoff & Lamberty, 2017; C. S. Kay, 2021) and sensation-seeking (van Prooijen et al., 2022) that characterize conspiracy believers.

3.2.2 Cognitive biases

In Study 1, we explored the impact of both social media and cognitive biases on driving traffic towards CTs and politically biased websites. We found that both factors drive people to search for information in highly ideological websites. In Study 4, we replicated these analyses with a better design and with a larger and more fine-grained classified set of websites so to provide a comprehensive and comparative assessment of the traffic leading to ideological websites. We found that both social media and cognitive biases contribute to drive people towards these websites. More specifically, we found that people endorsing high ideologies access information in a way suggesting confirmation bias, which is a defensive mechanism that allows to preserve a worldview by seeking confirmation of own prior beliefs while avoiding challenges. The impact of cognitive biases is more evident for users accessing conspiracy websites followed by conservative and liberal websites. Our analyses revealed that users of CTs and conservative websites showed high similarities. This is consistent with the fact that the two worldviews share core values (see Section 1.2.3.2) and CTs widely spread on the right political spectrum (see Section 1.4.2.5). Conspiracy websites, however, differ from politically biased (both liberal and conservative) websites in one important aspect: the influence of social group decreases as websites' conspiratorial ideology increases.

Findings from our Study 4 have practical implications on fighting the spread of misinformation. Given that access to conspiracy websites is mostly driven by cognitive biases instead of influences from social groups, individual-level interventions should be prioritized over group-level interventions. For example, following the Capitol Hill storm in January 2021, the social media accounts of the ex-president Donald J. Trump were suspended and Twitter took down more than 70,000 accounts spreading misinformation (Tollefson, 2021). Yet, misinformation continued to spread (Lasser et al., 2022). Overall, social media have low impact in driving traffic towards misinformation and in fact it seems that preventing access to these resources does not reduce the spread of misinformation. Then, individual-level interventions, such as improving digital literacy (Guess & Munger, 2022; Mosleh et al., 2021), would be more effective in preventing access towards highly ideological websites. Critical and analytic thinking, for example, are protective factors against misinformation (see Section 1.4.2.3).

3.3 Relevance and impact of the research

Following the COVID-19 pandemic, the dangerous effects associated with the spread of CTs were made salient to both lay people and academics (see e.g., Figure 1.2). Highly popular figures such as Donald Trump, Jair Bolsonaro, and Matteo Salvini (to cite a few) spread CTs rejecting the efforts made by the scientific community, impacting crucial public decision-making during the pandemic. Strategic or not, their use of CTs —from spreading false and misleading information to rejecting scientific evidence— has contributed to create a novel information landscape leading to negative consequences in the social sphere. To understand it, the scientific community had joined in unprecedented levels of multidisciplinary commitments from psychology, health, political, communication, and computational *sciences*. It is within this context, by helping to understand the content and spread of CTs, that this thesis is placed.

While there are widely used and validated questionnaires to assess people's conspiracy mentality and belief in CTs (see section 1.1), there are no similar tools, so far, to extract measures of conspiracism from texts. Because CTs spread as texts, which are a materialized form of belief capable of influencing people's attitudes, measuring conspiracism in texts is paramount to track and eventually limiting the spread of CTs —at least online. Knowing that conspiracy narratives are similar to each other reflecting an overarching worldview (as discussed in section 3.1.2) helps moving towards this goal. Dictionaries (i.e., list of words) that assess the probability of a text being conspiratorial seem foreseeable. Differently from fake news (see e.g., Pérez-Rosas et al., 2018), no classification algorithm has been developed so far to detect conspiracism from texts. In this regard, such a dictionary would represent an important step in practically advancing the development of automatic models for detection of

CTs online.

Envisioning a possible collaboration with social media (e.g., Facebook, Twitter), future experimental trials could implement such a dictionary and test the effectiveness of inoculation on reducing virality of misinformation. Politicians' or highly visible figures' social media content could be automatically analyzed for conspiratorial content. This ultimately would help improving people's information ecosystem, rising informed social media users or, at best, an informed electorate.

3.4 What's next

Parallel to the theoretical advancement in understanding the content and spread of CTs, this thesis required developing and applying new methodologies from computational science to psychology. These methods, made freely accessible in public repositories (e.g., on the Open Science Framework, see https://osf.io/4ghv5/), have been replicated, applied to, and discussed in other works I co-authored beyond my thesis (Carrella et al., in press; Hills & Miani, in press; Mayor & Miani, in review). The development of state-of-the-art skills in computational approaches linked to theory-driven psychological research provides a great research groundwork to undertake future works to move on better understanding the content and spread of CTs. Pragmatic extensions to the current field of investigation are foreseeable.

Upon completion of the PhD, in March 2023, I will spend a short research stay at the *Centre for Language Evolution Studies* (University of Torun, Poland) with Slawomir Wacewicz. In this stay, we will work on the evolution of CTs applying methods from cultural transmission studies (see e.g., Jagiello & Hills, 2018). We plan to explore how conspiratorial mentality affects the transmission of information, potentially generating CTs.

I have been awarded with a two-year SNSF Postdoc.Mobility fellowship, starting in May 2023, under the supervision of Stephan Lewandowsky at the *School of Psychological Science* (Bristol, UK). The project will extend on the possibility that in/coherence (see Study 2) is a more general psychological phenomenon associated with the strength of a(ny) worldview. The stronger a worldview, the higher the chance of endorsing incoherent and/or incompatible beliefs. For example, if one person strongly believes that immigration is bad, then s/he will likely hold, simultaneously, that immigrants are too lazy to work and that immigrants steal jobs. The project is divided into three parts and investigates in/coherence via two complimentary approaches, focusing on narratives (via text analyses) and on people's beliefs (via questionnaires and behavioral tasks). This allows to extend and generalize our previous findings while providing robust and convergent evidence from different fields/methods of investigation. Moreover, developing side projects started during the PhD, an algorithm to detect conspiratorial language will be finalized (which already yielded promising results, with an accuracy of 92%) and tested on a corpus built following LOCO methodology on websites from Study 4 comprising 1.5 million documents.

4 References

Abalakina-Paap, M., Stephan, W. G., Craig, T., & Gregory, W. L. (1999). Beliefs in conspiracies. *Political Psychology*, *20*(3), 637–647.

Acerbi, A. (2019). Cognitive attraction and online misinformation. *Palgrave Communications*, *5*(1), 15. https://doi.org/10.1057/s41599-019-0224-y

Albarracín, D. (2022). Processes of persuasion and social influence in conspiracy beliefs. *Current Opinion in Psychology*, *48*, 101463. https://doi.org/10.1016/j.copsyc.2022.101463

Alfano, M., Fard, A. E., Carter, J. A., Clutton, P., & Klein, C. (2021). Technologically scaffolded atypical cognition: The case of YouTube's recommender system. *Synthese*, *199*(1-2), 835–858. https://doi.org/10.1007/s11229-020-02724-x

Aliapoulios, M., Bevensee, E., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., & Zannettou, S. (2021). *An Early Look at the Parler Online Social Network* (No. arXiv:2101.03820). arXiv. https://arxiv.org/abs/2101.03820

Allen, J. (2008). No consensus on who was behind Sept 11: Global poll. *Reuters*. https://www.reuters.com/article/us-sept11-qaeda-poll-idUSN1035876620080910

Allington, D., & Joshi, T. (2020). "What others dare not say": An antisemitic conspiracy fantasy and its YouTube audience. *Journal of Contemporary Antisemitism*, *3*(1), 35–54. https://doi.org/10.26613/jca/3.1.42

Alper, S. (2022). There are higher levels of conspiracy beliefs in more corrupt countries. *European Journal of Social Psychology*, ejsp.2919. https://doi.org/10.1002/ejsp.2919

Aupers, S. (2012). "Trust no one": Modernization, paranoia and conspiracy culture. *European Journal of Communication*, *27*(1), 22–34. https://doi.org/10.1177/0267323111433566

AVAAZ. (2020). Facebook's Algorithm: A Major Threat to Public Health. In *Avaaz*. https://secure.avaaz.org/campaign/en/facebook_threat_health/.

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, *149*(8), 1608–1613. https://doi.org/10.1037/xge0000729

Ball, P. (2020). Anti-vaccine movement could undermine efforts to end coronavirus pandemic, researchers warn. *Nature*, *581*(7808), 251. https://doi.org/10.1038/d41586-020-01423-4

Bangerter, A., Wagner-Egger, P., & Delouvée, S. (2020). How Conspiracy Theories Spread. In M. Butter & P. Knight (Eds.), *Routledge Handbook of Conspiracy Theories* (First, pp. 206–218). Routledge. https://doi.org/10.4324/9780429452734-2_5

Barkun, M. (2017). President Trump and the "Fringe." *Terrorism and Political Violence*, *29*(3), 437–443. https://doi.org/10.1080/09546553.2017.1313649

Barlow, D. H., Durand, V. M., & Hofmann, S. G. (2017). *Abnormal Psychology: An Integrative Approach* (8th edition). Cengage Learning.

Barron, D., Furnham, A., Weis, L., Morgan, K. D., Towell, T., & Swami, V. (2018). The relationship between schizotypal facets and conspiracist beliefs via cognitive processes. *Psychiatry Research*, *259*, 15–20. https://doi.org/10.1016/j.psychres.2017.10.001

Barron, D., Morgan, K., Towell, T., Altemeyer, B., & Swami, V. (2014). Associations between schizotypy and belief in conspiracist ideation. *Personality and Individual Differences*, *70*, 156–159. https://doi.org/10.1016/j.paid.2014.06.040

Barron, F. (1993). Controllable Oddness as a Resource in Creativity. *Psychological Inquiry*, *4*(3), 182–184. https://doi.org/10.1207/s15327965pli0403_3

Bartlett, J., & Miller, C. (2010). *The power of unreason: Conspiracy theories, extremism and counter-terrorism*. Demos London.

Basit, A. (2021). Conspiracy theories and violent extremism. *Counter Terrorist Trends and Analyses*, *13*(3), 1–9.

Bavel, J. J. V., & Pereira, A. (2018). The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences*, *22*(3), 213–224. https://doi.org/10.1016/j.tics.2018.01.004

Berger, J., & Milkman, K. L. (2012). What Makes Online Content Viral? *Journal of Marketing Research*, *49*(2), 192–205. https://doi.org/10.1509/jmr.10.0353

Bessi, A. (2016). Personality traits and echo chambers on facebook. *Computers in Human Behavior*, *65*, 319–324. https://doi.org/10.1016/j.chb.2016.08.016

Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Science vs conspiracy: Collective narratives in the age of misinformation. *PLOS ONE*, *10*(2), e0118093. https://doi.org/10.1371/journal.pone.0118093

Bessi, A., Scala, A., Rossi, L., Zhang, Q., & Quattrociocchi, W. (2014). The economy of attention in the age of (mis)information. *Journal of Trust Management*, *1*(1), 12. https://doi.org/10.1186/s40493-014-0012-y

Bessi, A., Zollo, F., Del Vicario, M., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Trend of Narratives in the Age of Misinformation. *PLOS ONE*, *10*(8), e0134641. https://doi.org/10.1371/journal.pone.0134641

Betsch, C., Ulshöfer, C., Renkewitz, F., & Betsch, T. (2011). The Influence of Narrative v. Statistical Information on Perceiving Vaccination Risks. *Medical Decision Making*, *31*(5), 742–753. https://doi.org/10.1177/0272989X11400419

Biddlestone, M., Azevedo, F., & van der Linden, S. (2022). Climate of conspiracy: A meta-analysis of the consequences of belief in conspiracy theories about climate change. *Current Opinion in Psychology*, *46*, 101390. https://doi.org/10.1016/j.copsyc.2022.101390

Biddlestone, M., Green, R., Cichocka, A., Douglas, K., & Sutton, R. (2022). *A systematic review and meta-analytic synthesis of the motives associated with conspiracy beliefs*. PsyArXiv. https://doi.org/10.31234/osf.io/rxjqc

Biddlestone, M., Green, R., & Douglas, K. M. (2020). Cultural orientation, power, belief in conspiracy theories, and intentions to reduce the spread of COVID-19. *British Journal of Social Psychology*, *59*(3), 663–673. https://doi.org/10.1111/bjso.12397

Bierwiaczonek, K., Gundersen, A. B., & Kunst, J. R. (2022). The role of conspiracy beliefs for COVID-19 health responses: A meta-analysis. *Current Opinion in Psychology*, *46*, 101346. https://doi.org/10.1016/j.copsyc.2022.101346

Bierwiaczonek, K., Kunst, J. R., & Pich, O. (2020). Belief in COVID-19 Conspiracy Theories Reduces Social Distancing over Time. *Applied Psychology: Health and Well-Being*, *12*(4), 1270–1285. https://doi.org/10.1111/aphw.12223

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3*(null), 993–1022.

Bogart, L. M., & Bird, S. T. (2003). Exploring the relationship of conspiracy beliefs about HIV/AIDS to sexual behaviors and attitudes among African-American adults. *Journal of the National Medical Association*, *95*(11), 1057–1065.

Bogart, L. M., Wagner, G., Galvan, F. H., & Banks, D. (2010). Conspiracy Beliefs About HIV Are Related to Antiretroviral Treatment Nonadherence Among African American Men With HIV. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *53*(5), 648–655. https://doi.org/10.1097/QAI.0b013e3181c57dbc

Borum, R. (2014). Psychological Vulnerabilities and Propensities for Involvement in Violent Extremism. *Behavioral Sciences & the Law*, *32*(3), 286–305. https://doi.org/10.1002/bsl.2110

Bowes, S. M., Costello, T. H., Ma, W., & Lilienfeld, S. O. (2021). Looking under the tinfoil hat: Clarifying the personological and psychopathological correlates of conspiracy beliefs. *Journal of Personality*, *89*(3), 422–436. https://doi.org/10.1111/jopy.12588

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought* (Nachdr.). Basic Books.

Briand, S. C., Cinelli, M., Nguyen, T., Lewis, R., Prybylski, D., Valensise, C. M., Colizza, V., Tozzi, A. E., Perra, N., Baronchelli, A., Tizzoni, M., Zollo, F., Scala, A., Purnat, T., Czerniak, C., Kucharski, A. J., Tshangela, A., Zhou, L., & Quattrociocchi, W. (2021). Infodemics: A new challenge for public health. *Cell*, *184*(25), 6010–6014. https://doi.org/10.1016/j.cell.2021.10.031

Břízová, L., Gerbec, K., Šauer, J., & Šlégr, J. (2018). Flat Earth theory: An exercise in critical thinking. *Physics Education*, *53*(4), 045014. https://doi.org/10.1088/1361-6552/aac053

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*, *8*(1), 108–117. https://doi.org/10.1037/h0101832

Brooks, B. P., DiFonzo, N., & Ross, D. S. (2013). The GBN-dialogue model of outgroup-negative rumor transmission: Group membership, belief, and novelty. *Nonlinear Dynamics, Psychology, and Life Sciences*, *17*(2), 269–293.

Brotherton, R. (2016). *Suspicious minds: Why we believe conspiracy theories* (Paperback edition). Bloomsbury Sigma.

Brotherton, R., & French, C. C. (2014). Belief in Conspiracy Theories and Susceptibility to the Conjunction Fallacy. *Applied Cognitive Psychology*, *28*(2), 238–248. https://doi.org/10.1002/acp.2995

Brotherton, R., & French, C. C. (2015). Intention Seekers: Conspiracist Ideation and Biased Attributions of Intentionality. *PLOS ONE*, *10*(5), e0124125. https://doi.org/10.1371/journal.pone.0124125

Brotherton, R., French, C. C., & Pickering, A. D. (2013). Measuring Belief in Conspiracy Theories: The Generic Conspiracist Beliefs Scale. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00279

Bruder, M., Haffke, P., Neave, N., Nouripanah, N., & Imhoff, R. (2013). Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy mentality questionnaire. *Frontiers in Psychology*, *4*(April), 225. https://doi.org/10.3389/fpsyg.2013.00225

Bruder, M., & Manstead, A. S. R. (2009). *Questionnaire on conspiracy theories*. http://www.conspiracytheory.martinbruder.com/en/scale1.php.

Brugnoli, E., Cinelli, M., Quattrociocchi, W., & Scala, A. (2019). Recursive patterns in online echo chambers. *Scientific Reports*, *9*(1), 20118. https://doi.org/10.1038/s41598-019-56191-7

Butter, M., & Knight, P. (Eds.). (2023). *Covid conspiracy theories in global perspective* (1 Edition). Routledge.

Carrella, F., Miani, A., & Lewandowsky, S. (in press). IRMA: The 335-million-word italian coRpus for studying MisinformAtion. *The 17th Conference of the European Chapter of the Association for Computational Linguistics*.

Cassileth, B. R. (1984). Contemporary Unorthodox Treatments in Cancer Medicine. *Annals of Internal Medicine*, *101*(1), 105. https://doi.org/10.7326/0003-4819-101-1-105

Castelo, S., Santos, A., Almeida, T., Pham, K., Freire, J., Elghafari, A., & Nakamura, E. (2019). A topic-agnostic approach for identifying fake news pages. *The Web Conference 2019 - Companion of the World Wide Web Conference, WWW 2019*. https://doi.org/10.1145/3308560.3316739

Cichocka, A., Marchlewska, M., & Golec de Zavala, A. (2016). Does self-love or self-hate predict conspiracy beliefs? Narcissism, self-esteem, and the endorsement of conspiracy theories. *Social Psychological and Personality Science*, *7*(2), 157–166.

Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, *118*(9), e2023301118. https://doi.org/10.1073/pnas.2023301118

Clarke, S. (2007). Conspiracy Theories and the Internet: Controlled Demolition and Arrested Development. *Episteme*, *4*(2), 167–180. https://doi.org/10.3366/epi.2007.4.2.167

Colbert, S. M., & Peters, E. R. (2002). NEED FOR CLOSURE AND JUMPING-TO-CONCLUSIONS IN DELUSION-PRONE INDIVIDUALS: *The Journal of Nervous and Mental Disease*, *190*(1), 27–31. https://doi.org/10.1097/00005053-200201000-00007

Council of Canadian Academies. (2023). *Fault lines*. Expert Panel on the Socioeconomic Impacts of Science; Health Misinformation. https://cca-reports.ca/wp-content/uploads/2023/01/Report-Fault-Lines-digital-1.pdf

Craft, S., Ashley, S., & Maksl, A. (2017). News media literacy and conspiracy theory endorsement. *Communication and the Public*, *2*(4), 388–401. https://doi.org/10.1177/2057047317725539

Cropley, A. (2006). In Praise of Convergent Thinking. *Creativity Research Journal*, *18*(3), 391–404. https://doi.org/10.1207/s15326934crj1803_13

Dagnall, N., Denovan, A., Drinkwater, K., Parker, A., & Clough, P. J. (2017). Urban Legends and Paranormal Beliefs: The Role of Reality Testing and Schizotypy. *Frontiers in Psychology*, *8*, 942. https://doi.org/10.3389/fpsyg.2017.00942

Darwin, H., Neave, N., & Holmes, J. (2011). Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Personality and Individual Differences*, *50*(8), 1289–1293. https://doi.org/10.1016/j.paid.2011.02.027

Dave, D., McNichols, D., & Sabia, J. (2021). *Political violence, risk aversion, and non-localized disease spread: Evidence from the u.s. Capitol riot* (p. w28410). National Bureau of Economic Research. http://www.nber.org/papers/w28410.pdf

Davies, G., Wu, E., & Frank, R. (2021). A Witch's Brew of Grievances: The Potential Effects of COVID-19 on Radicalization to Violent Extremism. *Studies in Conflict & Terrorism*, 1–24. https://doi.org/10.1080/1057610X.2021.1923188

Davies, P. (2002). Antivaccination activists on the world wide web. *Archives of Disease in Childhood*, *87*(1), 22–25. https://doi.org/10.1136/adc.87.1.22

Davis, J., Wetherell, G., & Henry, P. J. (2018). Social devaluation of African Americans and race-related conspiracy theories. *European Journal of Social Psychology*, *48*(7), 999–1010. https://doi.org/10.1002/ejsp.2531

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo Chambers: Emotional Contagion and Group Polarization on Facebook. *Scientific Reports*, *6*(1), 37825. https://doi.org/10.1038/srep37825

Denovan, A., Dagnall, N., Drinkwater, K., Parker, A., & Neave, N. (2020). Conspiracist beliefs, intuitive thinking, and schizotypal facets: A further evaluation. *Applied Cognitive Psychology*, *34*(6), 1394–1405. https://doi.org/10.1002/acp.3716

DeVito, M. A. (2017). From Editors to Algorithms. *Digital Journalism*, *5*(6), 753–773. https://doi.org/10.1080/21670811.2016.1178592

Dieguez, S., Wagner-Egger, P., & Gauvrit, N. (2015). Nothing Happens by Accident, or Does It? A Low Prior for Randomness Does Not Explain Belief in Conspiracy Theories. *Psychological Science*, *26*(11), 1762–1770. https://doi.org/10.1177/0956797615598740

Dillard, J. (2001). Persuasion and the structure of affect. Dual systems and discrete emotions as complementary models. *Human Communication Research*, *27*(1), 38–68. https://doi.org/10.1093/hcr/27.1.38

Doosje, B., Moghaddam, F. M., Kruglanski, A. W., Wolf, A. de, Mann, L., & Feddes, A. R. (2016). Terrorism, radicalization and de-radicalization. *Current Opinion in Psychology*, *11*, 79–84. https://doi.org/10.1016/j.copsyc.2016.06.008

Douglas, K. M., & Sutton, R. M. (2023). What Are Conspiracy Theories? A Definitional Approach to Their Correlates, Consequences, and Communication. *Annual Review of Psychology*, *74*(1), annurev-psych-032420-031329. https://doi.org/10.1146/annurev-psych-032420-031329

Douglas, K. M., Sutton, R. M., & Cichocka, A. (2017). The psychology of conspiracy theories. *Current Directions in Psychological Science*, *26*(6), 538–542. https://doi.org/10.1177/0963721417718261

Douglas, K. M., Uscinski, J., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding conspiracy theories. *Political Psychology*, *40*(S1), 3–35. https://doi.org/10.1111/pops.12568

Dyrendal, A., Kennair, L. E. O., & Bendixen, M. (2021). Predictors of belief in conspiracy theory: The role of individual differences in schizotypal traits, paranormal beliefs, social dominance orientation, right wing authoritarianism and conspiracy mentality. *Personality and Individual Differences*, *173*, 110645. https://doi.org/10.1016/j.paid.2021.110645

Edelson, J., Alduncin, A., Krewson, C., Sieja, J. A., & Uscinski, J. (2017). The Effect of

Conspiratorial Thinking and Motivated Reasoning on Belief in Election Fraud. *Political Research Quarterly*, *70*(4), 933–946. https://doi.org/10.1177/1065912917721061

Eicher, V., & Bangerter, A. (2015). Social representations of infectious diseases. In G. Sammut, E. Andreouli, G. Gaskell, & J. Valsiner (Eds.), *The Cambridge Handbook of Social Representations* (First, pp. 385–396). Cambridge University Press. https://doi.org/10.1017/CBO9781107323650.031

Einstein, K. L., & Glick, D. M. (2015). Do I Think BLS Data are BS? The Consequences of Conspiracy Theories. *Political Behavior*, *37*(3), 679–701. https://doi.org/10.1007/s11109-014-9287-z

Elvevåg, B., Foltz, P. W., Weinberger, D. R., & Goldberg, T. E. (2007). Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophrenia Research*, *93*(1-3), 304–316. https://doi.org/10.1016/j.schres.2007.03.001

Enders, A. M., Smallpage, S. M., & Lupton, R. N. (2018). Are All "Birthers" Conspiracy Theorists? On the Relationship Between Conspiratorial Thinking and Political Orientations. *British Journal of Political Science*, 1–18. https://doi.org/10.1017/s0007123417000837

Ettinger, U., Meyhöfer, I., Steffens, M., Wagner, M., & Koutsouleris, N. (2014). Genetics, Cognition, and Neurobiology of Schizotypal Personality: A Review of the Overlap with Schizophrenia. *Frontiers in Psychiatry*, *5*. https://doi.org/10.3389/fpsyt.2014.00018

Eysenck, H. J. (2003). Creativity, personality and the convergent-divergent continuum. In *Critical creative processes* (pp. 95–114). Hampton Press.

Faasse, K., Chatman, C. J., & Martin, L. R. (2016). A comparison of language use in pro- and anti-vaccination comments in response to a high profile Facebook post,. *Vaccine*, *34*(47), 5808–5814. https://doi.org/10.1016/j.vaccine.2016.09.029

Faddoul, M., Chaslot, G., & Farid, H. (2020). *A Longitudinal Analysis of YouTube's Promotion of Conspiracy Videos*. https://doi.org/10.48550/ARXIV.2003.03318

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Fiagbenu, M. E. (2022). The stock market is rigged? Conspiracy beliefs and distrust predict lower stock market participation. *Applied Cognitive Psychology*, acp.3981. https://doi.org/10.1002/acp.3981

Fong, A., Roozenbeek, J., Goldwert, D., Rathje, S., & Linden, S. van der. (2021). The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations*, *24*(4), 606–623. https://doi.org/10.1177/1368430220987596

Franks, B., Bangerter, A., & Bauer, M. W. (2013). Conspiracy theories as quasi-religious mentality: An integrated account from cognitive science, social representations theory, and frame theory. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00424

Franks, B., Bangerter, A., Bauer, M. W., Hall, M., & Noort, M. C. (2017). Beyond "Monologicality"? Exploring Conspiracist Worldviews. *Frontiers in Psychology*, *8*, 861. https://doi.org/10.3389/fpsyg.2017.00861

Freeman, D., & Bentall, R. P. (2017). The concomitants of conspiracy concerns. *Social Psychiatry and Psychiatric Epidemiology*, *52*(5), 595–604. https://doi.org/10.1007/s00127-017-1354-4

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., Kuipers, E., Ray, K., & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. *British Journal of Psychiatry*, *186*(5), 427–435. https://doi.org/10.1192/bjp.186.5.427

Freeman, D., Waite, F., Rosebrock, L., Petit, A., Causier, C., East, A., Jenner, L., Teale, A.-L., Carr, L., Mulhall, S., Bold, E., & Lambe, S. (2022). Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychological Medicine*, *52*(2), 251–263. https://doi.org/10.1017/S0033291720001890

Friggeri, A., Adamic, L., Eckles, D., & Cheng, J. (2014). Rumor cascades. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*, 101–110.

Fu, L. Y., Zook, K., Spoehr-Labutta, Z., Hu, P., & Joseph, J. G. (2016). Search Engine Ranking, Quality, and Content of Web Pages That Are Critical Versus Noncritical of Human Papillomavirus Vaccine. *Journal of Adolescent Health*, *58*(1), 33–39. https://doi.org/10.1016/j.jadohealth.2015.09.016

Furnham, A., & Grover, S. (2022). Do you have to be mad to believe in conspiracy theories? Personality disorders and conspiracy theories. *International Journal of Social Psychiatry*, *68*(7), 1454–1461. https://doi.org/10.1177/00207640211031614

Gangarosa, E. J., Galazka, A. M., Wolfe, C. R., Phillips, L. M., Miller, E., Chen, R. T., & Gangarosa, R. E. (1998). Impact of anti-vaccine movements on pertussis control: The untold story. *The Lancet*, *351*(9099), 356–361. https://doi.org/10.1016/S0140-6736(97)04334-1

Garry, A., Walther, S., Rukaya, R., & Mohammed, A. (2021). QAnon conspiracy theory: Examining its evolution and mechanisms of radicalization. *Journal for Deradicalization*, *26*, 152–216.

Georgiou, N., Delfabbro, P., & Balzan, R. (2019). Conspiracy beliefs in the general population: The importance of psychopathology, cognitive style and educational attainment. *Personality and Individual Differences*, *151*, 109521. https://doi.org/10.1016/j.paid.2019.109521

Georgiou, N., Delfabbro, P., & Balzan, R. (2021). Conspiracy-beliefs and receptivity to disconfirmatory information: A study using the BADE task. *SAGE Open*, *11*(1), 215824402110061. https://doi.org/10.1177/21582440211006131

Geys, B., Heggedal, T.-R., & Sørensen, R. J. (2022). Age and vote choice: Is there a conservative shift among older voters? *Electoral Studies*, *78*, 102485. https://doi.org/10.1016/j.electstud.2022.102485

Goertzel, T. (1994). Belief in conspiracy theories. *Political Psychology*, *15*(4), 731. https://doi.org/10.2307/3791630

Golec de Zavala, A., & Cichocka, A. (2012). Collective narcissism and anti-semitism in poland. *Group Processes & Intergroup Relations*, *15*(2), 213–229. https://doi.org/10.1177/1368430211420891

Golec de Zavala, A., Cichocka, A., Eidelson, R., & Jayawickreme, N. (2009). Collective narcissism and its social consequences. *Journal of Personality and Social Psychology*, *97*(6), 1074–1096. https://doi.org/10.1037/a0016904

Golec de Zavala, A., Dyduch-Hazar, K., & Lantos, D. (2019). Collective Narcissism: Political Consequences of Investing Self-Worth in the Ingroup's Image. *Political Psychology*, *40*(S1), 37–74. https://doi.org/10.1111/pops.12569

Goodwin, J. (2019). Sophistical refutations in the climate change debates. *Journal of Argumentation in Context*, *8*(1), 40–64. https://doi.org/10.1075/jaic.18008.goo

Goreis, A., & Voracek, M. (2019). A systematic review and meta-analysis of psychological research on conspiracy beliefs: Field characteristics, measurement instruments, and associations with personality traits. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00205

Green, R., & Douglas, K. M. (2018). Anxious attachment and belief in conspiracy theories. *Personality and Individual Differences*, *125*, 30–37. https://doi.org/10.1016/j.paid.2017.12.023

Greenberg, J., & Jonas, E. (2003). Psychological motives and political orientation–The left, the right, and the rigid: Comment on Jost et al. (2003). *Psychological Bulletin*, *129*(3), 376–382. https://doi.org/10.1037/0033-2909.129.3.376

Grimes, D. R. (2016). On the Viability of Conspiratorial Beliefs. *Plos One*, *11*(1), e0147905. https://doi.org/10.1371/journal.pone.0147905

Grzesiak-Feldman, M. (2007). Conspiracy Thinking and State-Trait Anxiety in Young Polish Adults. *Psychological Reports*, *100*(1), 199–202. https://doi.org/10.2466/pr0.100.1.199-202

Grzesiak-Feldman, M. (2013). The Effect of High-Anxiety Situations on Conspiracy Thinking. *Current Psychology*, *32*(1), 100–118. https://doi.org/10.1007/s12144-013-9165-6

Guess, A., & Munger, K. (2022). Digital literacy and online political behavior. *Political Science Research and Methods*, 1–19. https://doi.org/10.1017/psrm.2022.17

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on facebook. *Science Advances*, *5*(1). https://doi.org/10.1126/sciadv.aau4586

Guess, A., Nyhan, B., Lyons, B., & Reifler, J. (2018). Avoiding the echo chamber about echo chambers. *Knight Foundation*, *2*, 1–25.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council*, *9*(3), 4.

Guess, A., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*, *4*(5), 472–480. https://doi.org/10.1038/s41562-020-0833-x

Harambam, J., & Aupers, S. (2017). "I Am Not a Conspiracy Theorist": Relational Identifications in the Dutch Conspiracy Milieu. *Cultural Sociology*, *11*(1), 113–129. https://doi.org/10.1177/1749975516661959

Hart, J., & Graether, M. (2018). Something's Going on Here: Psychological Predictors of Belief in Conspiracy Theories. *Journal of Individual Differences*, *39*(4), 229–237. https://doi.org/10.1027/1614-0001/a000268

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–588. https://doi.org/10.1037/a0015701

Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology*, *81*(6), 1028–1041. https://doi.org/10.1037/0022-3514.81.6.1028

Hills, T. (2019). The Dark Side of Information Proliferation. *Perspectives on Psychological Science*, *14*(3), 323–330. https://doi.org/10.1177/1745691618803647

Hills, T., & Miani, A. (in press). A short primer on historical natural language processing. In T. Hills & G. Pogrebna (Eds.), *Cambridge handbook of behavioral data science* (Cambridge University Press).

Hoffman, B. L., Felter, E. M., Chu, K.-H., Shensa, A., Hermann, C., Wolynn, T., Williams, D., & Primack, B. A. (2019). It's not all about autism: The emerging landscape of anti-vaccination sentiment on Facebook. *Vaccine*, *37*(16), 2216–2223. https://doi.org/10.1016/j.vaccine.2019.03.003

Hofstadter, R. (1964). *The paranoid style in American politics, and other essays* (1st Harvard University Press pbk. ed). Harvard University Press.

Hornsey, M. J., Bierwiaczonek, K., Sassenberg, K., & Douglas, K. M. (2022). Individual, intergroup and nation-level influences on belief in conspiracy theories. *Nature Reviews Psychology*. https://doi.org/10.1038/s44159-022-00133-0

Hornsey, M. J., & Pearson, S. (2022). Cross-national differences in willingness to believe conspiracy theories. *Current Opinion in Psychology*, *47*, 101391. https://doi.org/10.1016/j.copsyc.2022.101391

Hoseini, M., Melo, P., Benevenuto, F., Feldmann, A., & Zannettou, S. (2021). *On the Globalization of the QAnon Conspiracy Theory Through Telegram* (No. arXiv:2105.13020). arXiv. https://arxiv.org/abs/2105.13020

Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1), 1–27. https://doi.org/10.1145/3392854

Imhoff, R. (2022). Conspiracy Theories Through a Cross-Cultural Lens. *Online Readings in Psychology and Culture*, *5*(3). https://doi.org/10.9707/2307-0919.1175

Imhoff, R., Bertlich, T., & Frenken, M. (2022). Tearing apart the "evil" twins: A general conspiracy mentality is not the same as specific conspiracy beliefs. *Current Opinion in Psychology*, *46*, 101349. https://doi.org/10.1016/j.copsyc.2022.101349

Imhoff, R., Dieterle, L., & Lamberty, P. (2021). Resolving the puzzle of conspiracy worldview and political activism: Belief in secret plots decreases normative but increases nonnormative political engagement. *Social Psychological and Personality Science*, *12*(1), 71–79. https://doi.org/10.1177/1948550619896491

Imhoff, R., & Lamberty, P. K. (2017). Too special to be duped: Need for uniqueness motivates conspiracy beliefs. *European Journal of Social Psychology*, *00*. https://doi.org/10.1002/ejsp.2265

Imhoff, R., Lamberty, P., & Klein, O. (2018). Using Power as a Negative Cue: How Conspiracy Mentality Affects Epistemic Trust in Sources of Historical Knowledge. *Personality and Social Psychology Bulletin*, *44*(9), 1364–1379. https://doi.org/10.1177/0146167218768779

Imhoff, R., Zimmer, F., Klein, O., António, J. H. C., Babinska, M., Bangerter, A., Bilewicz, M., Blanuša, N., Bovan, K., Bužarovska, R., Cichocka, A., Delouvée, S., Douglas, K. M., Dyrendal, A., Etienne, T., Gjoneska, B., Graf, S., Gualda, E., Hirschberger, G., ... Prooijen, J.-W. van. (2022). Conspiracy mentality and political orientation across 26 countries. *Nature Human Behaviour*, *6*(3), 392–403. https://doi.org/10.1038/s41562-021-01258-7

Jagiello, R. D., & Hills, T. (2018). Bad News Has Wings: Dread Risk Mediates Social Amplification in Risk Communication. *Risk Analysis*, *38*(10), 2193–2207. https://doi.org/10.1111/risa.13117

Jolley, D., & Douglas, K. M. (2014a). The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint. *British Journal of Psychology*, *105*(1), 35–56. https://doi.org/10.1111/bjop.12018

Jolley, D., & Douglas, K. M. (2014b). The Effects of Anti-Vaccine Conspiracy Theories on Vaccination Intentions. *PLoS ONE*, *9*(2), e89177. https://doi.org/10.1371/journal.pone.0089177

Jolley, D., Douglas, K. M., Leite, A. C., & Schrader, T. (2019). Belief in conspiracy theories and intentions to engage in everyday crime. *British Journal of Social Psychology*, *58*(3), 534–549. https://doi.org/10.1111/bjso.12311

Jolley, D., & Paterson, J. L. (2020). Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *British Journal of Social Psychology*, *59*(3), 628–640. https://doi.org/10.1111/bjso.12394

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003a). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375. https://doi.org/10.1037/0033-2909.129.3.339

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003b). Exceptions that prove the rule–Using a theory of motivated social cognition to account for ideological incongruities and political anomalies: Reply to Greenberg and Jonas (2003). *Psychological Bulletin*, *129*(3), 383–393. https://doi.org/10.1037/0033-2909.129.3.383

Juárez-Ramos, V., Rubio, J. L., Delpero, C., Mioni, G., Stablum, F., & Gómez-Milán, E. (2014). Jumping to Conclusions bias, BADE and Feedback Sensitivity in schizophrenia and schizotypy. *Consciousness and Cognition*, *26*, 133–144. https://doi.org/10.1016/j.concog.2014.03.006

Kaplan, J. (2021). A Conspiracy of Dunces: Good Americans vs. A Cabal of Satanic Pedophiles? *Terrorism and Political Violence*, *33*(5), 917–921. https://doi.org/10.1080/09546553.2021.1932342

Kauk, J., Kreysa, H., & Schweinberger, S. R. (2021). Understanding and countering the spread of conspiracy theories in social networks: Evidence from epidemiological models of Twitter data. *PLOS ONE*, *16*(8), e0256179. https://doi.org/10.1371/journal.pone.0256179

Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J., & Laurin, K. (2008). God and the government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology*, *95*(1), 18–35. https://doi.org/10.1037/0022-3514.95.1.18

Kay, C. S. (2021). The targets of all treachery: Delusional ideation, paranoia, and the need for uniqueness as mediators between two forms of narcissism and conspiracy beliefs. *Journal of Research in Personality*, *93*, 104128. https://doi.org/10.1016/j.jrp.2021.104128

Keeley, B. L. (1999). Of Conspiracy Theories. *The Journal of Philosophy*, *96*(3), 109. https://doi.org/10.2307/2564659

Klein, C., Clutton, P., & Dunn, A. G. (2019). Pathways to conspiracy: The social and linguistic precursors of involvement in Reddit's conspiracy theory forum. *PLOS ONE*, *14*(11), e0225098. https://doi.org/10.1371/journal.pone.0225098

Klein, C., Clutton, P., & Polito, V. (2018). Topic Modeling Reveals Distinct Interests within an Online Conspiracy Forum. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00189

Knobloch-Westerwick, S., & Meng, J. (2009). Looking the other way: Selective exposure to attitude-consistent and counterattitudinal political information. *Communication Research*, *36*(3), 426–448. https://doi.org/10.1177/0093650209333030

Koltko-Rivera, M. E. (2004). The psychology of worldviews. *Review of General Psychology*, *8*(1), 3–58. https://doi.org/10.1037/1089-2680.8.1.3

Kruglanski, A. W., Molinario, E., Ellenberg, M., & Di Cicco, G. (2022). Terrorism and conspiracy theories: A view from the 3N model of radicalization. *Current Opinion in Psychology*, *47*, 101396. https://doi.org/10.1016/j.copsyc.2022.101396

Kuhn, S. A. K., Lieb, R., Freeman, D., Andreou, C., & Zander-Schellenberg, T. (2021). Coronavirus conspiracy beliefs in the german-speaking general population: Endorsement rates and links to reasoning biases and paranoia. *Psychological Medicine*, 1–15. https://doi.org/10.1017/S0033291721001124

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Kwon, S., Cha, M., & Jung, K. (2017). Rumor Detection over Varying Time Windows. *PLOS ONE*, *12*(1), e0168344. https://doi.org/10.1371/journal.pone.0168344

Lantian, A., Muller, D., Nurra, C., & Douglas, K. M. (2016). Measuring belief in conspiracy theories: Validation of a French and English single-item scale. *International Review of Social Psychology*, *29*(1), 1–14.

Lantian, A., Muller, D., Nurra, C., & Douglas, K. M. (2017). "I Know Things They Don't Know!" *Social Psychology*, *48*(3), 160–173. https://doi.org/10.1027/1864-9335/a000306

Lantian, A., Muller, D., Nurra, C., Klein, O., Berjot, S., & Pantazi, M. (2018). Stigmatized beliefs: Conspiracy theories, anticipated negative evaluation of the self, and fear of social exclusion. *European Journal of Social Psychology*, *48*(7), 939–954. https://doi.org/10.1002/ejsp.2498

Larsen, E. M., Donaldson, K. R., Liew, M., & Mohanty, A. (2021). Conspiratorial Thinking During COVID-19: The Roles of Paranoia, Delusion-Proneness, and Intolerance of Uncertainty. *Frontiers in Psychiatry*, *12*, 698147. https://doi.org/10.3389/fpsyt.2021.698147

Lasser, J., Aroyehun, S. T., Simchon, A., Carrella, F., Garcia, D., & Lewandowsky, S. (2022). Social media sharing of low-quality news sources by political elites. *PNAS Nexus*, *1*(4), pgac186. https://doi.org/10.1093/pnasnexus/pgac186

Lawson, M. A., & Kakkar, H. (2021). Of pandemics, politics, and personality: The role of conscientiousness and political ideology in the sharing of fake news. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001120

Lazarus, J. V., Ratzan, S. C., Palayew, A., Gostin, L. O., Larson, H. J., Rabin, K., Kimball, S., & El-Mohandes, A. (2020). A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine*. https://doi.org/10.1038/s41591-020-1124-9

Leman, P. J., & Cinnirella, M. (2013). Beliefs in conspiracy theories and the need for cognitive closure. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00378

Lewandowsky, S., Cook, J., & Lloyd, E. (2018). The 'alice in wonderland' mechanics of the rejection of (climate) science: Simulating coherence by conspiracism. *Synthese*, *195*(1), 175–196. https://doi.org/10.1007/s11229-016-1198-6

Lewandowsky, S., & Oberauer, K. (2021). Worldview-motivated rejection of science and the norms of science. *Cognition*, *215*, 104820. https://doi.org/10.1016/j.cognition.2021.104820

Liekefett, L., Christ, O., & Becker, J. C. (2021). Can conspiracy beliefs be beneficial? Longitudinal linkages between conspiracy beliefs, anxiety, uncertainty aversion, and existential threat. *Personality and Social Psychology Bulletin*, 014616722110609. https://doi.org/10.1177/01461672211060965

Linden, S. van der. (2015). The conspiracy-effect: Exposure to conspiracy theories (about global warming) decreases pro-social behavior and science acceptance. *Personality and Individual Differences*, *87*, 171–173. https://doi.org/10.1016/j.paid.2015.07.045

Linden, S., Panagopoulos, C., Azevedo, F., & Jost, J. T. (2021). The Paranoid Style in American Politics Revisited: An Ideological Asymmetry in Conspiratorial Thinking. *Political Psychology*, *42*(1), 23–51. https://doi.org/10.1111/pops.12681

Lobato, E., Mendoza, J., Sims, V., & Chin, M. (2014). Examining the Relationship Between Conspiracy Theories, Paranormal Beliefs, and Pseudoscience Acceptance Among a University Population: Relationship between unwarranted beliefs. *Applied Cognitive Psychology*, *28*(5), 617–625. https://doi.org/10.1002/acp.3042

Lu, J. G., Jin, P., & English, A. S. (2021). Collectivism predicts mask use during COVID-19. *Proceedings of the National Academy of Sciences*, *118*(23), e2021793118. https://doi.org/10.1073/pnas.2021793118

Lukić, P., Žeželj, I., & Stanković, B. (2019). How (ir)rational is it to believe in contradictory conspiracy theories? *Europe's Journal of Psychology*, *15*(1), 94–107. https://doi.org/10.5964/ejop.v15i1.1690

Marchlewska, M., Cichocka, A., & Kossowska, M. (2018). Addicted to answers: Need for cognitive closure and the endorsement of conspiracy beliefs. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2308

Marchlewska, M., Cichocka, A., Łozowski, F., Górska, P., & Winiewski, M. (2019). In search of an imaginary enemy: Catholic collective narcissism and the endorsement of *Gender Conspiracy Beliefs*. *The Journal of Social Psychology*, *159*(6), 766–779. https://doi.org/10.1080/00224545.2019.1586637

Marie, A., & Petersen, M. B. (2022). Political conspiracy theories as tools for mobilization and signaling. *Current Opinion in Psychology*, *48*, 101440. https://doi.org/10.1016/j.copsyc.2022.101440

Mashuri, A., & Zaduqisti, E. (2014). We believe in your conspiracy if we distrust you: The role of intergroup distrust in structuring the effect of Islamic identification, competitive victimhood, and group incompatibility on belief in a conspiracy theory. *Journal of Tropical Psychology*, *4*, e11. https://doi.org/10.1017/jtp.2014.11

Mathew, B., Dutt, R., Goyal, P., & Mukherjee, A. (2019). Spread of Hate Speech in Online Social Media. *Proceedings of the 10th ACM Conference on Web Science*, 173–182. https://doi.org/10.1145/3292522.3326034

Matthews, D. (2017). *Donald Trump has tweeted climate change skepticism 115 times.*

https://www.vox.com/policy-and-politics/2017/6/1/15726472/trump-tweets-global-warming-paris-climate-ag
Mayor, E., & Miani, A. (in review). "Will omicron ruin christmas?" A topic models analysis of the press coverage of the new SARS-Cov2 variant in the united kingdom. *BMC Public Health*.

McCauley, C., & Moskalenko, S. (2008). Mechanisms of Political Radicalization: Pathways Toward Terrorism. *Terrorism and Political Violence*, *20*(3), 415–433. https://doi.org/10.1080/09546550802073367

McCright, A. M., & Dunlap, R. E. (2011). Cool dudes: The denial of climate change among conservative white males in the United States. *Global Environmental Change*, *21*(4), 1163–1172. https://doi.org/10.1016/j.gloenvcha.2011.06.003

Midlarsky, M. I. (2011). *Origins of political extremism: Mass violence in the twentieth century and beyond*. Cambridge University Press.

Miller, J. M. (2020). Do COVID-19 Conspiracy Theory Beliefs Form a Monological Belief System? *Canadian Journal of Political Science*, *53*(2), 319–326. https://doi.org/10.1017/S0008423920000517

Min, S. J. (2021). Who Believes in Conspiracy Theories? Network Diversity, Political Discussion, and Conservative Conspiracy Theories on Social Media. *American Politics Research*, *49*(5), 415–427. https://doi.org/10.1177/1532673X211013526

Mitra, T., Counts, S., & Pennebaker, J. (2021). Understanding Anti-Vaccination Attitudes in Social Media. *Proceedings of the International AAAI Conference on Web and Social Media*, *10*(1), 269–278. https://doi.org/10.1609/icwsm.v10i1.14729

Mocanu, D., Rossi, L., Zhang, Q., Karsai, M., & Quattrociocchi, W. (2015). Collective attention in the age of (mis)information. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2015.01.024

Mosleh, M., Pennycook, G., Arechar, A. A., & Rand, D. G. (2021). Cognitive reflection correlates with behavior on Twitter. *Nature Communications*, *12*(1), 921. https://doi.org/10.1038/s41467-020-20043-0

Newheiser, A.-K., Farias, M., & Tausch, N. (2011). The functional nature of conspiracy beliefs: Examining the underpinnings of belief in the Da Vinci Code conspiracy. *Personality and Individual Differences*, *51*(8), 1007–1011. https://doi.org/10.1016/j.paid.2011.08.011

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and Mystery: The Cultural Selection of Minimally Counterintuitive Narratives. *Cognitive Science*, *30*(3), 531–553. https://doi.org/10.1207/s15516709cog0000_68

Nyhan, B. (2017). Why More Democrats Are Now Embracing Conspiracy Theories. *The New York Times*.

Oliver, J. E., & Wood, T. (2014). Medical Conspiracy Theories and Health Behaviors in the United States. *JAMA Internal Medicine*, *174*(5), 817. https://doi.org/10.1001/jamainternmed.2014.190

Olshansky, A., Peaslee, R. M., & Landrum, A. R. (2020). Flat-Smacked! Converting to Flat Eartherism. *Journal of Media and Religion*, *19*(2), 46–59. https://doi.org/10.1080/15348423.2020.1774257

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021).

Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter. *American Political Science Review*, *115*(3), 999–1015. https://doi.org/10.1017/S0003055421000290

Oswald, S. (2016). Conspiracy and bias: Argumentative features and persuasiveness of conspiracy theories. *OSSA Conference Archive*, *168*, 1–16.

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. penguin UK.

Pearson, G. D. H., & Knobloch-Westerwick, S. (2019). Is the confirmation bias bubble larger online? Pre-election confirmation bias in selective exposure to online versus print political information. *Mass Communication and Society*, *22*(4), 466–486. https://doi.org/10.1080/15205436.2019.1599956

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865–1880. https://doi.org/10.1037/xge0000465

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science*, *31*(7), 770–780. https://doi.org/10.1177/0956797620939054

Pennycook, G., & Rand, D. G. (2019a). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, *116*(7), 2521–2526. https://doi.org/10.1073/pnas.1806781116

Pennycook, G., & Rand, D. G. (2019b). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Perez, J. C., & Montagnier, L. (2020). Covid-19, Sars And Bats Coronaviruses Genomes Peculiar Homologous RNA Sequences. *International Journal of Research -GRANTHAALAYAH*, *8*(7), 217–263. https://doi.org/10.29121/granthaalayah.v8.i7.2020.678

Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2018). Automatic detection of fake news. *Proceedings of the 27th International Conference on Computational Linguistics*, 3391–3401. https://aclanthology.org/C18-1287

Peterson, J. C., Smith, K. B., & Hibbing, J. R. (2020). Do People Really Become More Conservative as They Age? *The Journal of Politics*, *82*(2), 600–611. https://doi.org/10.1086/706889

Pigden, C. (1995). Popper Revisited, or What Is Wrong With Conspiracy Theories? *Philosophy of the Social Sciences*, *25*(1), 3–34. https://doi.org/10.1177/004839319502500101

Piltch-Loeb, R., Silver, D., Kim, Y., Norris, H., McNeill, E., & Abramson, D. (2021). Determinants of the COVID-19 Vaccine Hesitancy Spectrum. *medRxiv*, 2021.08.05.21261675. https://doi.org/10.1101/2021.08.05.21261675

Pipes, D. (1999). *Conspiracy: How the paranoid style flourishes and where it comes from*. Simon and Schuster.

Pisl, V., Volavka, J., Chvojkova, E., Cechova, K., Kavalirova, G., & Vevera, J. (2021). Dissociation, Cognitive Reflection and Health Literacy Have a Modest Effect on Belief in Conspiracy Theories about COVID-19. *International Journal of Environmental Research* *and Public Health*, *18*(10), 5065. https://doi.org/10.3390/ijerph18105065

Pratt, R. (2003). Theorizing conspiracy. *Theory and Society*, *32*(2), 255–271. https://doi.org/10.1023/A:1023996501425

Prooijen, J.-W. van, Krouwel, A. P. M., Boiten, M., & Eendebak, L. (2015). Fear among the extremes. *Personality and Social Psychology Bulletin*, *41*(4), 485–497. https://doi.org/10.1177/0146167215569706

Prooijen, J.-W. van, & Kuijper, S. M. H. C. (2020). A comparison of extreme religious and political ideologies: Similar worldviews but different grievances. *Personality and Individual Differences*, *159*, 109888. https://doi.org/10.1016/j.paid.2020.109888

Pummerer, L. (2022). Belief in conspiracy theories and non-normative behavior. *Current Opinion in Psychology*, 101394. https://doi.org/10.1016/J.COPSYC.2022.101394

Pummerer, L., Böhm, R., Lilleholt, L., Winter, K., Zettler, I., & Sassenberg, K. (2022). Conspiracy Theories and Their Societal Effects During the COVID-19 Pandemic. *Social Psychological and Personality Science*, *13*(1), 49–59. https://doi.org/10.1177/19485506211000217

Raab, M. H., Auer, N., Ortlieb, S. A., & Carbon, C.-C. (2013). The Sarrazin effect: The presence of absurd statements in conspiracy theories makes canonical information less plausible. *Frontiers in Psychology*, *4*(JUL), 1–8. https://doi.org/10.3389/fpsyg.2013.00453

Raab, M. H., Ortlieb, S. A., Auer, N., Guthmann, K., & Carbon, C.-C. (2013). Thirty shades of truth: Conspiracy theories as stories of individuation, not of pathological delusion. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00406

Radnitz, S., & Underwood, P. (2017). Is Belief in Conspiracy Theories Pathological? A Survey Experiment on the Cognitive Roots of Extreme Suspicion. *British Journal of Political Science*, *47*(1), 113–129. https://doi.org/10.1017/S0007123414000556

Raemdonck, N. van. (2019). The echo chamber of anti-vaccination conspiracies: Mechanisms of radicalization on Facebook and Reddit. *Institute for Policy, Advocacy and Governance (IPAG) Knowledge Series, Forthcoming*.

Rathje, S., Roozenbeek, J., Traberg, C. S., Van Bavel, J. J., & Linden, S. van der. (2022). *Letter to the Editors of Psychological Science: Meta-Analysis Reveals that Accuracy Nudges Have Little to No Effect for U.S. Conservatives: Regarding Pennycook et al. (2020)* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/945na

Ren, Z. (Bella)., Dimant, E., & Schweitzer, M. (2023). Beyond belief: How social engagement motives influence the spread of conspiracy theories. *Journal of Experimental Social Psychology*, *104*, 104421. https://doi.org/10.1016/j.jesp.2022.104421

Rezaii, N., Walker, E., & Wolff, P. (2019). A machine learning approach to predicting psychosis using semantic density and latent content analysis. *Npj Schizophrenia*, *5*(1), 9. https://doi.org/10.1038/s41537-019-0077-9

Runco, M. A. (2010). Divergent Thinking, Creativity, and Ideation. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (First, pp. 413–446). Cambridge University Press. https://doi.org/10.1017/CBO9780511763205.026

Salmon, D. A., Moulton, L. H., Omer, S. B., DeHart, M. P., Stokley, S., & Halsey, N. A. (2005). Factors associated with refusal of childhood vaccines among parents of school-aged children: A case-control study. *Archives of Pediatrics & Adolescent Medicine*, *159*(5), 470–476. https://doi.org/10.1001/archpedi.159.5.470

Salvador Casara, B. G., Suitner, C., & Jetten, J. (2022). The impact of economic inequality on conspiracy beliefs. *Journal of Experimental Social Psychology*, *98*, 104245. https://doi.org/10.1016/j.jesp.2021.104245

Samory, M., & Mitra, T. (2018a). Conspiracies online: User discussions in a conspiracy community following dramatic events. *12th International AAAI Conference on Web and Social Media, ICWSM 2018*.

Samory, M., & Mitra, T. (2018b). "The Government Spies Using Our Webcams:" The Language of Conspiracy Theories in Online Discussions. *Proceedings of the ACM on Human-Computer Interaction*, *2*. https://doi.org/10.1145/3274421

Schomaker, J., & Meeter, M. (2015). Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neuroscience & Biobehavioral Reviews*, *55*, 268–279. https://doi.org/10.1016/j.neubiorev.2015.05.002

Sharon Parsons, William Simmons, Fr. (1999). A TEST OF THE GRAPEVINE: AN EMPIRICAL EXAMINATION OF CONSPIRACY THEORIES AMONG AFRICAN AMERICANS. *Sociological Spectrum*, *19*(2), 201–222. https://doi.org/10.1080/027321799280235

Silvia, P. J. (2007). *How to Write a Lot: A Practical Guide to Productive Academic Writing*. American Psychological Association.

Simmons, W. P., & Parsons, S. (2005). Beliefs in Conspiracy Theories Among African Americans: A Comparison of Elites and Masses*. *Social Science Quarterly*, *86*(3), 582–598. https://doi.org/10.1111/j.0038-4941.2005.00319.x

Smallpage, S. M., Enders, A. M., & Uscinski, J. E. (2017). The partisan contours of conspiracy theory beliefs. *Research & Politics*, *4*(4), 205316801774655. https://doi.org/10.1177/2053168017746554

Smith, N., & Graham, T. (2019). Mapping the anti-vaccination movement on Facebook. *Information, Communication & Society*, *22*(9), 1310–1327. https://doi.org/10.1080/1369118X.2017.1418406

Stasielowicz, L. (2022). Who believes in conspiracy theories? A meta-analysis on personality correlates. *Journal of Research in Personality*, *98*, 104229. https://doi.org/10.1016/j.jrp.2022.104229

Sternisko, A., Cichocka, A., & Bavel, J. J. V. (2020). The dark side of social movements: Social identity, non-conformity, and the lure of conspiracy theories. *Current Opinion in Psychology*, *35*, 1–6. https://doi.org/10.1016/j.copsyc.2020.02.007

Stieger, S., Gumhalter, N., Tran, U. S., Voracek, M., & Swami, V. (2013). Girl in the cellar: A repeated cross-sectional investigation of belief in conspiracy theories about the kidnapping of Natascha Kampusch. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00297

Stroud, N. J. (2014). *Selective exposure theories* (K. Kenski & K. H. Jamieson, Eds.; Vol. 1). Oxford University Press. https://academic.oup.com/edited-volume/34621/chapter/294951323

Strozier, C. B., Terman, D. M., Jones, J. W., & Boyd, K. A. (2010). *The fundamentalist mindset: Psychological perspectives on religion, violence, and history*. Oxford University Press.

Sunstein, C. R., & Vermeule, A. (2009). Conspiracy Theories: Causes and Cures. *Journal of Political Philosophy*, *17*(2), 202–227. https://doi.org/10.1111/j.1467-9760.2008.00325.x

Sutton, R. M., & Douglas, K. M. (2022). Rabbit hole syndrome: Inadvertent, accelerating, and entrenched commitment to conspiracy beliefs. *Current Opinion in Psychology*, *48*, 101462. https://doi.org/10.1016/j.copsyc.2022.101462

Swami, V., Barron, D., Weis, L., & Furnham, A. (2018). To Brexit or not to Brexit: The roles of Islamophobia, conspiracist beliefs, and integrated threat in voting intentions for the United Kingdom European Union membership referendum. *British Journal of Psychology*, *109*(1), 156–179. https://doi.org/10.1111/bjop.12252

Swami, V., Chamorro-Premuzic, T., & Furnham, A. (2010). Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs. *Applied Cognitive Psychology*, *24*(6), 749–761. https://doi.org/10.1002/acp.1583

Swami, V., Coles, R., Stieger, S., Pietschnig, J., Furnham, A., Rehim, S., & Voracek, M. (2011). Conspiracist ideation in britain and austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories: Conspiracist ideation. *British Journal of Psychology*, *102*(3), 443–463. https://doi.org/10.1111/j.2044-8295.2010.02004.x

Swami, V., Furnham, A., Smyth, N., Weis, L., Lay, A., & Clow, A. (2016). Putting the stress on conspiracy theories: Examining associations between psychological stress, anxiety, and belief in conspiracy theories. *Personality and Individual Differences*, *99*, 72–76. https://doi.org/10.1016/j.paid.2016.04.084

Swami, V., Voracek, M., Stieger, S., Tran, U. S., & Furnham, A. (2014). Analytic thinking reduces belief in conspiracy theories. *Cognition*, *133*(3), 572–585. https://doi.org/10.1016/j.cognition.2014.08.006

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Tangherlini, T. R., Shahsavari, S., Shahbazi, B., Ebrahimzadeh, E., & Roychowdhury, V. (2020). An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, Pizzagate and storytelling on the web. *PLOS ONE*, *15*(6), e0233879. https://doi.org/10.1371/journal.pone.0233879

Teličák, P., & Halama, P. (2021). Maladaptive Personality Traits, Religiosity and Spirituality as Predictors of Epistemically Unfounded Beliefs. *Studia Psychologica*, *63*(2), 175–189. https://doi.org/10.31577/sp.2021.02.820

Tempel, J. van der, & Alcock, J. E. (2015). Relationships between conspiracy mentality, hyperactive agency detection, and schizotypy: Supernatural forces at work? *Personality and Individual Differences*, *82*, 136–141. https://doi.org/10.1016/j.paid.2015.03.010

Theocharis, Y., Cardenal, A., Jin, S., Aalberg, T., Hopmann, D. N., Strömbäck, J., Castro, L., Esser, F., Van Aelst, P., de Vreese, C., Corbu, N., Koc-Michalska, K., Matthes, J., Schemer, C., Sheafer, T., Splendore, S., Stanyer, J., Stępińska, A., & Štětka, V. (2021). Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. *New Media & Society*, 146144482110456. https://doi.org/10.1177/14614448211045666

Thomas, S. B., & Quinn, S. C. (1991). The Tuskegee Syphilis Study, 1932 to 1972: Implications for HIV education and AIDS risk education programs in the black community. *American Journal of Public Health*, *81*(11), 1498–1505. https://doi.org/10.2105/AJPH.81.11.1498

Tian, K. T., Bearden, W. O., & Hunter, G. L. (2001). Consumers' Need for Uniqueness: Scale Development and Validation. *Journal of Consumer Research*, *28*(1), 50–66. https://doi.org/10.1086/321947

Tilley, J., & Evans, G. (2014). Ageing and generational effects on vote choice: Combining cross-sectional and panel data to estimate APC effects. *Electoral Studies*, *33*, 19–27. https://doi.org/10.1016/j.electstud.2013.06.007

Tollefson, J. (2021). Tracking QAnon: How trump turned conspiracy-theory research upside down. *Nature*. https://doi.org/10.1038/d41586-021-00257-y

Törnberg, P., & Törnberg, A. (2022). Inside a White Power echo chamber: Why fringe digital spaces are polarizing politics. *New Media & Society*, 146144482211229. https://doi.org/10.1177/14614448221122915

Turska-Kawa, A., Pilch, I., Wardawy, P., & Olszanecka-Marmola, S., Agata. (in press). Contemporary trends in psychological research on conspiracy beliefs. A systematic review. *Frontiers in Psychology*.

Tuters, M., Jokubauskaitė, E., & Bach, D. (2018). Post-Truth Protest: How 4chan Cooked Up the Pizzagate Bullshit. *M/C Journal*, *21*(3). https://doi.org/10.5204/mcj.1422

Uscinski, J., DeWitt, D., & Atkinson, M. D. (2018). A web of conspiracy? Internet and conspiracy theory. In A. Dyrendal, D. G. Robertson, & E. Asprem (Eds.), *Handbook of conspiracy theory and contemporary religion* (pp. 106–130). Brill.

Uscinski, J., Enders, A. M., Klofstad, C., & Stoler, J. (2022). Cause and effect: On the antecedents and consequences of conspiracy theory beliefs. *Current Opinion in Psychology*, *47*, 101364. https://doi.org/10.1016/j.copsyc.2022.101364

Uscinski, J., Enders, A., Diekman, A., Funchion, J., Klofstad, C., Kuebler, S., Murthi, M., Premaratne, K., Seelig, M., Verdear, D., & Wuchty, S. (2022). The psychological and political correlates of conspiracy theory beliefs. *Scientific Reports*, *12*(1), 21672. https://doi.org/10.1038/s41598-022-25617-0

Uscinski, J., Enders, A., Klofstad, C., Seelig, M., Drochon, H., Premaratne, K., & Murthi, M. (2022). Have beliefs in conspiracy theories increased over time? *PLOS ONE*, *17*(7), e0270429. https://doi.org/10.1371/journal.pone.0270429

Uscinski, J., & Parent, J. M. (2014). *American Conspiracy Theories*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199351800.001.0001

Uscinski, J., Parent, J. M., & Torres, B. (2011). Conspiracy Theories Are for Losers. *APSA 2011 Annual Meeting Paper*.

van Elk, M., & Lodder, P. (2018). Experimental Manipulations of Personal Control do Not Increase Illusory Pattern Perception. *Collabra: Psychology*, *4*(1), 19. https://doi.org/10.1525/collabra.155

van Prooijen, J.-W. (2017). Why Education Predicts Decreased Belief in Conspiracy Theories: Education and Conspiracy Beliefs. *Applied Cognitive Psychology*, *31*(1), 50–58. https://doi.org/10.1002/acp.3301

van Prooijen, J.-W. (2020). An Existential Threat Model of Conspiracy Theories. *European*

Psychologist, *25*(1), 16–25. https://doi.org/10.1027/1016-9040/a000381

van Prooijen, J.-W., & Acker, M. (2015). The Influence of Control on Belief in Conspiracy Theories: Conceptual and Applied Extensions: Control and conspiracy belief. *Applied Cognitive Psychology*, *29*(5), 753–761. https://doi.org/10.1002/acp.3161

van Prooijen, J.-W., & Douglas, K. M. (2017). Conspiracy theories as part of history: The role of societal crisis situations. *Memory Studies*, *10*(3), 323–333. https://doi.org/10.1177/1750698017701615

van Prooijen, J.-W., & Douglas, K. M. (2018). Belief in conspiracy theories: Basic principles of an emerging research domain. *European Journal of Social Psychology*, *48*(7), 897–908. https://doi.org/10.1002/ejsp.2530

van Prooijen, J.-W., Douglas, K. M., & De Inocencio, C. (2018). Connecting the dots: Illusory pattern perception predicts belief in conspiracies and the supernatural: Illusory pattern perception. *European Journal of Social Psychology*, *48*(3), 320–335. https://doi.org/10.1002/ejsp.2331

van Prooijen, J.-W., & Krouwel, A. P. (2015). Mutual suspicion at the political extremes: How ideology predicts belief in conspiracy theories. In *The psychology of conspiracy* (pp. 97–116). Routledge.

van Prooijen, J.-W., Krouwel, A. P. M., & Pollet, T. V. (2015). Political extremism predicts belief in conspiracy theories. *Social Psychological and Personality Science*, *6*(5), 570–578. https://doi.org/10.1177/1948550614567356

van Prooijen, J.-W., Ligthart, J., Rosema, S., & Xu, Y. (2022). The entertainment value of conspiracy theories. *British Journal of Psychology*, *113*(1), 25–48. https://doi.org/10.1111/bjop.12522

Vosoughi, S., Mohsenvand, M. 'Neo'., & Roy, D. (2017). Rumor Gauge: Predicting the Veracity of Rumors on Twitter. *ACM Transactions on Knowledge Discovery from Data*, *11*(4), 1–36. https://doi.org/10.1145/3070644

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

Wagner-Egger, P. (2022). *The Noises of Conspiracy: Psychology of Beliefs in Conspiracy Theories* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/fv52e

Wagner-Egger, P., Bronner, G., Delouvée, S., Dieguez, S., & Gauvrit, N. (2019). Why "Healthy Conspiracy Theories" Are (Oxy)morons. *Social Epistemology Review and Reply Collective*, *8*(3), 50–67.

Wagner-Egger, P., Delouvée, S., Gauvrit, N., & Dieguez, S. (2018). Creationism and conspiracism share a common teleological bias. *Current Biology*, *28*(16), R867–R868. https://doi.org/10.1016/j.cub.2018.06.072

Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., Berelowitz, M., Dhillon, A. P., Thomson, M. A., Harvey, P., Valentine, A., Davies, S. E., & Walker-Smith, J. A. (1998). RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, *351*(9103), 637–641. https://doi.org/10.1016/S0140-6736(97)11096-0

Wal, R. C. van der, Sutton, R. M., Lange, J., & Braga, J. P. N. (2018). Suspicious binds: Conspiracy thinking and tenuous perceptions of causal connections between co-occurring and spuriously correlated events. *European Journal of Social Psychology*, *48*(7), 970–989.

https://doi.org/10.1002/ejsp.2507

Walker, J. (2013). *The United States of paranoia: A conspiracy theory* (First edition). Harper.

Wang, H., & van Prooijen, J.-W. (2022). Stolen elections: How conspiracy beliefs during the 2020 American presidential elections changed over time. *Applied Cognitive Psychology*, acp.3996. https://doi.org/10.1002/acp.3996

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*(6), 1049–1062. https://doi.org/10.1037/0022-3514.67.6.1049

West, H. G. (2003). *Transparency and Conspiracy: Ethnographies of Suspicion in the New World Order* (T. Sanders, Ed.). Duke University Press. https://doi.org/10.1215/9780822384854

Westerwick, A., Johnson, B. K., & Knobloch-Westerwick, S. (2017). Confirmation biases in selective exposure to political online information: Source bias vs. Content bias. *Communication Monographs*, *84*(3), 343–364. https://doi.org/10.1080/03637751.2016.1272761

Wood, M. J. (2018). Propagating and Debunking Conspiracy Theories on Twitter During the 2015–2016 Zika Virus Outbreak. *Cyberpsychology, Behavior, and Social Networking*, *21*(8), 485–490. https://doi.org/10.1089/cyber.2017.0669

Wood, M. J., & Douglas, K. M. (2013). What about building 7?" A social psychological study of online discussion of 9/11 conspiracy theories. *Frontiers in Psychology*, *4*(JUL), 1–9. https://doi.org/10.3389/fpsyg.2013.00409

Wood, M. J., & Douglas, K. M. (2015). Online communication as a window to conspiracist worldviews. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00836

Wood, M. J., Douglas, K. M., & Sutton, R. M. (2012). Dead and alive: Beliefs in contradictory conspiracy theories. *Social Psychological and Personality Science*, *3*(6), 767–773. https://doi.org/10.1177/1948550611434786

Xu, Z., Ellis, L., & Umphrey, L. R. (2019). The Easier the Better? Comparing the Readability and Engagement of Online Pro- and Anti-Vaccination Articles. *Health Education & Behavior*, *46*(5), 790–797. https://doi.org/10.1177/1090198119853614

Xu, Z., & Guo, H. (2018). Using Text Mining to Compare Online Pro- and Anti-Vaccine Headlines: Word Usage, Sentiments, and Online Popularity. *Communication Studies*, *69*(1), 103–122. https://doi.org/10.1080/10510974.2017.1414068

Yelbuz, B. E., Madan, E., & Alper, S. (2022). Reflective thinking predicts lower conspiracy beliefs: A meta-analysis. *Judgment and Decision Making*, *17*(4), 720–744. https://doi.org/10.1017/S1930297500008913

Zander-Schellenberg, T., Kuhn, S. A. K., Möller, J., Meyer, A. H., Huber, C., Lieb, R., & Andreou, C. (2021). Is intuition allied with jumping to conclusions in decision-making? An intensive longitudinal study in patients with delusions and in non-clinical individuals. *PLOS ONE*, *16*(12), e0261296. https://doi.org/10.1371/journal.pone.0261296

Zannettou, S., Bradlyn, B., Cristofaro, E. D., Kwak, H., Sirivianos, M., Stringini, G., & Blackburn, J. (2018). What is Gab. *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*, 1007–1014. https://doi.org/10.1145/3184558.3191531

Zarefsky, D. (2014). Conspiracy Arguments in the Lincoln-Douglas Debates. In

Argumentation Library (pp. 195–209). Springer Nature. https://doi.org/10.1007/978-3-319-05485-8_16

Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, *395*(10225), 676. https://doi.org/10.1016/S0140-6736(20)30461-X

Zhang, Y., Wang, L., Zhu, J. J. H., & Wang, X. (2021). Conspiracy vs science: A large-scale analysis of online discussion cascades. *World Wide Web*, *24*(2), 585–606. https://doi.org/10.1007/s11280-021-00862-x

Zmigrod, L., Eisenberg, I. W., Bissett, P. G., Robbins, T. W., & Poldrack, R. A. (2021). The cognitive and perceptual correlates of ideological attitudes: A data-driven approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1822), 20200424. https://doi.org/10.1098/rstb.2020.0424

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2018). Cognitive underpinnings of nationalistic ideology in the context of Brexit. *Proceedings of the National Academy of Sciences*, *115*(19). https://doi.org/10.1073/pnas.1708960115

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

Zubiaga, A., Liakata, M., Procter, R., Hoi, G. W. S., & Tolmie, P. (2016). Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. *PLOS ONE*, *11*(3), e0150989. https://doi.org/10.1371/journal.pone.0150989

5 Appendix

5.1 Study 1

Behavior Research Methods
https://doi.org/10.3758/s13428-021-01698-z

LOCO: The 88-million-word language of conspiracy corpus

Alessandro Miani[1] · Thomas Hills[2,3] · Adrian Bangerter[1]

Accepted: 26 August 2021

Abstract

The spread of online conspiracy theories represents a serious threat to society. To understand the content of conspiracies, here we present the language of conspiracy (LOCO) corpus. LOCO is an 88-million-token corpus composed of topic-matched conspiracy (N = 23,937) and mainstream (N = 72,806) documents harvested from 150 websites. Mimicking internet user behavior, documents were identified using Google by crossing a set of seed phrases with a set of websites. LOCO is hierarchically structured, meaning that each document is cross-nested within websites (N = 150) and topics (N = 600, on three different resolutions). A rich set of linguistic features (N = 287) and metadata includes upload date, measures of social media engagement, measures of website popularity, size, and traffic, as well as political bias and factual reporting annotations. We explored LOCO's features from different perspectives showing that documents track important societal events through time (e.g., Princess Diana's death, Sandy Hook school shooting, coronavirus outbreaks), while patterns of lexical features (e.g., deception, power, dominance) overlap with those extracted from online social media communities dedicated to conspiracy theories. By computing within-subcorpus cosine similarity, we derived a subset of the most representative conspiracy documents (N = 4,227), which, compared to other conspiracy documents, display prototypical and exaggerated conspiratorial language and are more frequently shared on Facebook. We also show that conspiracy website users navigate to websites via more direct means than mainstream users, suggesting confirmation bias. LOCO and related datasets are freely available at https://osf.io/snpcg/.

Keywords ■■■

Introduction

Conspiracy theories (CTs) are narratives that attempt to explain significant social events as being secretly plotted by powerful and malicious elites at the expense of an unwitting population (Douglas et al., 2019; Samory & Mitra, 2018b). Belief in CTs is widespread. In 2013, it was estimated that over 50% of the US population believed in at least one CT (Oliver & Wood, 2014), while in 2020, in the middle of the COVID-19 pandemic, health-related misinformation attracted four times as much traffic as official health sources on social media (AVAAZ, 2020). The consequences associated with the circulation of such theories are not trivial, potentially leading to detrimental social action (Franks et al., 2013; Imhoff et al., 2021; Sternisko et al., 2020). Belief in CTs is linked to rejection of official information and science (Raab, Auer, et al., 2013a; Raab, Ortlieb, et al., 2013b; van der Linden, 2015), decreased intentions to adopt vaccines (Jolley & Douglas, 2014b; Lazarus et al., 2020; Salmon et al., 2005), resistance to COVID-19 containment measures and vaccination (Biddlestone et al., 2020; Lazarus et al., 2020), and reduced protection against sexually transmitted diseases (Bogart et al., 2010). CT belief is also related to general distrust and political alienation along with endorsement of nonnormative (vs. normative) political intentions (Einstein & Glick, 2015; Imhoff et al., 2021; Jolley & Douglas, 2014a). Such beliefs also provide justification for engaging in everyday crime (Jolley et al., 2019; Jolley & Paterson, 2020) and anti-Semitic and Islamophobic attitudes (Golec de Zavala & Cichocka, 2012; Swami et al., 2018). Therefore, within psychology, research has typically focused on motivational and contextual factors as well as individual differences underlying belief in CTs (Butter & Knight, 2020; Douglas et al., 2019; Douglas & Sutton, 2018).

✉ Alessandro Miani
alessandro.miani@unine.ch

1 Institute of Work and Organizational Psychology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

2 Department of Psychology, University of Warwick, University Road, Coventry CV47AL, UK

3 The Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, UK

Published online: 25 October 2021

A more complete understanding of CTs requires understanding how they spread. The current focus on individual beliefs, predispositions, and biases is of limited utility in this respect, for two reasons. First, beliefs are not straightforwardly connected to CT transmission. For example, a skeptical-minded individual may share a CT document within a debunking community for critical purposes (Franks et al., 2017), or a credulous individual may hesitate to share such a document in a science-oriented community for fear of being stigmatized (Lantian et al., 2018). Moreover, transmission can also be motivated strategically, independently of belief, to influence constituencies, such as when CTs or fake news is intentionally shared on social media to affect outcomes like voter behavior (Bangerter et al., 2020; Douglas et al., 2019). Second, CT beliefs do not spread per se. Rather, CTs spread as materialized forms of belief, conveyed as narratives in the form of written text (e.g., from webpages or social media posts), video (e.g., from video-sharing platforms such as YouTube), images (e.g., internet memes), or eventually audio (e.g., podcasts or the recent audio-based social media Clubhouse). Regardless of their form, CT beliefs emerge in the minds of recipients when they interact with such content. For whatever reasons CT narratives are created, they circulate, sticking in the mind of conspiracy-predisposed recipients and potentially motivating individual and collective action (Franks et al., 2013; Imhoff et al., 2021; Jolley & Paterson, 2020). Therefore, to understand the spread of CTs and their outcomes, research should investigate the content of CT narratives.

On the internet, misinformation spreads faster, farther, and deeper within groups of like-minded individuals (Del Vicario et al., 2016a, b; Vosoughi et al., 2018, but see Clarke, 2007; Uscinski et al., 2018 for a critical view). The internet constitutes a system of information proliferation by which many people form opinions in regard to political parties, social issues, and health-related information (Betsch et al., 2011). In the Web 2.0 version of the internet, information is produced and consumed in a horizontal fashion, allowing anyone to create and share content, with few editorial filters (Aupers, 2012; Bessi et al., 2015a, b). CT texts may thus have the same epistemological weight for many users as mainstream texts, and compete with them for attention (Bessi et al., 2014; Eicher & Bangerter, 2015; Hills, 2019). The perceived credibility of epistemic sources is also a function of belief in CTs (Imhoff et al., 2018). This makes epistemic authority difficult to evaluate, especially when conspiratorial narratives are promoted by political leaders (Barkun, 2017), by scholars in prestigious journals (Wakefield et al., 1998), and by Nobel Prize winners (Perez & Montagnier, 2020).

Research on the content and circulation of CTs on the internet has focused on user-generated texts such as comments and posts gathered from social media such as Twitter (Mitra et al., 2016; Wood, 2018), Facebook (Bessi, 2016; Bessi, Zollo, et al., 2015b; Brugnoli et al., 2019; Smith & Graham, 2019), Reddit (Klein et al., 2018, 2019; Samory & Mitra, 2018a, 2018b), Gab (Zannettou et al., 2018), or comment sections of news websites (Wood & Douglas, 2013, 2015). This approach has the advantage of exploring large, ecologically valid samples of text as a complement to psychological investigations of CT beliefs. However, it is difficult to reliably extract measures of individual belief from comments embedded within the noisy and heterogeneous discussion threads of conspiracy believers and debunkers (Wood & Douglas, 2013, 2015; but see Klein et al., 2019). Moreover, discussion threads limit the utility of extracted text because the comments and posts are brief, are contextualized in the discussion in which they are embedded, and are incapable of spreading independently from the whole thread. As a matter of fact, discussion threads are not conspiracy narratives per se. While the comments and posts they contain might instill curiosity, reinforce existing beliefs, or support conversion, they often do not constitute the actual source through which CTs are transmitted.

Towards corpora of CT texts

One valuable source of CTs for academic research aimed at understanding the content and transmission of CTs is CT websites. Although social media sites engage more traffic and are overall more popular than other websites (Facebook, Twitter, and Instagram are respectively ranked as the third, fourth, and fifth most popular websites following Google and YouTube, according to similarweb.com, accessed on 20 March 2021), websites provide more in-depth and elaborated discourse than posts and tweets on social media, which are nevertheless crucial for the spread of webpages. Conspiracy websites, specifically, are specialized sources created for the purpose of developing, collecting, and spreading CTs. These websites provide ample page space to showcase arguments that discredit official narratives and function as trustworthy epistemic sources for CT believers. Analysis of CT webpages offers a series of advantages. Webpages constitute standalone, structured texts nested within sources (i.e., websites) and as such are accompanied by paratextual (i.e., metadata) information. As standalone documents, webpages can easily be shared on social media and so can provide measures of spread. Appropriately identified webpages, therefore, would be beneficial for studying the content of CTs, and a large corpus of such CTs could provide a solid grounding for CT research.

Only a handful of studies, focused on related phenomena such as anti-vaccine movements, rumors, and fake news, have built corpora from online material (discussed below). Yet, to our knowledge, the field lacks a corpus specifically focused on online CTs. In this section, we describe published work, stressing its strengths and weaknesses.

General-purpose linguistic corpora such as the WaCky corpus (Baroni et al., 2009) and the British National Corpus

(BNC, Aston & Burnard, 1998), although composed of large collections of texts, do not generally allow researchers to focus on either the source type (e.g., conspiracy vs. mainstream) or a specific topic (e.g., an event that has generated a CT). It should be noted, however, that documents in WaCky were gathered from a set of 2000 seeds consisting of randomly chosen pairs of content words selected from the BNC, meaning that seeds were used as keywords to retrieve webpages. This represents a useful approach which we use here: by creating ad hoc seeds, data collection can be directed to a particular (set of) topic(s) encapsulated in the seed.

Other corpora focus on specific themes. In the field of online anti-vaccine movements, a few studies have collected webpages gathered through search engines (Fu et al., 2016; Okuhara et al., 2017; Sak et al., 2015). This approach is convenient, as it allows researchers to obtain data by mimicking how users retrieve information. However, because these corpora were collected manually, sample sizes are limited, therefore reducing generalizability. To a different degree, the CORPS corpus (Guerini et al., 2013) is composed of 3600 political speeches gathered from the web. Nonetheless, being composed of only one genre without any (matched) control group, CORPS does not enable comparisons beyond descriptive analyses.

Focusing on rumors and fake news, other studies have built corpora that include a control group that allows between-group comparisons (Castelo et al., 2019; Kwon et al., 2017). Kwon et al. (2017), for example, built two Twitter subcorpora from a list of rumor and non-rumor events (henceforth RumTweet). Castelo et al. (2019), on the other hand, collected material (fake news vs. mainstream) via lists of reliable and unreliable websites compiled by independent fact checkers (henceforth FNweb). This approach is useful because it reduces selection bias during data collection. However, although these two studies allow us to narrow the sampling method to obtain either two-group sources (websites) or two-group themes (rumors), they present an important limitation. In fact, to systematically study phenomena related to language (e.g., conspiracy, or fake news, or rumors) through a corpus, many forms of analysis are likely to require subcorpora matched by topic. This allows researchers to compare different versions of the same event to identify discriminating features. If not matched, the two subcorpora are treated as bags of words (as in Castelo et al., 2019, and Kwon et al., 2017), ignoring the inherent structure emerging from different themes or sources. Although some changes are expected to emerge systematically from a bag-of-words approach, there may also be important topic-specific differences. For example, differences between CT and mainstream accounts of Princess Diana's death are likely to differ in informative ways from CT and mainstream accounts of COVID-19. These differences can only emerge from a topic-matched corpus.

Overcoming this limitation, the PHEME dataset (Zubiaga et al., 2016) focuses on predefined events that have generated rumors, allowing researchers to compare rumor with non-rumor tweets around specific events. Yet, the only work, to our knowledge, that has focused on CTs is that of Uscinski and collaborators (2011), who gathered 100,000 published letters to the editor of *The New York Times* (henceforth NYT) from 1897 to 2010. After the data were collected, each document was manually coded as either referring to a conspiracy or not, of which 800 were identified as conspiracies. In addition, differently from the other works we reviewed above, the authors coded several groups of actors post hoc (e.g., left/right/foreign political actors, capitalists/communists, media, government institutions, and other).

Here, we present LOCO,[1] our Language Of COnspiracy corpus that was built upon the strengths and weaknesses of the reviewed corpora.

Table 1 shows the corpora comparison, including LOCO, and summarizes each corpus' key features, including the focus (i.e., the goal, e.g., general-purpose language or fake news), the source of the material gathered (e.g., from webpages or twitter), size expressed in number of documents and tokens (i.e., non-unique words), presence of topics (e.g., events or themes) or grouping (e.g., rumors vs. non-rumors) structures, the date range of documents expressed in years, and whether the material is freely available.

The language of conspiracy (LOCO) corpus

LOCO is a multilevel topic-matched corpus composed of standalone documents (N = 96,743) gathered via ready-made lists of conspiracy and mainstream websites (see LOCO's key feature in Table 2). LOCO has been built as a freely available text source from which researchers can extract features and/or generate predictive and classification models. Previous studies of CT textual data have extracted lexical features (Del Vicario, Vivaldo, et al., 2016b; Faasse et al., 2016; Klein et al., 2019; Mitra et al., 2016; Samory & Mitra, 2018a; Wood & Douglas, 2015), topic distributions (Bessi, Zollo, et al., 2015b; Klein et al., 2018; Mitra et al., 2016; Samory & Mitra, 2018b), and narrative patterns (Samory & Mitra, 2018b). Such analyses can be replicated and extended with LOCO due to its rich metadata.

[1] The acronym LOCO might suggest the idea that conspiracy theories and theorists are all crazy. Far from this position, we rather highlight the polarizing phenomenon by which, regardless of the belief position, the "others" are considered crazy. People with beliefs in CTs feel, and often are, stigmatized (Lantian et al., 2018). On the other hand, nonbelievers are also in some instances mocked as "globetards," "vaxholes," or "covidiots." The last expression is emblematic as it is used by both sides of the belief spectrum to refer to people who either believe or do not believe in COVID-19. And, of course, some conspiracy theories are true.

Table 1 Key features of eight corpora relevant to conspiracy theory content

Resource	BNC	WaCky	CORPS	FNweb	RumTweet	PHEME	NYT	LOCO
Focus	Language	language	Political speeches	Fake news	Rumors	Rumors	Conspiracy	Conspiracy
Obtained from	Printed material	Web pages	Web pages	Webpages from list of websites	Twitter	Twitter	Newspaper	Webpages from list of websites
Number of documents	4 K	2.69 M	3.6 K	14 K (7 K fake)	192 K tweets (61 K rumor)	7.5 K threads (35 K rumor tweets)	100 K (800 conspir-acy)	96 K (24 K conspiracy)
Number of tokens	100 M	1.9 B	7.9 M	7 M*	2.8 M*	100 K*		88 M
Topic structure	NO	2 K seeds	NO	NO	YES 111 events (60 rumors, 51 non-rumors)	YES 9 events	YES	YES 47 seeds 600 topics
Grouping structure	NO	NO	NO	YES	YES	YES (matched)	YES	YES (matched)
Year range			1917 2010	2013 2018	2006 2009	Events around 2014–2015	1897 2010	1853 2020
Freely available	YES	YES	YES	YES	YES	YES	NO	YES

Note. Resources: BNC (Aston & Burnard, 1998); WaCky (Baroni et al., 2009); CORPS (Guerini et al., 2013); FNweb (Castelo et al., 2019); RumTweet (Kwon et al., 2017); PHEME (Zubiaga et al., 2016); NYT (Uscinski et al., 2011). *Number of tokens calculated from studies' freely available datasets

The main goal of LOCO is to shed light on the language of conspiracy. To this aim, LOCO is built on documents that revolve around CTs. Because we do not yet know what the language of conspiracy is, i.e., to what extent conspiracy language differs from non-conspiracy language, selecting documents (e.g., from webpages) based on an a priori definition of conspiracy would be difficult. At best, selecting documents based on their content would result in both a limited sample size (due to manual coding, see e.g. Fu et al., 2016; Okuhara et al., 2017; Sak et al., 2015) and limited heterogeneity (due to selection criteria based on a specific linguistic/rhetoric style). We therefore chose to categorize document selection starting from the source (i.e., websites).

Not all content from conspiracy websites will contain CTs. Intuitively, it is unlikely that all ~93,000 webpages in www.globalresearch.ca contain CTs, and some content might come from neighboring genres such as rumors, fake news, urban legends, and pseudoscience. To provide an estimate of how well conspiracy and mainstream documents reflect their true labels and how well the two sources can be distinguished from each other, we have blindly coded a subset of LOCO's documents (60 documents from each subcorpus, see Section SM1 in supplemental material) as being either conspiratorial or not. With an overall accuracy of .88 (Cohen's k = .77), we have correctly classified as conspiracy 85% of documents and correctly classified as

Table 2 Summary statistics of mainstream, conspiracy, and all documents in LOCO

	Mainstream	Conspiracy	Whole corpus
No. of documents	72,806	23,937	96,743
No. of websites	92	58	150
Range of years	1853–2020	2004–2020	1853–2020
	M (*SD*) [range]	*M* (*SD*) [range]	*M* (*SD*) [range]
No. of words per document	805.94 (939) [97–9507]	1236.32 (1307) [100–9428]	912.43 (1059) [97–9507]
Total no. of words	*58,677,322*	*29,593,678*	*88,271,000*
No. of sentences per document	37.92 (47.89) [1–1087]	59.63 (69.58) [1–1047]	43.29 (54.88) [1–1087]
Total no. of sentences	*2,760,789*	*1,427,397*	*4,188,186*
No. of paragraphs per document	16.56 (19.30) [1–829]	24.51 (32.83) [1–905]	18.53 (23.64) [1–905]
Total no. of paragraphs	*1,205,904*	*586,748*	*1,792,652*

mainstream 92% of documents. The lower classification performance on conspiracy documents suggests that not all documents from conspiracy websites are in fact CTs, while mainstream documents are less ambiguously classified as non-conspiracy. An alternative explanation is that conspiracy texts are difficult to distinguish from mainstream texts, at least via human inspection (meaning that future algorithms might find features that help improve on human classification). These results also suggest that conspiracy and mainstream texts overlap to some extent (suggesting a continuum).

The multilevel structure of LOCO allows us to take into consideration natural hierarchical grouping of documents cross-nested within websites and topics. At the document, webpage, and website levels, LOCO's metadata[2] allow researchers to create subsets of documents or to add covariates during analyses. In Table 3, we summarized the key variable types we provide with LOCO for each level. For example, each document is associated with topic labels that summarize its semantic content. These labels refer to the topics that have the highest probability (among all topics extracted from LOCO) of describing the document's content (see "Topic extraction" section). This is useful for tracking differences (e.g., in lexical features) between conspiracy and mainstream texts within a specific topic (e.g., Princess Diana's death), within a set of related topics (e.g., coronavirus outbreak in China, coronavirus outbreak in the United States), between topics (e.g., Pizzagate vs. moon landing), or within and between topics, e.g., by using a 2 (e.g., Princess Diana, coronavirus) × 2 (conspiracy, mainstream) factorial design. Similar analyses can be performed using the data LOCO provides on website information about political bias, factual reporting, and website category. For most (~67%) webpages, we gathered information about their upload/creation date (see "Date" section). This allows researchers to test time-related hypotheses such as the evolution through time of topics or lexical features (e.g., coronavirus topics over time). Other crucial features of LOCO are the spread and popularity metrics associated with both websites and webpages. These metrics allow researchers to test hypotheses about social media transmission, for example, testing webpages' spread and engagement while correcting for the website's popularity. Last but not least, LOCO is provided with a set of almost 300 lexical features (e.g., psychological processes associated with words) derived from two widely used and validated text-analysis programs based on word-within-category counting.

[2] Note that we make a distinction between documents, webpages, and websites' metadata. For document, we refer to the text and its intrinsic features such as title, topic, lexical features, etc. Differently, for webpage, we refer to a set of paratextual information related to the webpage (that contains the text) such as the URL, date, spread, and the website host. Websites' metadata refer to the second level of paratextual information such as website's political bias, size, and popularity.

Method

Seed selection

Similar to the construction of the WaCky corpus (Baroni et al., 2009), we used seeds (i.e., keywords) to retrieve the webpages that provide the texts for LOCO. Seeds were extracted from the items of two CT-based surveys: a national poll (Jensen, 2013, Source 1, e.g., "*Do you believe that Lee Harvey Oswald acted alone in killing President Kennedy, or was there some larger conspiracy at work?*") and the 17-item "endorsement of conspiracy theories" from Douglas and Sutton (2011, Source 2, e.g., "*The American moon landings were faked*"). We extracted the seeds from these surveys for two reasons. Firstly, these surveys on CTs encompass a broad set of well-known CTs, since they are supposed to measure specific beliefs from a wide range of people. Secondly, these surveys condense each theory within a short space, usually a sentence. These two surveys were chosen because, while they measure specific theories, they are broad in scope, and encompass a large and heterogeneous set of CTs. Items from both surveys were grouped to obtain a unique seed (e.g., "*Princess Diana faked her own death so she and Dodi could retreat into isolation*," "*Princess Diana's death was an accident*," and "*One or more rogue 'cells' in the British Secret Service constructed and carried out a plot to kill Princess Diana*" were merged as "Princess Diana's death").

We further broadened the pool of seeds by manually adding 20 seeds corresponding to popular (e.g., Illuminati, genetically modified organisms, Pizzagate) and current (e.g., coronavirus, Bill Gates, 5G) CTs missing from Sources 1 and 2. Note that seeds such as "chemtrails," when applied to mainstream documents, in most if not all cases return documents referring to CTs. We keep these documents in LOCO so as to have a broad mainstream pool and allow users to create subsets of texts prior to analyses (e.g., by removing mainstream documents that mention CTs, see "Mentioning "Conspiracy" and Effect of mentioning conspiracy" sections). In order to include events that might be associated with different spellings, for some seeds we used synonyms (e.g., big pharma, drug companies, and pharmaceutical industry; new world order and NWO; climate change and global warming). In Table 4, we show the full set of seeds used to retrieve documents and the final document count in LOCO by source type. Note that the seed count is larger than the number of documents. This is because a single webpage can be returned by a Google search using different keywords. For example, if a document relates to Princess Diana's death due to an Illuminati plot, then this document would be returned twice for both "Princess diana death" and "illuminati" searches.

Note that although we used seeds as keywords to retrieve webpages, we do not intend seeds to serve as proxies for document content. This is because a webpage is returned by

Table 3 Types of variables included in LOCO

Level	Variable type	Example of variable	Section
1. Document	Raw content	Document ID	Table 6
		Title	3.4
		Text	3.4
	Features	Number of words, sentences, paragraphs	3.8.2
	Semantic content	Topic	3.6
		Lexical features	3.5
	Conspiracy content	Representativeness	3.7
		Mention of conspiracy	3.8.1
2. Webpage	Information	Website host	3.2
		URL	3.3
		Date	3.8.3
		Seeds	3.1
	Spread	Facebook shares, comments, and reactions	3.8.4
3. Website	Classification	Political orientation, factual reporting, category	3.8.5
	Size	Number of webpages	3.8.6
	Popularity	Visits, traffic, and rank	3.8.6
	Spread	Facebook shares, comments, and reactions	3.8.4

Google if the seed is present in the webpage (but note, not necessarily in the main text) at least once. The seed presence in the webpage, however, does not necessarily indicate that the seed reflects the main topic of the document's text, because the seed can be contained in boilerplate texts or in the comments section of the webpage. Instead, we remind the user that for a more precise content of documents, we offer a more fine-grained measure of document content (extracted from the cleaned text), namely topics (see "Topic extraction" section). We include the seed variable in the LOCO dataset, believing it might be useful for answering other questions, e.g., regarding webpage indexing.

Website lists

Following previous work (Pennycook & Rand, 2019), we gathered a list of conspiracy websites from mediabiasfactcheck (MBFC).[3] Websites are labeled by MBFC as conspiracy if they publish unverifiable information related to known conspiracies such as the New World Order, Illuminati, false flags, aliens, anti-vaccination propaganda, etc. (for further details, see category descriptions in "Website category" section). From the whole list of 241 conspiracy websites, we selected (in December 2019) those that scored the highest on the conspiracy rating (i.e., "tin foil hat," N = 68[4]). This increased the chances of obtaining highly conspiratorial texts, limiting contamination by mainstream or less conspiratorial texts.

The mainstream list of websites was created (in June 2020) in a data-driven fashion by extracting the websites returned by Google for each seed. While maximizing data acquisition, this approach also mimics users' online behavior. We proceeded as follows. For each seed, we created a Google query, gathered the resulting top 40 URLs, and extracted the websites' domains.[5] We repeated this operation with different IPs, mimicking the searches from the UK (London), USA (New Jersey), and Australia (Melbourne) to maximize English language domains as well as the heterogeneity of websites. This procedure returned a total of 1453 unique domains. All domain counts were aggregated, and we computed two popularity metrics per domain: (1) the number of times a domain appears overall for all seeds (absolute frequency), and (2) the number of unique seeds associated with a specific domain (relative frequency). These two metrics were chosen to obtain a large portion of pages (absolute method) and a wide coverage of seeds (relative method). The top 120 domains for each metric were visually inspected to remove potential conspiracy

[3] https://mediabiasfactcheck.com/conspiracy/

[4] Note that the final number of conspiracy websites in LOCO is 58. This is because during the data cleaning process for some websites we did not obtain any webpages (e.g., stormfront.org, learntherisk.org). Other websites were excluded because they were either collections of tweets and videos or were CT search engines (e.g., qanon.pub, disclose.tv, and alternativenews.com).

[5] E.g., telegraph.co.uk from the URL https://www.telegraph.co.uk/news/uknews/1577644/MMR-vaccine-doesnt-cause-autism-says-study.html.

Table 4 List of seeds

seed source	No. of conspiracy documents	No. of mainstream documents	
5g	m	702	1664
aids	2	1025	2428
alien	1, 2	813	1715
barack obama	1	496	1485
big foot	1	708	2019
big pharma	1	716	1758
bill gates	m	717	1623
cancer	m	839	2098
chemtrails	1	744	549
cia cocaine	1	552	1030
climate change	1, 2	889	2166
coronavirus	m	1104	2588
covid 19	m	1004	2395
drug companies	1	1024	2356
ebola	m	626	2140
elvis death	m	188	1386
elvis presley	m	132	1258
flat earth	m	605	1646
fluoride water	1	395	1384
george bush	1	844	1737
george soros	m	735	1178
global warming	1, 2	896	1793
gmo	m	620	1924
illuminati	m	804	1479
jfk assassination	1, 2	607	1344
jonestown suicide	2	42	594
mh370	m	167	1086
michael jackson death	m	616	1564
mind control	1	949	2036
moon landing	1, 2	349	1579
new world order	1	1036	2162
nwo	1	814	1350
osama bin laden	1	645	1415
paul mccartney death	1	149	1190
pharmaceutical industry	1	828	1684
pizzagate	m	359	1012
planned parenthood	m	626	1434
population control	m	972	2295
princess diana death	2	309	1338
reptilian	1	494	1418
saddam hussein	1	677	1623
sandy hook	m	470	1500
september 11 attack	1, 2	939	2207
vaccine	1	803	2125
vaccine autism	1	531	1654
vaccine covid	m	923	2031
zika virus	m	473	1675

Note. Sources 1, 2, and m refer to: 1 = Jensen (2013); 2 = Douglas and Sutton (2011), and m = manual websites (none appeared), less relevant websites such as those not related to text content (YouTube, Amazon, Instagram, Pinterest, LinkedIn, Shutterstock), websites with user-generated content (Blogger, Facebook, Twitter), and other websites such as those related to movie reviews, private companies, and online courses. Following these exclusion criteria, a total of 19 domains were removed. Keeping all domains appearing in both metrics (N = 135), this list was visually inspected and subdomains were aggregated (e.g., keith.seas.harvard.edu, sitn.hms.harvard.edu, health.harvard.edu, hsph.harvard.edu aggregated to harvard.edu) while removing mistakenly extracted domains (e.g., www) and non-English domain suffixes (e.g., nationalgeographic.fr). This left us with 93 domains.[6]

URL extraction and cleaning

Once we had obtained the list of seeds and the two lists of websites, we proceeded with collecting the webpages' URLs through Google. Besides being the most popular search engine (ranked # 1 worldwide according to www.similarweb.com, accessed September 2020), we used Google Search because we were interested in mimicking user behavior. Importantly, while allowing us to automate URL extraction, this procedure also uses the same search criteria for all websites, without relying on website-specific search engines that might have biased results (e.g., by using the search bar within the website).

URL scraping was performed in R (R Core Team, 2019), using the *curl* package (Ooms, 2019). We formed Google queries by crossing each seed with each website to search for a specific seed within a specific website. For example, the Google query *site:bbc.com moon landing*[7] returned results about moon landing from the BBC website. The UK top-level domain "google.co.uk" was chosen over "google.com" to ensure English language searches (".com" in Switzerland—where the study was conducted—automatically returns results in either German, French, or Italian). We also prompted Google to extract results in the English language by adding "hl=en" to the query. For each query, we extracted the first 60 results. Data collection occurred between May 20th and July 4th, 2020 (see workflow in SM2).

Once the URL collection was complete ($N_{conspiracy}$ = 67,813; $N_{mainstream}$ = 163,488), we proceeded with removing duplicated and non-relevant URLs. This was performed by searching (with regular expressions) and removing the URLs that did not include the website searched, non-text files (pdf, pictures, videos), video and photo galleries, feeds, forums, and blogs, dynamic pages (e.g., URL ending with "php," "?"), collection pages and archives of links, shops and stores, and Wikipedia lists and discussions. This procedure left us with 29,885 conspiracy and 105,461 mainstream documents.

Text extraction and cleaning

To extract the HTML files and then the useful text from our list of URLs, we tested several Python packages. These scripts, called "boilerplate stripping," remove noise text from webpages such as navigation links, header and footer sections, etc. The Python *Goose* package returned the best performance (see SM3) and therefore was chosen for extracting the texts. Importantly, *Goose* can be set to return a series of meta-descriptions and tags from the raw HTML file. Therefore, along with the main body of the text, we used *Goose* to extract the title of the document, the language tag (further capturing non-English pages), and the date the file was uploaded on the website or created (see discussion in "Date" section).

Once all the texts were collected, we further cleaned the raw corpus using the following exclusion criteria: documents for which the HTML meta-tag language was not set as English, empty documents, exact duplicated texts, and texts shorter than 100 words.[8] In order to further remove non-English documents that did not contain the language HTML tag, we removed texts in which the percentage of top 1000 English words (Fry, 2000) was below 40% (threshold chosen after visual inspection). Finally, we also removed texts whose word count was 2.5 standard deviations above the mean of the whole corpus. This procedure left us with the final LOCO sample of 23,937 conspiracy and 72,806 mainstream documents (see Table 2 for details).

Lexical feature extraction

For each document in LOCO, we extracted measures of language use with two word-counting tools, namely LIWC (Linguistic Inquiry and Word Count, see Tausczik & Pennebaker, 2010) and Empath (Fast et al., 2016). Both tools have been used previously to investigate the language of conspiracy on social media (Fong et al., 2021; Klein et al., 2019). These tools work on the same principle: they analyze texts, word by word, and check whether the word is included in a predefined category; if so, the category value increases. To extract LIWC categories, we used the LIWC standalone application (version 2015), while for Empath we relied on CLA (Custom List Analyzer version 1.1.1, see Kyle et al., 2015), a standalone application that, along with the batch of texts, takes

[6] Note that this number is different from the final N = 92 for mainstream websites. This is because after the cleaning section, the website urbandictionary.com was no longer present.

[7] URL: https://www.google.co.uk/search?q=site%3Abbc.com+moon+landing&hl=en

[8] The discrepancy with Table 2, which shows the minimum word count as 97 words in a document, is due to the fact that at this stage (document cleaning) we counted words as portions of text separated by spaces, while LOCO's final word count was performed with TAACO (see "Text statistics" section).

as input an ad hoc list of dictionaries. Both tools provide standardized outputs, that is, the number of words in a given category divided by the total number of words from the text file. Note that the two tools provide different formats for their output: while LIWC returns percentages (range: 0–100), Empath returns ratios (range: 0–1).

Although these tools work on the same principle, they differ in how they were built, making them somewhat complementary. First, unlike Empath, LIWC detects grammatical categories such as articles, prepositions, pronouns, etc. Second, while LIWC construction relied on human coding, Empath categories were built in a data-driven fashion from a semantic database. For instance, by seeding terms such as "facebook" and "twitter," Empath generates the category labeled "social media." The two methods by which these tools were built explain why they compute slightly different values along their categories, as shown in between-dictionary correlations (see Section SM4).

Topic extraction

For each document in LOCO, we quantify the semantic content by providing a fine-grained topical distribution. This represents a vector containing the probabilities that each of a series of topics is associated with each document. This was achieved with Latent Dirichlet Allocation, (LDA; Blei et al., 2003, see SM5 for text preprocessing). LDA is an unsupervised probabilistic machine learning model capable of identifying co-occurring word patterns and extracting the underlying topic distribution for each text document. By setting a priori the number of topics in a given corpus, LDA computes, for each document in the corpus, the probabilities for all topics of being represented in the document. Meanwhile, each word of the corpus has a probability of being part of a topic. In other words, a word x has probability β of being part of topic k; a topic k has probability γ of being part of document n. The sum of all the word probabilities within one topic is 1, and the sum of all the topic probabilities within one document is 1.

In LDA, the "right" number of topics is determined by the goal of the task more than the data itself (Nguyen et al., 2020, but see also clustering algorithms in general; von Luxburg et al., 2012). LDA topics can be thought as the resolution of a microscope (Barron et al., 2018; Nguyen et al., 2020): if a fine-grained resolution is required, then a large number of topics is better; if the number of topics is small, these topics become more general (Allen & Murdock, 2020). Here, topic extraction was performed with the *topicmodels* R package (Grün & Hornik, 2011), using Gibbs sampling. We left the other LDA parameters set as default, while setting the same seed for reproducibility for all topic extractions. We performed topic extraction with three different levels of resolution, setting k at 100, 200, and 300 topics. As a consequence, summing all k topics, we obtained 600 topics (see Section SM6 for a thorough description of topics and Section SM7 for topic comparison between different ks). In Section SM7.1 of the supplemental material, we have suggested a way to assess topic specificity based on the position of a theme's keyword (e.g., "Diana" for Princess Diana) within the beta weight-ordered topic's terms, and the correlation with lexical features. If the theme is event-based (e.g., disappearance of Malaysia Airlines Flight 370 [MH370], 8 March 2014), we also suggest visually inspecting the gamma values plotted over time.

As a proxy for document topic, for each of the three sets of k topics, we extracted the topic that had the highest probability of representing the document, i.e., the highest gamma value within all topics within k, and included it in the LOCO dataset (see dataset description in "Data availability" section). This means that each document is associated with three topic labels, one for each k. We chose this option so as to offer LOCO users a way to perform analyses on a specific topic resolution. Note that we did not provide labels for document topics. Instead, we provide the top 15 words for each topic that, taken together, summarize the topic's content (Nguyen et al., 2020, and see also beta weight distributions by k in Section SM6.1).

We provide with LOCO the matrix containing all gamma values for each document and topic pairs (see "Data availability" section). This results in a matrix with a dimensionality of 96,743 documents × 600 topics. This is useful for obtaining a fine-grained topic description for each document. For example, if a document n has the topic with the highest $\gamma = .90$, then this topic has 90% probability of representing document n, while the remaining 10% is distributed among all other topics. Similarly, if the highest $\gamma = .10$, all the other topics, by exclusion, occupy the remaining 90% of probabilities. While in the first case we can say that document n is well represented by a topic k (where gamma is maximum), in the second case, the low gamma value shows that the document n is not well represented by a topic k. LOCO contains all γ values, allowing the user to select their own threshold when selecting documents based on topic.

Data associated with LOCO's topics

In order to facilitate topic exploration prior to data analysis, we attach additional files to LOCO that offer an in-depth description of topic content. The first one is a matrix that contains all gamma values for each topic for each document (topic_gamma.json). Because there are three sets of k topics (100, 200, and 300), we have named each topic adding the k resolution as prefix. For example, the fifth topic of k200 is labeled "k200_5," while the 134th topic of k300 is labeled "k300_134." Note that, because we merged the three sets of ks into a unique dataset, the sum of topic probabilities for each document is now 3 (1 for each k set of topics). The second file (topic_description.json, see also description in SM6) includes descriptions for each of the 600 topics. Descriptions include the top 15 terms ordered by beta weight, the number of documents in which the topic has the highest gamma, the highest correlation with other topics and highest correlation with lexical features (both LIWC and Empath). We also provide a series of plots (in the file "topic_by_time.pdf," see description in SM6), one for each topic, that track the evolution through time (from 1995 to 2020, see e.g. Fig. 2) of the gamma values. Each plot also includes the topic name and the list of the top 15 terms, ordered by beta values. We believe that these plots, along with the description of each topic (and the actual matrix with gamma values), will help researchers not only in exploring topic associations and lexical features, but also in visually inspecting topics prior to data analysis.

Representative conspiracy theories

Because one might be interested in what a prototypical conspiratorial language is, we aimed at extracting a set of the most representative CT documents on the basis of the most frequently occurring words within the conspiracy subcorpus. We believe that a set of representative documents may allow researchers to make inferences about CTs more generally. As such, a representative document should share more words with the conspiracy subcorpus compared to a less representative document. Recurrent word patterns such as "they are trying to KILL US!" (from document C01b90) or "know the truth" (document C073a0) might in fact be highly shared across conspiracy documents; hence they would be represented to a larger extent in the conspiracy universe.

Following this reasoning, we extracted the documents that were most similar to the entire conspiracy subcorpus. As a measure of representativeness, we computed the cosine similarity (CS) between words of each document against all words in the conspiracy subcorpus (for a similar procedure, see e.g. de Vries et al., 2018). Text preprocessing was the same as we used to extract LDA topics (see SM5). Documents' CS was computed using the `textstat_simil` function from the R package *quanteda*. Values range from 0 to 1, indicating either no overlap (0) or a perfect overlap (1) of terms. This returned a vector for each conspiracy document that indicated the similarity between it and all other conspiracy documents. We averaged this vector to obtain a single value for each document. We finally labeled as "conspiracy representative" the documents whose CS value was higher than one standard deviation above the mean. This resulted in a subset of 4,227 documents, that is, 17.66% of the conspiracy subcorpus. In Section SM8 of the supplemental material, we report the top five documents with the highest and lowest cosine similarity.

Metadata

Mentioning "conspiracy"

We marked documents that mentioned conspiracy in the text. This was done by searching, via regular expressions, and counting the occurrences of the word "conspir*."[9] This measure helps keep track of mainstream documents that mention conspiracy which may contaminate mainstream language with details about the corresponding conspiracy (e.g., Pizzagate or Illuminati, themes that rarely appear outside the context of CTs). Therefore, instead of removing these documents, as they represent a special case of mainstream media whose focus is on CTs, we left them in LOCO and annotated the number of instances of the word "conspir*."[10] In the conspiracy subcorpus, a total of 3520 documents mentioned conspiracies at least once, while in the mainstream subcorpus there were 5031 documents. On average, conspiracy documents show more instances of "conspir*" than mainstream documents (conspiracy: $M = 0.351$, $SD = 1.548$, range: 0–75; mainstream: $M = 0.211$, $SD = 1.735$, range: 0–182, $t_{(45246)} = 11.773$, $p < .001$, $d = 0.09$). However, when the instances were normalized per word count (i.e., divided by numbers of words in text), there were no differences, $t_{(49107)} = .993$, $p = .321$, $d = 0.01$.

Text statistics

For each document, we calculated the number of words, sentences, and paragraphs using the Tool for the Automatic Analysis of Cohesion, TAACO (Crossley et al., 2016, 2019), a freely available standalone application that allows batch processing of text files. Although LIWC also provides measures of word count, which correlates highly with TAACO word count, $r = .9996$, we relied on TAACO measures for two reasons. First, based on the Python Natural Language Toolkit (Bird et al., 2009), TAACO extracts the part of speech for each word, from which it derives a text word count as well as the number of sentences and paragraphs. This, we believe, is a more sophisticated way than merely counting instances of characters separated by spaces. Secondly, because the word-per-sentence measures of LIWC and TAACO correlate poorly, $r = .59$, we visually inspected documents with the highest discrepancy between the two tools. We discovered that LIWC performs poorly when full stop periods are missing from sentences, whereas TAACO considers the new line as a valid

[9] The word "conspir*" was chosen to be able to retrieve all conspiracy-related words (conspiracies, conspiracist, conspiracy, conspiration, conspirator, conspiratorial, conspiratorially, conspiratress, conspire, conspirer, and conspiring) but not others (e.g., conspicuous). This was checked on both American and British Oxford English dictionaries.

[10] From this count measure, a Boolean measure of "mentioning conspiracy" can easily be derived by simply stating "TRUE if mentions > 0."

sentence-separator marker. Therefore, in LOCO, we keep both LIWC and TAACO word counts, but for consistency with paragraph and sentence counts, we report here (see Table 2) only the TAACO word count.

Date

Information about document date was obtained primarily from the *Goose* package, which extracts the upload date directly from the raw HTML document. When date was not available (i.e., *Goose* returned an empty cell), we extracted the upload date with regular expressions from the URL of the document (e.g., "http://[…]/2018/01/23/[…].html" was coded as 23 January 2018). In LOCO, date data are provided for 63,868 documents (67% of the entire corpus; 56.67% conspiracy and 69.09% mainstream), see distribution of documents by date in Fig. 1.

It must be noted that date values reflect either the upload date or the authoring date. Both types of information would be informative for different purposes: texts that were authored on the same date are based on a similar level of available information/evidence; texts that were published on the same date compete for audience attention.[11] While dates before the internet era (e.g., 1853) refer unambiguously to the authoring date, this is less clear for more recent documents. We believe that this information might be nevertheless useful, and therefore we provide all dates available in LOCO. We warn researchers to be aware of date ambiguity before testing any time-related hypothesis. Researchers can either set a threshold for documents' dates to keep (e.g., after the internet became widespread or another arbitrary cutoff) or develop a method to disentangle the two. However, although documents' dates may refer to either authoring or upload date, we show in Fig. 2 that documents' dates are nevertheless linked to the social events discussed in documents.

Lastly, date range differs between mainstream and conspiracy subcorpora, see Table 2. We do not know the reason for this difference, considering that our Google search was independent from documents' upload date. One possible explanation is that conspiracy websites, being less popular (see Table 5), are also developed with less standardized protocols (see e.g., www.w3c.org). This might have resulted in a less methodical use of HTML meta-tags and therefore the lack of date in some documents. This might also explain the higher percentage of missing dates in conspiracy documents (56.67%). If this is the case, some documents predating 2004 (i.e., the oldest conspiracy document in LOCO) might be in this corpus yet lacking the date. Alternatively (or complementarily), conspiracy websites might be younger, overall, than mainstream websites. For example, the infowars.com domain was registered on 1999-03-07 (data obtained from https://who.is), 911truth.org on 2003-01-14, ahtribune.com (less popular in terms of monthly visits among LOCO's conspiracy websites) on 2015-08-23, and worldaffairsbrief.com (most popular) on 2004-04-06. In contrast, scientificamerican.com was created on 1997-05-02, sciencemag.org on 1996-04-28, cnn.com on 1993-09-22, and bcc.com on 1989-07-15. Although not tested systematically, those few observations suggest that, overall, conspiracy websites in LOCO might be younger than mainstream ones, therefore explaining the different date ranges.

Facebook shares

For each webpage, we obtained information about spread from the web tool sharedcount.com (SC). Via an application programming interface, SC retrieves from Facebook[12] the number of shares, comments, and reactions for each webpage URL. According to the website, SC reports "all time statistics," which means that values refer to the overall shares since the creation of the URL tracked. All data from SC were collected in September 2020.

Besides single URL shares, we also computed an estimation of the total number of shares from the observed data we collected for each website. To this end, we computed the sum of all webpage Facebook shares for each website and divided them by the proportion of sampled LOCO webpage for each website. For instance, in LOCO, there are 967 documents extracted from the website www.infowars.com. Infowars has 15,500 webpages indexed on Google (see "Website metrics" section), which means that LOCO contains 6.24% of all Infowars webpages. The aggregated total Facebook shares of all 967 Infowars documents in LOCO is 89,639. By dividing the total shares (89,639) for the proportion of LOCO documents (0.0624), we obtain an estimation of total website shares, which in this case is 1,436,820 times, a rough estimation of the grand total of shares of all Infowars webpages. Once this measure was computed for all websites, we then tested the correlations of this measure with other spread measures. The estimated Facebook shares correlates with website global rank (r = -.81) and with website monthly visits (r = .81, see SM9 for more details).

Website category

We relied on MBFC for obtaining metrics of political side and factual reporting for each website. MBFC contains manual annotations and bias analyses for over 2,000—mostly news—websites. According to the MBFC method,[13] each website's bias is evaluated on four criteria, including biased

[11] We thank the anonymous reviewers for this suggestion.

[12] see: https://developers.facebook.com/tools/debug/)

[13] https://mediabiasfactcheck.com/methodology/

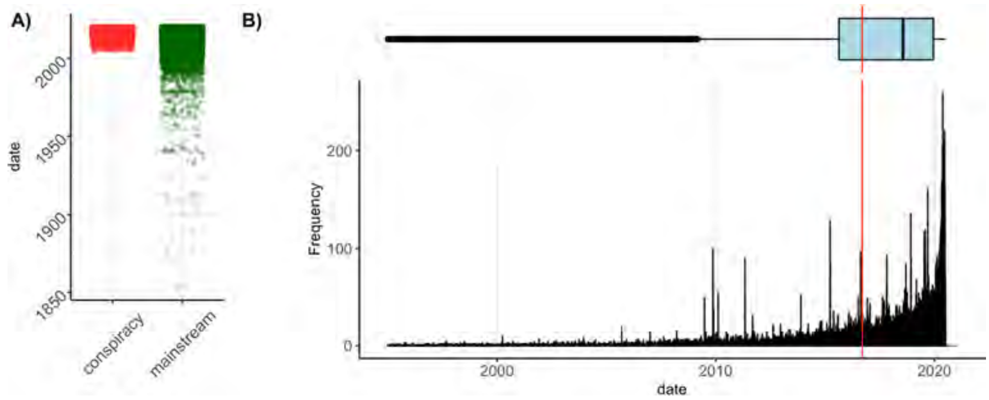


Fig. 1 Distribution of documents in LOCO by date. Distribution for **a** each subcorpora (red: conspiracy; green: mainstream) and **b** all documents from 1995 to the time of data collection (the red vertical line represents the mean, the boxplot on top displays the median and the interquartile ranges)

wording headlines (e.g., the source uses loaded words to convey emotion to sway the reader), factual sourcing (e.g., the source reports factually and backs up claims with well-sourced evidence), story choices (e.g., the source reports news from both sides), and political affiliation (e.g., the source endorses a particular political ideology). Factual reporting is based on the factual sourcing used for assessing bias. For each website, a minimum of 10 headlines and 5 news stories are assessed by MBFC experts. Low and very low factual reporting sources are those that need to be fact-checked for intentional fake news, conspiracy, and propaganda. Although MBFC states that their methodology has been not tested scientifically, they nevertheless adhere to the International Fact-Checking Network fact-checkers' Code of Principles[14] and strive for transparency. Furthermore, MBFC annotations have been used by other researchers to study fake news and conspiracy websites (Baly et al., 2018; Cinelli et al., 2021; Pennycook & Rand, 2019; Risius et al., 2019).

For each of the LOCO websites that was reviewed in MBFC, we extracted measures of political orientation (left, left center, least biased, right center, and right), factual reporting (from "very low" to "very high"), pseudoscience level (provided by MBFC only for conspiracy websites), and whether the website was labeled as pro-science (i.e., relying on legitimate science or evidence based on credible scientific sourcing). Note that pro-science websites do not have political orientation labels. Data from MBFC were collected in July 2020.

Website metrics

We have extracted a series of website metrics that, overall, offer an idea of popularity, engagement, and size for each website. From the web tool similarweb.com[15] (SW), we collected data about monthly total visits, global rank, and category. We also collected information about the type of incoming traffic. Expressed in percentage, these metrics partition each website's incoming traffic into direct (when a user reaches the website directly by typing the URL on the web browser or recalling it from bookmarks), from a search engine (when a website is reached through a search engine, e.g., Google), and from social media (when a website is reached through social media, e.g., a post on Facebook or Twitter). Other types of incoming traffic offered by SW, which we did not collect, are referrals, mail, and display, which overall account for about 7% (SD = 6.38) of remaining incoming traffic in our dataset (computed by summing direct, search engine, and social media traffic and subtracting it from 100).

SW was chosen over Alexa.com (a web tool that provides similar services), mainly because SW updates its statistics every month, whereas Alexa provides daily updates. While the latter appears to be more fine-grained, it nevertheless poses some limitations in terms of data collection (which manually spans several days) due to daily statistical fluctuations. In addition, SW offers a wide range of free features, otherwise accessible in Alexa upon a monthly subscription, and, importantly, the SW database is composed of ~50 million websites (vs. ~30 million websites in Alexa). These data were collected in July 2020.

In addition, in order to obtain an estimation of the website size, we extracted the total number of webpages per website indexed by Google. This was done by querying Google with "*site:*" followed by the website.[16] This data was collected in March 2021.

Data availability

LOCO's data is freely available at https://osf.io/snpcg and includes:

[14] https://www.poynter.org/ifcn/
[15] https://www.similarweb.com/corp/ourdata/
[16] E.g., https://www.google.co.uk/search?q=site%3Abbc.com

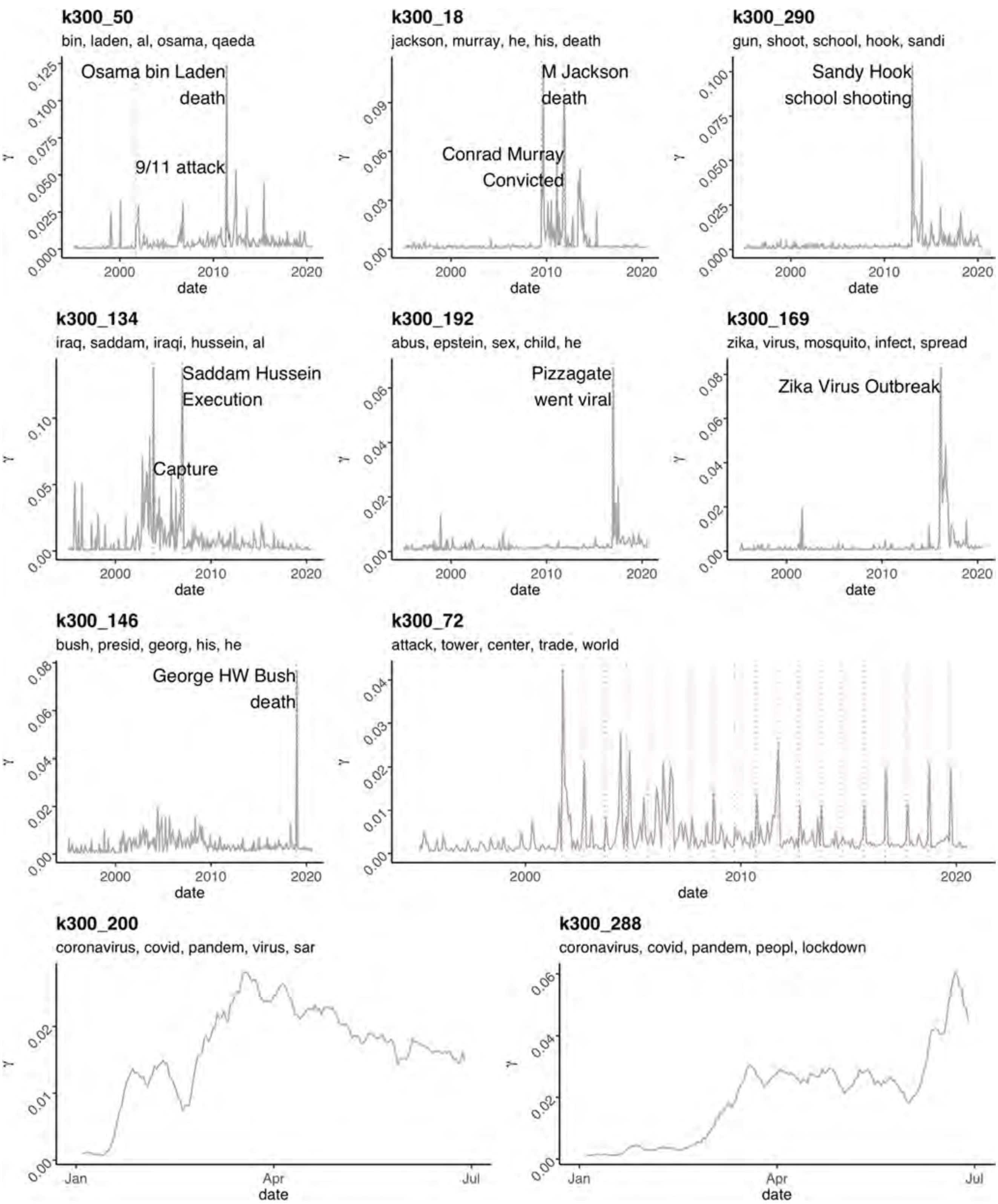


Fig. 2 LDA topic gamma values over time**.** The red dotted vertical lines represent the occurrences of significant events associated with the topic. In the 9/11 topic, each vertical line represents September 11th in each year, starting from 2001. Coronavirus topics (bottom) are distributed over the year 2020 (from January to July, when LOCO data collection ended).

Table 5 Differences between conspiracy and mainstream website metrics

	Mainstream			Conspiracy			*t*-test statistics (raw)			*t*-test statistics (log)		
	M	*(SD)*	*N*	*M*	*(SD)*	*N*	*t*	*p*	*d*	*t*	*p*	*d*
Total monthly visits	102,285,513	(191,306,614)	92	965,242	(2,115,315)	28	5.08	***	1.10	14.00	***	3.02
Global rank	7313	(21,765)	89	211,904	(168,890)	28	−6.39	***	1.39	–17.11	***	3.71
Website size	6,844,908	(16,049,205)	92	6224	(12,918)	58	4.09	***	0.69	18.43	***	3.09
FB projected shares	3,213,458,353	(9,348,074,961)	92	27,11,190	(14,540,274)	58	3.29	**	0.55	12.78	***	2.14
Traffic, direct†	28.95	(13.45)	92	57.55	(21.57)	28	–6.63	***	1.43			
Traffic, search†	56.83	(17.49)	92	13.82	(10.44)	28	16.00	***	3.45			
Traffic, social†	8.08	(5.58)	92	18.4	(19.28)	28	–2.8	**	0.60			

Note. Differences tested with Welch's unequal variances *t*-test. Log transformation was applied to highly skewed variables after having added a constant 1 to avoid -Infinite values when the raw score was zero. † Values expressed as percentages and not log-transformed. *d*: Cohen's *d*. FB: Facebook. Website size is expressed in number of webpages

1. **LOCO.json** (587.6 MB): a JSON (JavaScript Object Notation) file containing the LOCO corpus itself. 96,746 rows (documents) × 20 columns (see Table 6)
2. **website_metadata.json** (55.3 KB): a JSON file containing websites' metadata. 150 rows (websites) × 18 columns (see Table 7)
3. **LOCO_LFs.json** (573.1 MB): a JSON file containing the full set of lexical features. 96,746 rows (documents) × 288 columns (N_{Empath} = 194; N_{LIWC} = 93)
4. **topic_gamma.json** (963.7 MB): a JSON file containing topics' gamma values. 96,746 rows (documents) × 600 columns (topics)
5. **topic_by_time.pdf** (169.6 MB): a PDF file containing plots of topics' gamma values over time (from 1995 to 2020). It contains 600 pages.
6. **topic_description.json** (188.2 KB): a JSON file containing detailed descriptions of topics. 600 rows (topics) × 12 columns (see SM6)

Exploring LOCO's features

In this section, we explore LOCO's features and provide examples on how to handle LOCO's variables and subset corpus. Some of these analyses are descriptive in nature and offer a way to visually explore to what extent LOCO's data relate to the external world, such as visualizing the evolution of LDA topics through time (see "Topic analyses" section) or exploring to what extent the language used in LOCO's documents overlaps with the language used in social media (see "Towards corpora of CT texts" section). Other analyses are more explorative, such as testing whether mentioning conspiracy in mainstream documents affects lexical features (see "Effect of mentioning conspiracy" section) or whether conspiracy-representative documents are in fact different from other conspiracy documents in terms of lexical features and spread (i.e., Facebook shares, see "Properties of representative conspiracy documents" section). Lastly, we also explore to what extent LOCO's higher-level metadata might provide insights into psychological processes by analyzing the behavior of websites' users (see "Website incoming traffic" section). Overall, these analyses not only suggest how to use LOCO, but also offer insights on the language of conspiracy and the psychology of conspiracy websites users.

Topic analyses

Each document in LOCO is associated with a vector that encapsulates and quantifies the semantic content, namely the LDA topics. While in the main dataset (LOCO.json) we provide for each document only the label of the most prevalent topic (one for each level of topic resolution, that is k = 100, k = 200, and k = 300), in a separate dataset (topic_gamma.json) each document is associated with the gamma values for all 600 LDA topics extracted. In this section, we explore how LDA topics reflects real-world events by visually inspecting how these LDA topics develop through time for documents whose date was recorded. This reasoning is supported by the fact that, because texts are capable of showing cultural patterns (Lansdall-Welfare et al., 2017; Li et al., 2019, 2020; Michel et al., 2011), a significant social event should be reflected in the texts' topic time series. To explore this possibility, we selected a set of topics that are associated with a specific event (instead of non-event-specific topics such as AIDS or Illuminati) such as the death of societally significant people: Osama bin Laden (2011-05-02, topic k300_50), Michael Jackson (2009-06-25, k300_18), George H.W. Bush (2018-11-30, k300_146), and Saddam Hussein (2006-12-30, k300_134); outbreaks of pandemics such as Zika virus (2016-02-01, k300_169) and coronavirus

Table 6 LOCO dataset variables description

Variable name (% empty/missing values, if any)	Variable description
doc_id	Six-character hexadecimal sequence of document unique identification number. The first character stores the source: C stands for conspiracy (e.g., C0004d) and M stands for mainstream (e.g., M095eb)
URL	URL associated with the document
Website	The website from which the document was extracted
seeds (2.26%)	The seeds we used to gather documents. The page was returned by all the keywords listed in this variable (N=47)
date (33.98%)	The date the webpage was uploaded or uploaded (format: YYYY-MM-DD)
subcorpus	Either conspiracy or mainstream ($N_{conspiracy}$=23,937; $N_{mainstream}$=72,806).
title (0.11%)	Title of the document
txt	Document text (see text statistics in Table 2)
txt_nwords	Number of words
txt_nsentences	Number of sentences
txt_nparagraphs	Number of paragraphs
topic_k100	The topic ID with highest gamma value within k100 LDA (N=100 unique, e.g., k100_24)
topic_k200	The topic ID with highest gamma value within k200 LDA (N=200 unique, e.g., k200_75)
topic_k300	The topic ID with highest gamma value within k300 LDA (N=300 unique, e.g., k300_192)
mention_conspiracy	Occurrences count for the word "conspir*" in text, see "Mentioning conspiracy" " section
conspiracy_representative	Logical. TRUE (N=4227) if the conspiracy document is representative
cosine_similarity	Cosine similarity values for conspiracy documents (values > mean + 1 SD are considered representative)
FB_shares (0.01%)	URL's Facebook shares
FB_comments (0.01%)	URL's Facebook comments
FB_reactions (0.01%)	URL's Facebook reactions

Note. Percentages of empty/missing values are calculated on the list of documents (N = 96,743)

Table 7 LOCO's website metadata variables description

Variable name (% empty/missing values, if any)	Variable description
Website	Website name (N=150)
URL	URL associated with the website domain
n_webpages	Overall number of webpages in website obtained by Google search (see "Website metrics" section)
MBFC_political_orientation (69%)	Political orientation. Left (N=4), left_center (N=19), least_biased (N=15), right_center (N=4), right (N=5)
MBFC_factual_reporting (21%)	Factual reporting. Very_low (N=10), low (N=43), mixed (N=16), mostly_factual (N=4), high (N=35), very_high (N=11)
MBFC_conspiracy	Logical. If TRUE (N=58), website is conspiracy
MBFC_pseudoscience (62%)	For conspiracy websites only. Zero (N=1), mild (N=2), moderate (N=9), strong (N=16), quackery (N=29)
MBFC_proscience	Logical. TRUE (N=16) if website is labeled as pro-science
SW_total_visits (20%)	Total visits, desktop and mobile web aggregated
SW_global_rank (22%)	Traffic rank of website, as compared to all other websites in the world
SW_Category (20%)	Website category (e.g., news_and_media, N=60; health, N=16, science_and_education, N=13)
SW_traffic_direct (20%)	Percentage of direct desktop incoming traffic (from typing the URL in a browser)
SW_traffic_search (20%)	Percentage of search desktop incoming traffic (from a search engine)
SW_traffic_social (20%)	Percentage of direct desktop incoming traffic (from a URL on social media)
FB_shares_homepage	Facebook shares of homepage (see discussion in SM9)
FB_shares_estimated	Estimated overall Facebook shares given total number of website's webpages (see "Facebook shares" section)

Note. Percentages of empty/missing values are calculated on the list of websites (N = 150)

(2020-03-11, k300_200 and k300_288); and other significant societal events such as the 9/11 terroristic attack (2001-09-11, k300_72), the Sandy Hook school shooting (2012-12-14, k300_290), and Pizzagate (2016-11-01, when it went viral, k300_192). In Fig. 2, for all documents in LOCO provided with upload/creation data, topic patterns (i.e., gamma values on the *Y* axis) are shown within a time span of 25 years, from 1995 to 2020 (first three rows) and for 2020 (fourth row for coronavirus-related topics) from January to July, when LOCO's data collection ended.

Lexical features

Overlap with Reddit users' language

Ideally, a corpus must be representative and replicable, meaning that the sampled data should represent the full range of variability of the population from which the sample is drawn. If our corpus successfully represents CTs, then its content should mirror the content of comments and threads posted by conspiracy believers on social media. To this aim, we compared the lexical features extracted from LOCO's documents (LOCO_LFs.json) with those extracted from comments on Reddit by Klein et al. (2019). Although user discussions on conspiracy forums are not conspiracy per se, we expect a certain overlap in language features with LOCO documents. This is because, while forums do not offer adequate space to fully develop argumentative discourses, a conspiracy believer can nevertheless express a conspiratorial worldview through language use (e.g., deception: "*They are hiding the cure from us for their own profit!!*"), even in discussion not related to conspiracy. In fact, Klein et al. (2019) compared language features of a group of users who posted in the r/conspiracy subreddit with those from a carefully matched control group of users who never posted in r/conspiracy. Although we do not know to what extent users who posted in the r/conspiracy subreddit endorse CTs, Klein and colleagues found language differences associated with a conspiratorial mindset (e.g., power, deception, dominance) that sees hidden powerful and malevolent enemies at work.

We proceeded with replicating the method of Klein et al. (2019) on LOCO by comparing our two subcorpora and explored whether the same patterns emerged. Similar to their work, we used the lexical features derived from Empath and tested differences between conspiracy and mainstream documents on the 194 Empath categories. Then, we used Welch's *t*-test and computed Cohen's *d* for each test on the variables that yielded a significant difference at $p < .00026$ (Bonferroni correction for 194 tests). Note that here we are not testing any particular hypothesis, but provide this as exploratory analysis to guide future research. Results are shown in Fig. 3. On the top (A), only variables that produced an effect size of $d > .20$ are displayed, arranged in decreasing order. On the bottom (B), each variable was scaled to *z* values, and mean values are shown for different website categories: conspiracy_representative (N = 4,227), other conspiracy (N = 19,710), biased_LR (aggregating documents biased towards either the left or right, N = 31,928), least-biased (N = 14,180), and pro-science (N = 11,440).[17]

Lexical differences between LOCO conspiracy and mainstream documents overlap with those between Reddit groups found by Klein and collaborators (Fig. 3a). Among the lexical categories characterizing conspiracy language (i.e., positive values in Fig. 3a), half of them emerged as overlapping between the two datasets. In LOCO, other lexical categories were higher in conspiracy (vs. mainstream), such as *divine* and *worship* that correlate with *religion* (r = .92, r = .95, respectively, in our dataset) found in Klein, and *kill* and hate that correlate with *death* (r = .72, r = .44) and *negative_emotion* (r = .71; r = .76) found in Klein but not in LOCO. It is also worth noting that representative conspiracy documents, on average, display an exaggeration of the "average" conspiratorial language as exemplified from the means further departing from zero (this will be further explored in "Properties of representative conspiracy documents" section).

Effect of mentioning conspiracy

We explored the possibility that mentioning conspiracy in the text would increase conspiratorial language (see "Mentioning "conspiracy" section). To this aim, we run multiple *t*-tests using as dependent variables the 31 lexical categories that yielded an effect size $d > .20$ from the previous section (see "Overlap with Reddit users'" section). In doing so, we used two different LOCO subcorpora: one (raw) which is based on the whole mainstream data ($N_{\text{mainstream}}$ = 72,806) and one (cleaned) from which we removed all mainstream documents containing at least one instance of the word "conspir*" (Final $N_{\text{mainstream}}$ = 67,775). Note that we removed mentions of conspiracy only in the mainstream documents because we aimed to test the difference between the subcorpora removing potential conspiratorial language from mainstream documents, so as to obtain a mainstream subcorpus cleaned of conspiratorial language. We reasoned that conspiracy documents deliver conspiracy language even without mentioning the word "conspir*." From each test, we extracted the effect size (Cohen's *d*) and then compared the changes in *d*, with a paired *t*-test, from the raw to the cleaned dataset. Results show an overall increase in the effect sizes, $t_{(30)} = 5.08$, $p < .001$, suggesting that cleaning the mainstream subcorpus from

[17] Note that the sum of pro-science, least-biased, and biased documents is 57,575 (and not 72,806, the total of mainstream documents). This is because not all websites are rated by MBFC, and therefore documents from these websites do not compare in plot B.

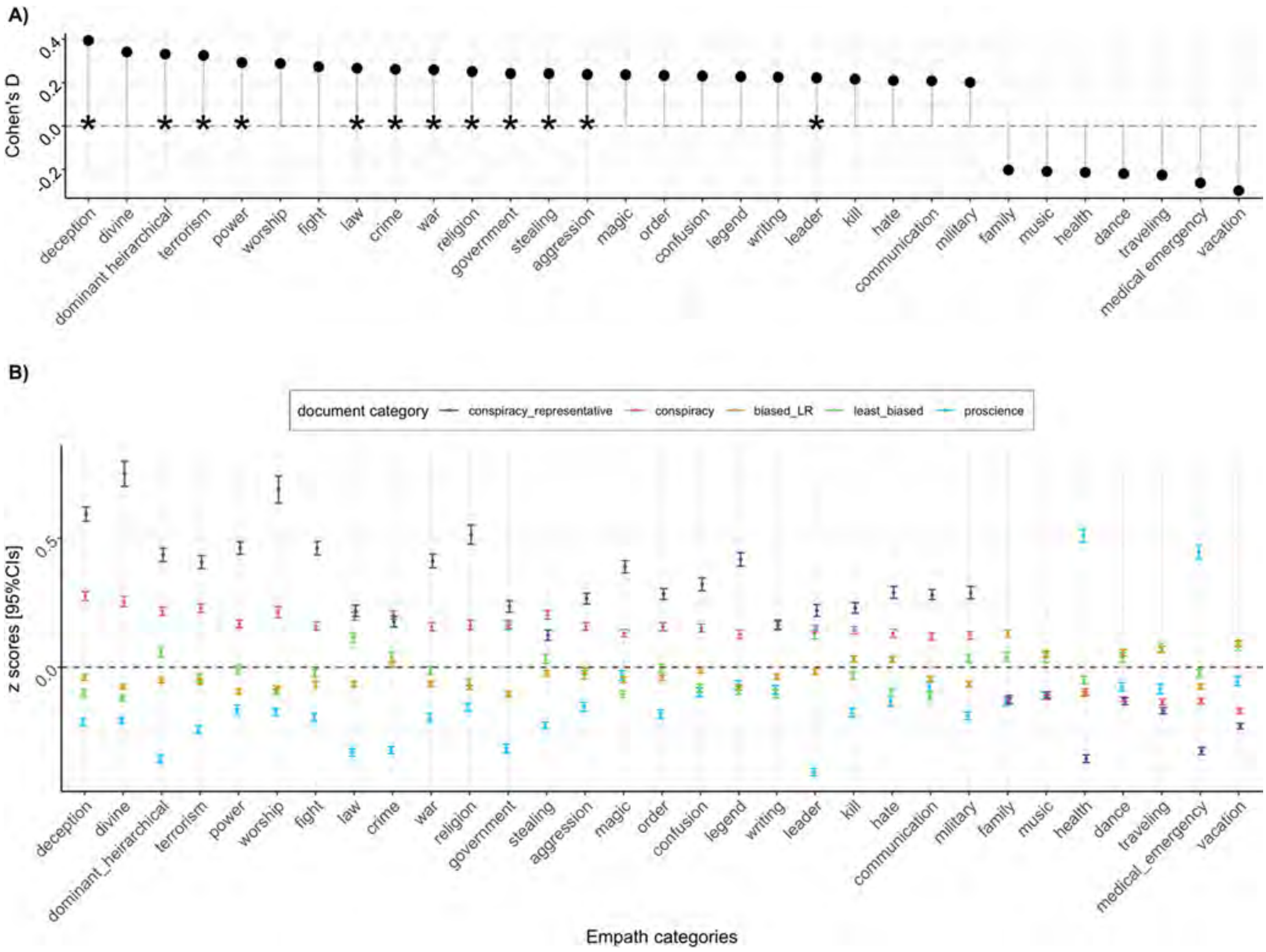


Fig. 3 Differences in lexical features between conspiracy and mainstream documents **a** Effect sizes that yielded a Cohen's $d > .20$ from t-tests between conspiracy and mainstream documents on Empath lexical categories. Positive effect sizes indicate that the category value is higher in conspiracy documents. A star indicates that the category emerged as having $d > .20$ also in Klein et al. (2019). **b** Comparison of means [and 95% CIs] for the same set of variables (scaled to z values) across different document categories

documents that mention conspiracies amplifies differences in language features between the two subcorpora (see SM10 for details).

We finally explored whether mentioning conspiracy had an effect on lexical features. To this aim, we extracted the top four Empath categories that in the previous analysis (see paragraph above) had yielded the largest changes in effect size, namely *crime*, *terrorism*, *deception*, and *stealing*, and tested the correlation (on log-transformed variables, but see SM10 for non-log-transformed results) between the number of mentions of conspiracy and lexical variables. The results showed a positive relationship: crime: $r = .31$; terrorism: $r = .33$; deception: $r = .21$; and stealing: $r = .13$. Overall, these tests show that mentioning conspiracy, even in mainstream documents, affects language features. Therefore, we suggest that researchers carefully evaluate whether or not to include mainstream documents containing the word "conspir*" in their analyses.

Properties of representative conspiracy documents

We explored to what extent the representative set of conspiracy documents ($N = 4{,}227$) differs from the other conspiracy documents ($N = 19{,}710$) in terms of lexical features. To this aim, after subsetting LOCO to only conspiracy documents, we run a series of linear mixed-effects models using the *lme4* (Bates et al., 2015) and the *lmerTest* (Kuznetsova et al., 2017) R packages. In each model, we specified as dependent variables the LIWC ($N = 93$) and Empath ($N = 194$) categories, and as fixed effects the dichotomous representativeness predictor. As random intercept, we specified both the websites from which documents were extracted and the topic label with the highest gamma value for $k = 100$ because, being less specific, it provides a more inclusive clustering that aggregates similar topics. In other words, while for $k = 300$ we would have had several LDA topics revolving around a theme, with a lower topic resolution, topics are more general (we replicated the same analyses with $k = 200$ topics, and results are not visibly different, see SM11). Before entering the model, the dependent variables were scaled to z values. Standardized β estimates are displayed in Fig. 4 for only the dependent variables that were significant at $p < .00017$ (Bonferroni correction for 287 tests) by the dichotomous predictor. Positive estimates indicate that the category is higher in the representative subset of conspiracy documents.

The representative conspiracy subset is generally more emotionally charged than the other documents, as displayed by the higher value for the category related to affective processes (LIWC category *affect*), and more specifically to negative emotions (LIWC categories *anger*, *swear*, *negemo*). Representative conspiracy documents, as compared with the non-representative conspiracy documents, display a prototypical language of conspiracy focused on power, dominance, and aggression (Empath categories *deception*, *dominant hierarchical*, *kill*, *hate*, *order*, *power*, *aggression*, and *rage*).

As for the rhetorical style used by the representative subset, we observe higher values for certainty (category *certain*), and interrogative (category *interrog*) language, along with higher use of question and exclamation marks (categories *Exclam*, *QMark*). This is in line with the observation that the rhetorical style of conspiracy narratives is built upon refutational strategies based on questioning the dubious version of the official story while highlighting the lack of answers from official sources (Oswald, 2016).

In line with research on social motives underlying belief in CTs (Douglas et al., 2019), the higher use of *we* and *they*, along with affiliative (LIWC category *affiliation*) and social (category *social*) language, suggests a process of social identification of the ingroup (*we*) by exclusion from the outgroup (*they*).

Overall, as already seen in Fig. 3 and in the work of Klein et al. (2019), the representative conspiracy documents seem to be an exaggerated version of an average conspiracy document, characterized by language of power, action, and dominance. They are at the same time less likely to display non-conspiratorial language as exemplified by lower values for categories such as *tourism*, *vacation*, *urban*, and *morning*. Interestingly, these patterns overlap with those found on Twitter, in which lexical differences between conspiracy and science influencers were identified in the use of negative emotion (e.g., anger) and a focus on topics such as death, religion, and power (Fong et al., 2021).

If the representative documents are rhetorically appealing and emotionally loaded, then we can expect that they will spread more successfully than the other, less representative documents. This reasoning is also in line with the fact that emotional content is a successful feature of narrative stickiness and transmission (Franks et al., 2013; Heath et al., 2001). Therefore, we tested whether the representative subset of conspiracy documents spread more than non-representative conspiracy documents. To this aim, we fit a linear mixed-effects model predicting Facebook shares (log-transformed). We set conspiracy representativeness as predictor along with website total visits as covariate while specifying a random intercept for websites. Results showed that the subset of representative

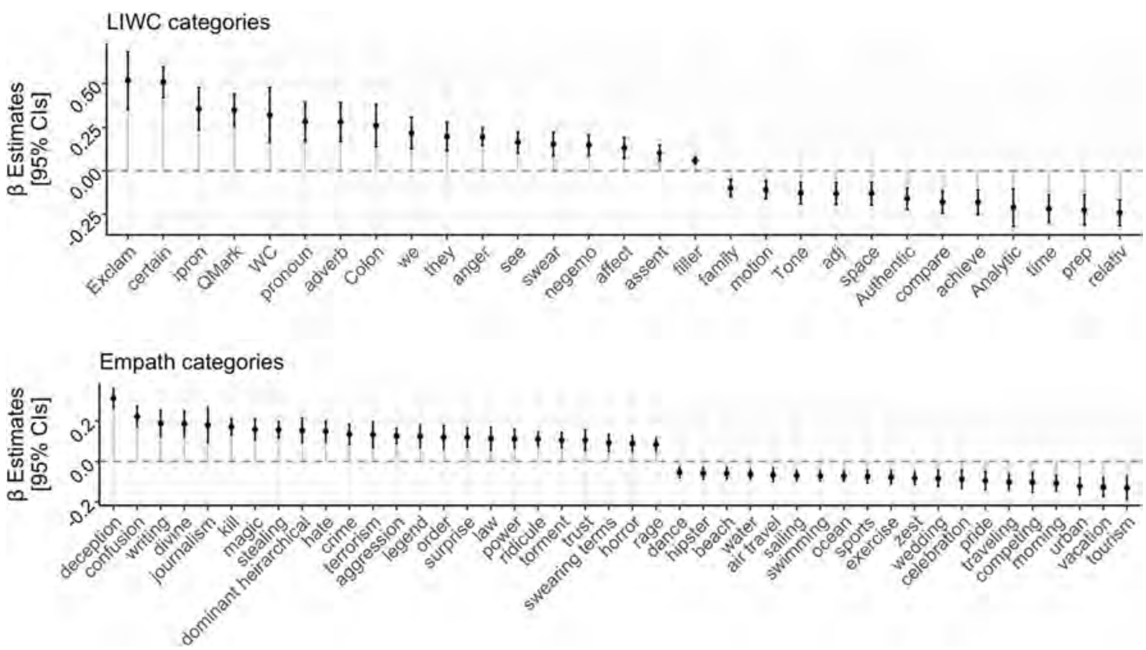


Fig. 4 Differences in lexical features between high- and low-representative conspiracy documents. Positive β estimates indicate that the category is higher among conspiracy documents that are more representative of the conspiracy corpus as measured by their document cosine similarity with other conspiracy documents in the corpus

conspiracy documents was more shared on Facebook compared to the other conspiracy documents, $\beta = 0.121$, $SE = 0.017$, $t = 7.075$, $p < .001$.

Website incoming traffic

Besides the texts themselves and the documents and their metadata, LOCO is also provided with higher-level metadata, namely information about its constituent websites. Such a set of variables (contained in the file website_metadata.json) might be useful for testing hypotheses at the website level. For example, here, we describe the behavior of conspiracy and non-conspiracy communities from websites traffic information.

Analysis of online social media shows that users tend to aggregate in echo chambers that are homogeneous clusters of communities of interest (Bessi, Coletto, et al., 2015a; Brugnoli et al., 2019; Del Vicario, Bessi, et al., 2016a). Such clustering is reinforced in online and offline social networks (Del Vicario et al., 2017) whereby a like-minded trusted node in the social network (a friend or a page followed on Facebook) shares content that adheres to a system of beliefs. Moreover, within online social networks, users access information through a narrower spectrum of sources compared to web searches (Nikolov et al., 2015), meaning that being embedded within a social bubble reduces exposure to different viewpoints. When users of conspiracy Facebook pages are exposed to debunking information, they increase traffic towards conspiracy-like content (Zollo et al., 2017). This behavior suggests a confirmation bias: people avoid cognitive dissonance while searching for reinforcement (Brugnoli et al., 2019; Hills, 2019).

Website incoming traffic provides similar information about user behavior. For example, direct traffic may indicate a certain level of loyalty or at least that the user knows the website or has learned about it through their social contacts (Pauwels et al., 2016). When a website is reached from a search engine, the website is not necessarily known to the user. Put differently, how people arrive at a website may indirectly reveal information about their prior knowledge, beliefs, and social community. If echo chambers provide links to belief confirming content, then a confirmation bias theory of conspiratorial thinking would predict that users of conspiratorial websites are more likely to arrive there via a bookmarked URL or through online social networks than through impartial search engines.

To explore this possibility, we analyze user behavior through website incoming traffic (see "Website metrics" section). Because of a link between confirmation bias and belief in CTs (Del Vicario et al., 2017; Del Vicario, Bessi, et al., 2016a; Marchlewska et al., 2018; Meppelink et al., 2019; Zollo et al., 2017), we expect that conspiracy websites display higher levels of direct traffic and lower levels of search traffic. Conspiracy ideas spread within homogeneous social media communities of like-minded believers who share conspiracy narratives; thus we also expect that traffic from social media (i.e., incoming traffic from a social media link) is higher in conspiracy compared to non-conspiracy websites. Moreover, because of known links between partisanship polarization and echo chambers (Stroud, 2010), confirmation bias (Westen et al., 2006), and belief in CTs (van Prooijen et al., 2015) we explored whether politically polarized websites (on both left and right sides of the spectrum) show patterns comparable to those of conspiracy websites compared to least biased websites.

We selected the websites (for which traffic data were collected) labeled as conspiracy ($N = 28$), least biased ($N = 15$), and pro-science ($N = 16$), and aggregated the websites leaning on either the left or right of the political spectrum, labeling them "biased_LR" ($N = 32$). Analysis of variance and post hoc comparisons using Tukey's honestly significant difference (HSD) test were used to test differences in traffic type between website categories. Direct traffic was highest for conspiracy ($M = 57.55$, $SD = 21.57$) and lowest for pro-science ($M = 13.35$, $SD = 12.12$), $F_{(3,87)} = 23.41$, $p < .001$. All post hoc differences among the four categories were significant at $p < .01$ except differences between least biased and biased websites ($p = .92$) and between pro-science and least biased websites ($p = .09$). As for traffic from search engines, the highest rate was on pro-science websites ($M = 70.80$, $SD = 14.80$) and the lower on conspiracy ones ($M = 13.82$, $SD = 10.44$), $F_{(3,87)} = 70.46$, $p < .001$. All differences were significant ($ps < .001$) except those between least biased and biased websites (p = .76). Incoming traffic from social media sites was higher in conspiracy ($M = 18.40$, $SD = 19.28$) than in pro-science ($M = 5.44$, $SD = 4.76$), $F_{(3,87)} = 4.93$, $p < .01$; all other differences were nonsignificant. Results are shown in Fig. 5.

These results suggest that CT websites are predominantly reached by the users typing the URL on their browser (or by recalling the URL from bookmarks) or following a link posted on social media. On the contrary, pro-science websites are mostly accessed from web searches. Differences in access routes between biased and least-biased websites were not significant. This indicates that though users of conspiracy websites are most similar to users of biased websites, they are nonetheless in a category of their own.

Discussion and conclusions

LOCO is a multilevel, richly annotated, topic-matched corpus of CTs composed of nearly 100,000 documents, with a total of 88 million words. This represents a rich source of data for better understanding the content and spread of CTs. LOCO is also freely available (https://osf.io/snpcg). Being for the

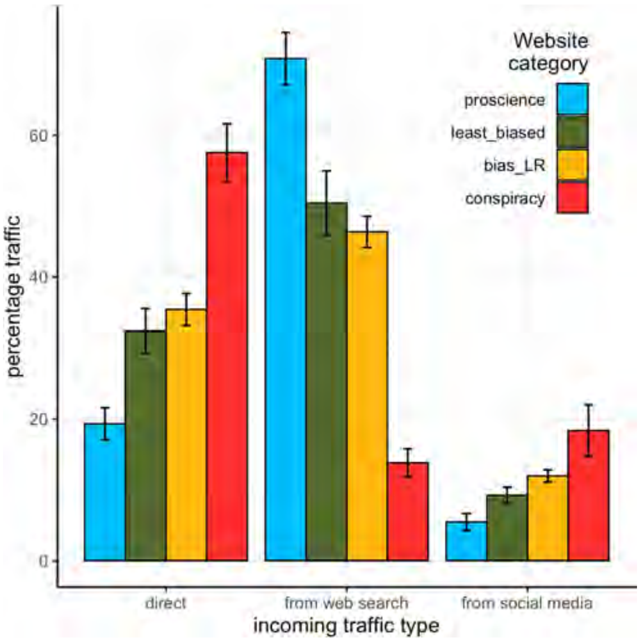


Fig. 5 Types of incoming traffic by website category. Average of websites' percentages of incoming traffic (direct, from web search, and from social media) by website categories. Error bars represent the standard error of the mean

most part composed of texts with additional metadata and lexical features, LOCO is conceptualized as a turnkey resource from which researchers can test hypotheses, further extract features, and/or build classification and predictive models. To this end, while building LOCO, we aimed at obtaining a large yet representative set of documents that also provided a set of metadata that could be used ad hoc to partition LOCO prior to analysis.

A large portion of the present paper has focused on thoroughly describing the methodology on which LOCO was built. As we have built LOCO upon previous works' strengths and weaknesses, we believe that a meticulous method description will also allow future research to benefit from LOCO's strengths and weaknesses, opening up possibilities for further data collection in the field of CT studies.

Our analysis of LOCO demonstrates its potential by making a number of contributions to the conspiracy research literature: (1) By mapping topics on document dates, we show that LOCO's documents track important social events. (2) We replicated the lexical analysis of previous work, finding an overlap between LOCO documents and comments on online social media. (3) We find that mainstream documents that mention conspiracy display conspiratorial language. (4) We have extracted and analyzed the language of prototypical conspiracy documents and find that these amplify features of conspiratorial language. (5) We find a pattern of website traffic indicating active online social media communities and the potential for confirmation bias via direct traffic. And (6) we find that conspiracy websites show statistically different patterns of web traffic than biased (politically left or right) websites, suggestive of a difference in their users. At the same time, we have provided suggestions on how to use LOCO to make new contributions.

Because we relied on a multitude of heterogeneous methods, we also believe that each of our corpus construction stages can benefit data collection for text analysis research in general. While we built LOCO on a specific narrative genre, namely CTs, the same methodology, or part of it, can be employed for other purposes. For example, researchers may be interested in comparing a list of websites against another one, or comparing webpages returned by specific sets of seeds, or, as we have done, do both at the same time by crossing lists of websites and seeds. We have also shown that it is possible to rely on several tools to enrich a web-based set of text with meta-data, such as political biases and fact accuracy (from MBFC), measures of spread (from SC), and popularity and traffic (from SW). Other freely available tools we have employed are available for text extraction (*Goose*) and analyses such as Empath, TAACO, as well as the *quanteda* and the *topicmodels* packages.

Because we also provided the URLs associated with each document, it is potentially possible to extract HTML data in order to analyze web markup features as previous work has done on fake news (Castelo et al., 2019). Moreover, different sets of psycholinguistic measures can be extracted from LOCO's texts, such as word norms for valence, arousal, and dominance (Warriner et al., 2013), imageability (Cortese & Fugett, 2004), frequency (Brysbaert & New, 2009), concreteness (Brysbaert et al., 2014), and age of acquisition (Kuperman et al., 2012).

In conclusion, LOCO is a rich source that helps to better understand the content of CTs. Here, we have explored how CT users behave online and which language features are associated with the spread of documents over social media, and we sketched a preliminary overview of the lexical fingerprint of the (prototypical) conspiratorial language. Therefore, LOCO's contribution is multiple: while providing data mainly for lexical analysis and document spread, it can also help to reveal psychological processes. For the sake of global public interest, given the detrimental potential consequences associated with the endorsement of CTs, understanding how CTs spread is critical to ultimately limiting their negative consequences.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13428-021-01698-z.

Funding Open Access funding provided by Université de Neuchâtel.

Declarations

Financial Disclosure None of the authors have any financial relationships relevant to this article to disclose.

References

Allen, C., & Murdock, J. (2020). *LDA Topic Modeling: Contexts for the History & Philosophy of Science*. http://philsci-archive.pitt.edu/17261/

Aston, G., & Burnard, L. (1998). The BNC Handbook: Exploring the British National Corpus with SARA. In: *English Language and Linguistics*. Edinburgh University Press.

Aupers, S. (2012). 'Trust no one': Modernization, paranoia and conspiracy culture. *European Journal of Communication*, *27*(1), 22–34. https://doi.org/10.1177/0267323111433566

AVAAZ. (2020). *Facebook's Algorithm: A Major Threat to Public Health*. https://secure.avaaz.org/campaign/en/facebook_threat_health/

Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting Factuality of Reporting and Bias of News Media Sources. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 3528–3539. https://doi.org/10.18653/v1/D18-1389

Bangerter, A., Wagner-Egger, P., & Delouvée, S. (2020). The Spread of Conspiracy Theories. In M. Butter & P. Knight (Eds.), *Routledge Handbook of Conspiracy Theories* (pp. 2016–2218). Routledge.

Barkun, M. (2017). President Trump and the "Fringe". *Terrorism and Political Violence*, *29*(3), 437–443. https://doi.org/10.1080/09546553.2017.1313649

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, *43*(3), 209–226. https://doi.org/10.1007/s10579-009-9081-4

Barron, A. T. J., Huang, J., Spang, R. L., & DeDeo, S. (2018). Individuals, institutions, and innovation in the debates of the French Revolution. *Proceedings of the National Academy of Sciences*, *115*(18), 4607–4612. https://doi.org/10.1073/pnas.1717729115

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bessi, A. (2016). Personality traits and echo chambers on facebook. *Computers in Human Behavior*, *65*, 319–324. https://doi.org/10.1016/j.chb.2016.08.016

Bessi, A., Scala, A., Rossi, L., Zhang, Q., & Quattrociocchi, W. (2014). The economy of attention in the age of (mis)information. *Journal of Trust Management*, *1*(1), 12. https://doi.org/10.1186/s40493-014-0012-y

Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015a). Science vs Conspiracy: Collective Narratives in the Age of Misinformation. *PLOS ONE*, *10*(2), e0118093. https://doi.org/10.1371/journal.pone.0118093

Bessi, A., Zollo, F., Del Vicario, M., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015b). Trend of Narratives in the Age of Misinformation. *PLOS ONE*, *10*(8), e0134641. https://doi.org/10.1371/journal.pone.0134641

Betsch, C., Ulshöfer, C., Renkewitz, F., & Betsch, T. (2011). The Influence of Narrative v. Statistical Information on Perceiving Vaccination Risks. *Medical Decision Making*, *31*(5), 742–753. https://doi.org/10.1177/0272989X11400419

Biddlestone, M., Green, R., & Douglas, K. M. (2020). Cultural orientation, power, belief in conspiracy theories, and intentions to reduce the spread of COVID-19. *British Journal of Social Psychology*, *59*(3), 663–673. https://doi.org/10.1111/bjso.12397

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc."

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bogart, L. M., Wagner, G., Galvan, F. H., & Banks, D. (2010). Conspiracy Beliefs About HIV Are Related to Antiretroviral Treatment Nonadherence Among African American Men With HIV. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *53*(5), 648–655. https://doi.org/10.1097/QAI.0b013e3181c57dbc

Brugnoli, E., Cinelli, M., Quattrociocchi, W., & Scala, A. (2019). Recursive patterns in online echo chambers. *Scientific Reports*, *9*(1), 20118. https://doi.org/10.1038/s41598-019-56191-7

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Butter, M., & Knight, P. (2020). *Routledge Handbook of Conspiracy Theories* (M. Butter & P. Knight (eds.)). Routledge.

Castelo, S., Santos, A., Almeida, T., Pham, K., Freire, J., Elghafari, A., & Nakamura, E. (2019). A topic-agnostic approach for identifying fake news pages. *The Web Conference 2019 - Companion of the World Wide Web Conference, WWW 2019*. https://doi.org/10.1145/3308560.3316739

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, *118*(9), e2023301118. https://doi.org/10.1073/pnas.2023301118

Clarke, S. (2007). Conspiracy Theories and the Internet: Controlled Demolition and Arrested Development. *Episteme*, *4*(2), 167–180. https://doi.org/10.3366/epi.2007.4.2.167

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 384–387. https://doi.org/10.3758/BF03195585

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, *48*(4), 1227–1237. https://doi.org/10.3758/s13428-015-0651-7

Crossley, S. A., Kyle, K., & Dascalu, M. (2019). The Tool for the Automatic Analysis of Cohesion 2.0: Integrating semantic similarity and text overlap. *Behavior Research Methods*, *51*(1), 14–27. https://doi.org/10.3758/s13428-018-1142-4

de Vries, E., Schoonvelde, M., & Schumacher, G. (2018). No Longer Lost in Translation: Evidence that Google Translate Works for Comparative Bag-of-Words Text Applications. *Political Analysis*, *26*(4), 417–430. https://doi.org/10.1017/pan.2018.26

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016a). The spreading of misinformation online. *Proceedings of the National Academy of*

Sciences, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016b). Echo Chambers: Emotional Contagion and Group Polarization on Facebook. *Scientific Reports*, *6*(1), 37825. https://doi.org/10.1038/srep37825

Del Vicario, M., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2017). Modeling confirmation bias and polarization. *Scientific Reports*, *7*(1), 40391. https://doi.org/10.1038/srep40391

Douglas, K. M., & Sutton, R. M. (2011). Does it take one to know one? Endorsement of conspiracy theories is influenced by personal willingness to conspire. *British Journal of Social Psychology*, *50*(3), 544–552.

Douglas, K. M., & Sutton, R. M. (2018). Why conspiracy theories matter: A social psychological analysis. *European Review of Social Psychology*, *29*(1), 256–298. https://doi.org/10.1080/10463283.2018.1537428

Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding Conspiracy Theories. *Political Psychology*, *40*(S1), 3–35. https://doi.org/10.1111/pops.12568

Eicher, V., & Bangerter, A. (2015). Social representations of infectious diseases. In G. Sammut, E. Andreouli, G. Gaskell, & J. Valsiner (Eds.), *The Cambridge Handbook of Social Representations* (pp. 385–396). Cambridge University Press. https://doi.org/10.1017/CBO9781107323650.031

Einstein, K. L., & Glick, D. M. (2015). Do I Think BLS Data are BS? The Consequences of Conspiracy Theories. *Political Behavior*, *37*(3), 679–701. https://doi.org/10.1007/s11109-014-9287-z

Faasse, K., Chatman, C. J., & Martin, L. R. (2016). A comparison of language use in pro- and anti-vaccination comments in response to a high profile Facebook post,. *Vaccine*, *34*(47), 5808–5814. https://doi.org/10.1016/j.vaccine.2016.09.029

Fast, E., Chen, B., & Bernstein, M. S. (2016). Empath. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4647–4657. https://doi.org/10.1145/2858036.2858535

Fong, A., Roozenbeek, J., Goldwert, D., Rathje, S., & van der Linden, S. (2021). The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations*, *24*(4), 606–623. https://doi.org/10.1177/1368430220987596

Franks, B., Bangerter, A., & Bauer, M. W. (2013). Conspiracy theories as quasi-religious mentality: an integrated account from cognitive science, social representations theory, and frame theory. *Frontiers in Psychology*, *4*, 1–12. https://doi.org/10.3389/fpsyg.2013.00424

Franks, B., Bangerter, A., Bauer, M. W., Hall, M., & Noort, M. C. (2017). Beyond "Monologicality"? Exploring Conspiracist Worldviews. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00861

Fry, E. (2000). *1000 instant words: the most common words for teaching reading, writing and spelling*. Teacher Created Resources.

Fu, L. Y., Zook, K., Spoehr-Labutta, Z., Hu, P., & Joseph, J. G. (2016). Search Engine Ranking, Quality, and Content of Web Pages That Are Critical Versus Noncritical of Human Papillomavirus Vaccine. *Journal of Adolescent Health*, *58*(1), 33–39. https://doi.org/10.1016/j.jadohealth.2015.09.016

Golec de Zavala, A., & Cichocka, A. (2012). Collective narcissism and anti-Semitism in Poland. *Group Processes & Intergroup Relations*, *15*(2), 213–229. https://doi.org/10.1177/1368430211420891

Grün, B., & Hornik, K. (2011). Topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, *40*(13), 1–30. https://doi.org/10.18637/jss.v040.i13

Guerini, M., Giampiccolo, D., Moretti, G., Sprugnoli, R., & Strapparava, C. (2013). The New Release of CORPS: A Corpus of Political Speeches Annotated with Audience Reactions. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 86–98). https://doi.org/10.1007/978-3-642-41545-6_8

Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology*, *81*(6), 1028–1041. https://doi.org/10.1037/0022-3514.81.6.1028

Hills, T. T. (2019). The Dark Side of Information Proliferation. *Perspectives on Psychological Science*, *14*(3), 323–330. https://doi.org/10.1177/1745691618803647

Imhoff, R., Lamberty, P., & Klein, O. (2018). Using Power as a Negative Cue: How Conspiracy Mentality Affects Epistemic Trust in Sources of Historical Knowledge. *Personality and Social Psychology Bulletin*, *44*(9), 1364–1379. https://doi.org/10.1177/0146167218768779

Imhoff, R., Dieterle, L., & Lamberty, P. (2021). Resolving the Puzzle of Conspiracy Worldview and Political Activism: Belief in Secret Plots Decreases Normative but Increases Nonnormative Political Engagement. *Social Psychological and Personality Science*, *12*(1), 71–79. https://doi.org/10.1177/1948550619896491

Jensen, T. (2013). *Democrats and Republicans differ on conspiracy theory beliefs*. Public Policy Polling. http://www.publicpolicypolling.com/polls/democrats-and-republicans-differ-on-conspiracy-theory-beliefs

Jolley, D., & Douglas, K. M. (2014a). The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint. *British Journal of Psychology*, *105*(1), 35–56. https://doi.org/10.1111/bjop.12018

Jolley, D., & Douglas, K. M. (2014b). The Effects of Anti-Vaccine Conspiracy Theories on Vaccination Intentions. *PLoS ONE*, *9*(2), e89177. https://doi.org/10.1371/journal.pone.0089177

Jolley, D., & Paterson, J. L. (2020). Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *British Journal of Social Psychology*, *59*(3), 628–640. https://doi.org/10.1111/bjso.12394

Jolley, D., Douglas, K. M., Leite, A. C., & Schrader, T. (2019). Belief in conspiracy theories and intentions to engage in everyday crime. *British Journal of Social Psychology*, *58*(3), 534–549. https://doi.org/10.1111/bjso.12311

Klein, C., Clutton, P., & Polito, V. (2018). Topic Modeling Reveals Distinct Interests within an Online Conspiracy Forum. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00189

Klein, C., Clutton, P., & Dunn, A. G. (2019). Pathways to conspiracy: The social and linguistic precursors of involvement in Reddit's conspiracy theory forum. *PLOS ONE*, *14*(11), e0225098. https://doi.org/10.1371/journal.pone.0225098

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Kwon, S., Cha, M., & Jung, K. (2017). Rumor Detection over Varying Time Windows. *PLOS ONE*, *12*(1), e0168344. https://doi.org/10.1371/journal.pone.0168344

Kyle, K., Crossley, S. A., & Kim, Y. J. (2015). Native language identification and writing proficiency. *International Journal of Learner Corpus Research*, *1*(2), 187–209. https://doi.org/10.1075/ijlcr.1.2.01kyl

Lansdall-Welfare, T., Sudhahar, S., Thompson, J., Lewis, J., & Cristianini, N. (2017). Content analysis of 150 years of British periodicals. *Proceedings of the National Academy of Sciences*, *114*(4), E457–E465. https://doi.org/10.1073/pnas.1606380114

Lantian, A., Muller, D., Nurra, C., Klein, O., Berjot, S., & Pantazi, M. (2018). Stigmatized beliefs: Conspiracy theories, anticipated negative evaluation of the self, and fear of social exclusion. *European Journal of Social Psychology*, *48*(7), 939–954. https://doi.org/10.1002/ejsp.2498

Lazarus, J. V., Ratzan, S. C., Palayew, A., Gostin, L. O., Larson, H. J., Rabin, K., Kimball, S., & El-Mohandes, A. (2020). A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine*. https://doi.org/10.1038/s41591-020-1124-9

Li, Y., Engelthaler, T., Siew, C. S. Q., & Hills, T. T. (2019). The Macroscope: A tool for examining the historical structure of language. *Behavior Research Methods*, *51*(4), 1864–1877. https://doi.org/10.3758/s13428-018-1177-6

Li, Y., Hills, T., & Hertwig, R. (2020). A brief history of risk. *Cognition*, *203*, 104344. https://doi.org/10.1016/j.cognition.2020.104344

Marchlewska, M., Cichocka, A., & Kossowska, M. (2018). Addicted to answers: Need for cognitive closure and the endorsement of conspiracy beliefs. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2308

Meppelink, C. S., Smit, E. G., Fransen, M. L., & Diviani, N. (2019). "I was Right about Vaccination": Confirmation Bias and Health Literacy in Online Health Information Seeking. *Journal of Health Communication*, *24*(2), 129–140. https://doi.org/10.1080/10810730.2019.1583701

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, *331*(6014), 176–182. https://doi.org/10.1126/science.1199644

Mitra, T., Counts, S., & Pennebaker, J. W. (2016). Understanding anti-vaccination attitudes in social media. *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*.

Nguyen, D., Liakata, M., DeDeo, S., Eisenstein, J., Mimno, D., Tromble, R., & Winters, J. (2020). How We Do Things With Words: Analyzing Text as Social and Cultural Data. *Frontiers in Artificial Intelligence*, *3*. https://doi.org/10.3389/frai.2020.00062

Nikolov, D., Oliveira, D. F. M., Flammini, A., & Menczer, F. (2015). Measuring online social bubbles. *PeerJ Computer Science*, *1*(12), e38. https://doi.org/10.7717/peerj-cs.38

Okuhara, T., Ishikawa, H., Okada, M., Kato, M., & Kiuchi, T. (2017). Readability comparison of pro- and anti-HPV-vaccination online messages in Japan. *Patient Education and Counseling*, *100*(10), 1859–1866. https://doi.org/10.1016/j.pec.2017.04.013

Oliver, J. E., & Wood, T. J. (2014). Conspiracy Theories and the Paranoid Style(s) of Mass Opinion. *American Journal of Political Science*, *58*(4), 952–966. https://doi.org/10.1111/ajps.12084

Ooms, J. (2019). *curl: A Modern and Flexible Web Client for R*. https://cran.r-project.org/package=curl

Oswald, S. (2016). Conspiracy and bias: argumentative features and persuasiveness of conspiracy theories. *OSSA Conference Archive*, *168*, 1–16.

Pauwels, K., Demirci, C., Yildirim, G., & Srinivasan, S. (2016). The impact of brand familiarity on online and offline media synergy. *International Journal of Research in Marketing*, *33*(4), 739–753. https://doi.org/10.1016/j.ijresmar.2015.12.008

Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, *116*(7), 2521–2526. https://doi.org/10.1073/pnas.1806781116

Perez, J. C., & Montagnier, L. (2020). Covid-19, Sars And Bats Coronaviruses Genomes Peculiar Homologous RNA Sequences. *International Journal of Research -GRANTHAALAYAH*, *8*(7), 217–263. https://doi.org/10.29121/granthaalayah.v8.i7.2020.678

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Raab, M. H., Auer, N., Ortlieb, S. A., & Carbon, C.-C. (2013a). The Sarrazin effect: the presence of absurd statements in conspiracy theories makes canonical information less plausible. *Frontiers in Psychology*, *4*, 1–8. https://doi.org/10.3389/fpsyg.2013.00453

Raab, M. H., Ortlieb, S. A., Auer, N., Guthmann, K., & Carbon, C.-C. (2013b). Thirty shades of truth: conspiracy theories as stories of individuation, not of pathological delusion. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00406

Risius, M., Aydinguel, O., & Haug, M. (2019). Towards an understanding of conspiracy echo chambers on Facebook. *Proceedings of the 27th European Conference on Information Systems (ECIS)*. https://aisel.aisnet.org/ecis2019_rip/36

Sak, G., Diviani, N., Allam, A., & Schulz, P. J. (2015). Comparing the quality of pro- and anti-vaccination online information: a content analysis of vaccination-related webpages. *BMC Public Health*, *16*(1), 38. https://doi.org/10.1186/s12889-016-2722-9

Salmon, D. A., Moulton, L. H., Omer, S. B., DeHart, M. P., Stokley, S., & Halsey, N. A. (2005). Factors associated with refusal of childhood vaccines among parents of school-aged children: a case-control study. *Archives of Pediatrics & Adolescent Medicine*, *159*(5), 470–476. https://doi.org/10.1001/archpedi.159.5.470

Samory, M., & Mitra, T. (2018a). Conspiracies online: User discussions in a conspiracy community following dramatic events. *12th International AAAI Conference on Web and Social Media, ICWSM 2018*.

Samory, M., & Mitra, T. (2018b). "The Government Spies Using Our Webcams:" The language of conspiracy theories in online discussions. *Proceedings of the ACM on Human-Computer Interaction*, *2*(152). https://doi.org/10.1145/3274421

Smith, N., & Graham, T. (2019). Mapping the anti-vaccination movement on Facebook. *Information, Communication & Society*, *22*(9), 1310–1327. https://doi.org/10.1080/1369118X.2017.1418406

Sternisko, A., Cichocka, A., & Van Bavel, J. J. (2020). The dark side of social movements: social identity, non-conformity, and the lure of conspiracy theories. *Current Opinion in Psychology*, *35*, 1–6. https://doi.org/10.1016/j.copsyc.2020.02.007

Stroud, N. J. (2010). Polarization and Partisan Selective Exposure. *Journal of Communication*, *60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x

Swami, V., Barron, D., Weis, L., & Furnham, A. (2018). To Brexit or not to Brexit: The roles of Islamophobia, conspiracist beliefs, and integrated threat in voting intentions for the United Kingdom European Union membership referendum. *British Journal of Psychology*, *109*(1), 156–179. https://doi.org/10.1111/bjop.12252

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Uscinski, J. E., Parent, J. M., & Torres, B. (2011). Conspiracy Theories Are for Losers. *APSA 2011 Annual Meeting Paper*. https://ssrn.com/abstract=1901755

Uscinski, J. E., DeWitt, D., & Atkinson, M. D. (2018). A Web of Conspiracy? Internet and Conspiracy Theory. In: *Handbook of Conspiracy Theory and Contemporary Religion* (pp. 106–130). BRILL. https://doi.org/10.1163/9789004382022_007

van der Linden, S. (2015). The conspiracy-effect: Exposure to conspiracy theories (about global warming) decreases pro-social behavior and science acceptance. *Personality and Individual Differences*, *87*, 171–173. https://doi.org/10.1016/j.paid.2015.07.045

van Prooijen, J.-W., Krouwel, A. P. M., & Pollet, T. V. (2015). Political Extremism Predicts Belief in Conspiracy Theories. *Social Psychological and Personality Science*, *6*(5), 570–578. https://doi.org/10.1177/1948550614567356

von Luxburg, U., Williamson, R. C., & Guyon, I. (2012). Clustering: Science or Art? In I. Guyon, G. Dror, V. Lemaire, G. Taylor, & D. Silver (Eds.), *Proceedings of ICML Workshop on Unsupervised and*

Transfer Learning (Vol. 27, pp. 65–79). PMLR. http://proceedings.mlr.press/v27/luxburg12a.html

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

Wakefield, A., Murch, S., Anthony, A., Linnell, J., Casson, D., Malik, M., Berelowitz, M., Dhillon, A., Thomson, M., Harvey, P., Valentine, A., Davies, S., & Walker-Smith, J. (1998). RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, *351*(9103), 637–641. https://doi.org/10.1016/S0140-6736(97)11096-0

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Westen, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural Bases of Motivated Reasoning: An fMRI Study of Emotional Constraints on Partisan Political Judgment in the 2004 U.S. Presidential Election. *Journal of Cognitive Neuroscience*, *18*(11), 1947–1958. https://doi.org/10.1162/jocn.2006.18.11.1947

Wood, M. J. (2018). Propagating and Debunking Conspiracy Theories on Twitter During the 2015–2016 Zika Virus Outbreak. *Cyberpsychology, Behavior, and Social Networking*, *21*(8), 485–490. https://doi.org/10.1089/cyber.2017.0669

Wood, M. J., & Douglas, K. M. (2013). What about building 7? A social psychological study of online discussion of 9/11 conspiracy theories. *Frontiers in Psychology*, *4*, 1–9. https://doi.org/10.3389/fpsyg.2013.00409

Wood, M. J., & Douglas, K. M. (2015). Online communication as a window to conspiracist worldviews. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00836

Zannettou, S., Bradlyn, B., De Cristofaro, E., Kwak, H., Sirivianos, M., Stringini, G., & Blackburn, J. (2018). What is Gab. *Companion of the Web Conference 2018 on The Web Conference 2018 - WWW '18*, 1007–1014. https://doi.org/10.1145/3184558.3191531

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

Zubiaga, A., Liakata, M., Procter, R., Wong Sak Hoi, G., & Tolmie, P. (2016). Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. *PLOS ONE*, *11*(3), e0150989. https://doi.org/10.1371/journal.pone.0150989

LOCO: the 88-million-word language of conspiracy corpus

SUPPLEMENTAL MATERIAL

Alessandro Miani[1], Thomas Hills[2,3], and Adrian Bangerter[1]

[1] *Institute of Work and Organizational Psychology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland*
[2] *Department of Psychology, University of Warwick, University Road, Coventry CV47AL, United Kingdom*
[3] *The Alan Turing Institute, British Library, 96 Euston Road, London, NW1 2DB, United Kingdom*

Contents

SM 1 Evaluation of conspiracy and mainstream labels

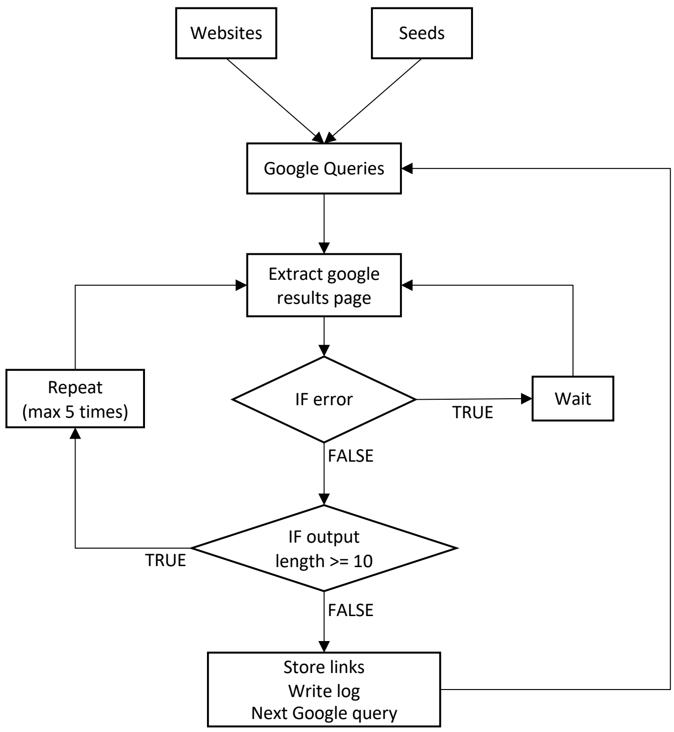
In order to estimate to what extent conspiracy and mainstream documents reflect their true labels, namely whether a document is in fact either conspiracy or mainstream, we have visually inspected and classified a subset of documents. We extracted a random sample of 60 conspiracy and 60 mainstream documents. One of the authors (AM) manually coded these documents either as conspiracy or mainstream. Such coding was performed blind, meaning that only unformatted text with no other information was available to the coder (besides a newly generated random ID to allow documents to match with true labels after the manual coding). Results show that ratings had an overall accuracy of .88 (95%CIs = .81 - .93, Cohen's k = .77), with a sensitivity (true positive rate: conspiracy correctly labelled as conspiracy; 9 misclassified) of .85 and specificity (true negative rate: mainstream correctly labelled as mainstream; 5 misclassified) of .92.

SM 2 URL extraction workflow

The URL extraction was performed via the following workflow (see Figure S1):

1. Aggregate seeds and websites into a unique query
2. Send query to Google and extract links from the first page
3. If error, wait (to avoid the "HTTP 429 too many requests" error) and repeat (no more than five times)
4. If more than 10 results are returned, go to the next page; if error, wait
5. If less than 10 results are returned, store results and move to the next query

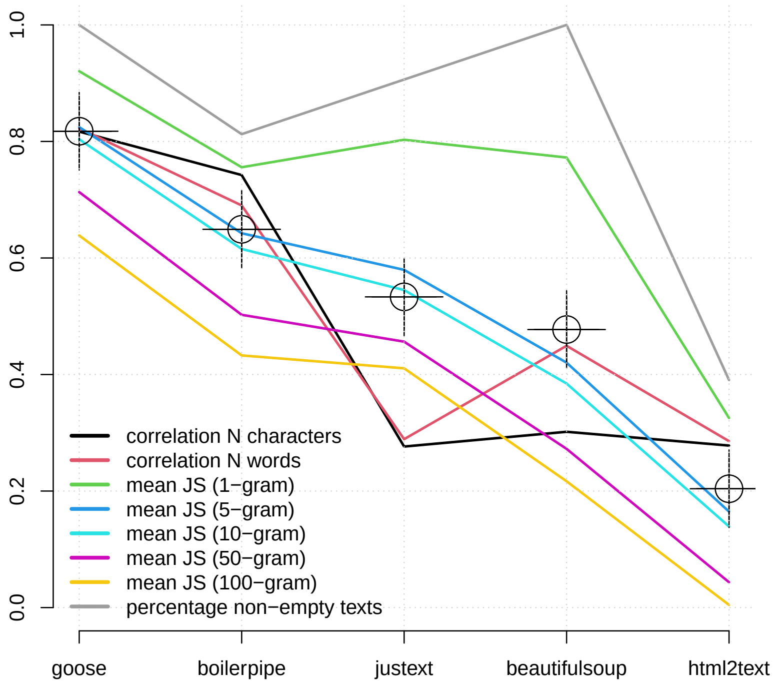
Figure S1 – URL extraction workflow



SM 3 Testing boilerplate stripping

In order to find the most reliable script (*i.e.*, the one that returns a cleaned text most similar to manual cleaning), we tested five different Python packages: *beautifulsoup*, *boilerpipe*, *Goose*, *HTML2text*, and *justext*. To test these scripts, we visually inspected a set of 100 random links and, when possible, we manually extracted the useful portions of text (*i.e.*, copying the text and pasting it to a new file). This was done on 64 webpages from 39 unique websites. Then, we ran each of the five scripts along the URL list. Doing so, we obtained six versions of the text per each webpage (one from manual extraction, and five from the scripts). We then tested how each script was similar to the manually extracted text on different metrics. These were: a) the correlation of number of characters and number of words, b) Jaccard similarity (*i.e.*, 1-distance) on different levels of ngrams (1, 5, 10, 50, 100), and c) the percentage of non-empty texts returned by the script. Per each set of metrics, we extracted values on the individual text-level and then we averaged them into a single value per metric per boilerplate. The Goose package (`https://github.com/grangier/python-goose`) returned the best performance (highest aggregated mean of all metrics, see Figure S2) and therefore was chosen for extracting the texts.

Figure S2 – Boilerplate stripping test results



Note. Comparisons for each of the 5 stripping packages we tested on the 8 different metrics described in the text. The circled crosses represent the mean of the 8 metrics within stripping package. JS: Jaccard similarity.

SM 4 LIWC and Empath correlations

Here we report the top 20 (positive and negative) highest and lowest absolute correlations between LIWC and Empath dictionaries. Correlations have been run at the document level (N = 96,743). See Table S1 below.

Table S1 – LIWC and Empath correlations

	highest			*lowest (abs)*			*lowest*	
LIWC	**Empath**	**r**	**LIWC**	**Empath**	**r**	**LIWC**	**Empath**	**r**
money	valuable	0.98	article	disappointment	0	Tone	violence	-0.58
money	payment	0.98	Colon	pet	0	body	leader	-0.58
health	health	0.98	discrep	aggression	0	tentat	leader	-0.58
money	banking	0.98	drives	lust	0	Tone	weapon	-0.59
money	money	0.97	netspeak	government	0	percept	gain	-0.59
family	family	0.97	Comma	exercise	0	body	government	-0.59
money	economics	0.97	swear	terrorism	0	Tone	suffering	-0.60
health	medical_emergency	0.97	space	fight	0	percept	government	-0.60
relig	worship	0.97	reward	noise	0	cause	party	-0.60
family	children	0.95	Colon	cooking	0	shehe	science	-0.61
money	wealthy	0.94	Dash	fire	0	number	speaking	-0.62
body	body	0.94	shehe	ugliness	0	male	science	-0.62
relig	religion	0.93	number	leisure	0	cause	home	-0.62
relig	divine	0.93	SemiC	banking	0	Analytic	confusion	-0.62
money	negotiate	0.92	health	restaurant	0	Sixltr	friends	-0.62
death	kill	0.91	social	hipster	0	cause	wedding	-0.63
work	business	0.91	female	anonymity	0	Sixltr	childish	-0.63
home	domestic_work	0.91	they	dispute	0	Analytic	speaking	-0.66
bio	health	0.90	nonflu	weather	0	Tone	aggression	-0.67
leisure	fun	0.90	ipron	death	0	Tone	kill	-0.68

SM 5 Text preprocessing

Text preprocessing for topic extraction relied on the corpus' document-term matrix (DTM), namely a bag of words for the whole corpus that stores the occurrences of each word (columns) for each document (rows). Before generating the DTM, it was needed to preprocess each text so to remove noise, reduce sparsity, while reducing computation time. This preprocessing was mostly done by removing the most frequent (*e.g.*, stopwords) and infrequent (*e.g.*, misspellings or extreme rare) words.

5.1 Identifying stopwords

Before running the cleaning pipeline, we first identified a set of stopwords suitable for our purpose. We started by merging the 175 English stopwords from the *stopwords* R package (Benoit et al., 2020) with the top 100 most frequent words in English (Fry, 2000) and single letters (N = 26). This resulted in a set of 229 words (see Table S2 below). Motivated from the literature, from this set, we decide to remove (*e.g.*, include in DTM) the following terms:

1. Pronouns: in line with research on social motives underlying belief in CTs (Douglas et al., 2019), the conspiratorial language heavily relies on pronouns such as we/they marking a process of social identification (we) by exclusion from the outgroup (they). This list (of pronouns) was compiled from the LIWC category pronouns.

2. Negations and interrogative words: crucial for the rhetoric of conspiracy (Oswald, 2016) based on refutation and questioning. Negations were extracted from the LIWC category negations, while wh-question words were manually excluded by looking at the remaining list of stopwords.

3. Ad hoc set of words (within the stopwords) that might be important for conspiracy language. This was done by visually inspecting the list of stopwords and removing those potentially important such as against (language of fighting), *oil* (related to economy), *people* (see *we/they* above) as well as *see* and *write* (reporting facts).

Note that we also removed stopwords that included contractions (*e.g.*, *I'm*, *you're*) because in in the text preprocessing we expanded these contractions. The reasoning to expand contractions is that by keeping them, the term *they're* is treated as a different term than *they.* Because we were interested in keeping pronouns, expanding contractions allows us to keep them in the DTM. Furthermore, we have no theoretical reason to prefer a contracted vs an expanded form. Contractions were expanded by using the dataset contractions (N = 70 entries) from the R package *qdapDictionaries* (Rinker, 2013).

5.2 Cleaning pipeline

Once the list of stopwords was compiled, we proceeded with the text cleaning pipeline. This was done by using the *quanteda* R package (Benoit et al., 2018). The pipeline was as follow:

1. Remove non-ASCII characters

2. Tolower (*i.e.*, convert upper case to lower case)

3. Tokenization
   (a) remove URLs
   (b) remove punctuation
   (c) remove numbers
   (d) remove separators
   (e) split hyphens
   (f) remove symbols
4. Expand contractions
5. Remove stopwords (see section SM 5.1)
6. Wordstem (words were reduced to their root, *e.g.*, "frequenc" for "frequency" and "frequencies")
7. Generate DTM
8. Trim the DTM to top 10,000 terms

The DTM was trimmed to the top 10,000 terms so to reduce sparsity and computation time. This reduced sparsity from 0.999 to 0.974. The DTM was finally composed of 96,743 documents and 10,000 terms, for a total of 47,523,008 types (without trimming, the DTM was 349,063 terms accounting for 50,547,244 types).

Table S2 – Stopword sets

Initial set (N = 229)	a, about, above, after, again, against, all, am, an, and, any, are, aren't, as, at, b, be, because, been, before, being, below, between, both, but, by, c, call, can, can't, cannot, come, could, couldn't, d, day, did, didn't, do, does, doesn't, doing, don't, down, during, e, each, f, few, find, first, for, from, further, g, get, go, h, had, hadn't, has, hasn't, have, haven't, having, he, he'd, he'll, he's, her, here, here's, hers, herself, him, himself, his, how, how's, i, i'd, i'll, i'm, i've, if, in, into, is, isn't, it, it's, its, itself, j, k, l, let's, like, long, look, m, made, make, many, may, me, more, most, mustn't, my, myself, n, no, nor, not, now, number, o, of, off, oil, on, once, one, only, or, other, ought, our, ours, ourselves, out, over, own, p, part, people, q, r, s, said, same, see, shan't, she, she'd, she'll, she's, should, shouldn't, so, some, such, t, than, that, that's, the, their, theirs, them, themselves, then, there, there's, these, they, they'd, they'll, they're, they've, this, those, through, time, to, too, two, u, under, until, up, use, v, very, w, was, wasn't, water, way, we, we'd, we'll, we're, we've, were, weren't, what, what's, when, when's, where, where's, which, while, who, who's, whom, why, why's, will, with, won't, word, would, wouldn't, write, x, y, you, you'd, you'll, you're, you've, your, yours, yourself, yourselves, z
To remove (N = 53)	**Pronouns**: i, me, my, myself, we, let, us, our, ours, ourselves, you, thou, thoust, thee, thy, thyself, your, yours, yourself, yourselves, she, he, her, hers, his, him, himself, hisself, herself, their, theirs, them, themself, themselves, they. **Negations and wh-question words**: not, cannot, no, nor, nope, none, what, when, where, which, who, how, why. **Ad hoc words**: against, oil, people, see, write
Final set (N = 137)	a, about, above, after, again, all, am, an, and, any, are, as, at, b, be, because, been, before, being, below, between, both, but, by, c, call, can, come, could, d, day, did, do, does, doing, down, during, e, each, f, few, find, first, for, from, further, g, get, go, h, had, has, have, having, here, if, in, into, is, it, its, itself, j, k, l, like, long, look, m, made, make, many, may, more, most, n, now, number, o, of, off, on, once, one, only, or, other, ought, out, over, own, p, part, q, r, s, said, same, should, so, some, such, t, than, that, the, then, there, these, this, those, through, time, to, too, two, u, under, until, up, use, v, very, w, was, water, way, were, while, whom, will, with, word, would, x, y, z

SM 6 Topic description

We extracted three sets of topic distributions based on different k (*i.e.*, number of topics) resolutions, namely $k = 100$, $k = 200$, and $k = 300$. As a consequence, LOCO is provided with a total of 600 topics. Besides the actual matrix with the gamma vales for each topic per each document (`topic_gamma.json`), we also attach to LOCO a dataset (with 600 entries, namely the number of topics) containing a description of each topic. Below, in Table S3, an excerpt of the file `topic_gamma.json`. Descriptions include:

1. **Topic ID**: the topic indexed as the sequential topic ID preceded by the k value, so for example "k100_4" refers to the fourth topic extracted with $k = 100$
2. **Top 15 words**: the top 15 words of the topic ordered by decreasing beta weight (see Figure S3)
3. **N (C/M)**: The total (N) number of documents whose such topic has the max gamma value across all topics within a k set. N_C and N_M refer to the number of documents within conspiracy and mainstream subcorpora. The variable "Prop conspiracy" refers to the proportion of conspiracy documents.
4. **Topic in K r**: It reports the topic with the highest correlation among all topics within the k set. The correlation is computed on the document level (N = 96,743).
5. **Topic all r**: same as above, but correlation is computed for all topics for all k sets (*i.e.*, N = 600) on the document level (N = 96,743).
6. **LF r**: the highest correlation with lexical features (LF) from Empath (E) and LIWC (L).

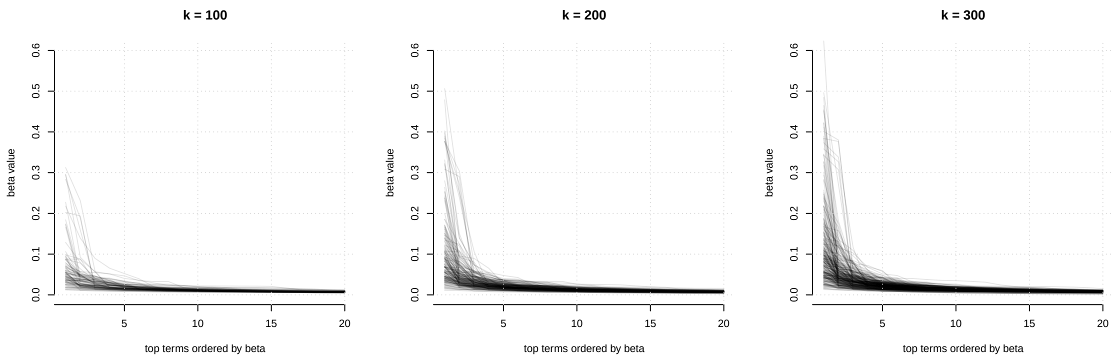
Table S3 – Excerpt from data frame `topic_description.json`

Topic ID	Top 15 words	N Total	N C	N M	Prop conspiracy	Topic In K	Topic In K r	Topic all	Topic all r	LF	LF r
k100_1	*studi, risk, increas, report, associ, case, group, age, data, effect, rate, result, factor, not, includ*	672	53	619	0.079	k100_42	0.15	k200_1	0.82	E_health	0.24
k100_2	*game, team, play, player, sport, world, year, win, event, their, club, who, season, match, footbal*	842	27	815	0.032	k100_89	0.12	k200_179	0.93	E_sports	0.42
k200_1	*studi, data, age, associ, rate, report, case, estim, among, popul, use, compar, differ, group, incid*	436	26	410	0.060	k200_196	0.14	k100_1	0.82	L_Parenth	0.22
k200_2	*gate, bill, foundat, world, microsoft, melinda, his, billion, billionair, founder, global, million, wealth, who, year*	766	157	609	0.205	k200_170	0.11	k300_294	0.85	L_home	0.32
k300_1	*inform, provid, requir, applic, avail, includ, not, servic, section, must, follow, assess, addit, date, request*	289	18	271	0.062	k300_300	0.11	k200_44	0.82	E_office	0.22
k300_2	*fluorid, teeth, dental, drink, tooth, health, decay, toothpast, dentist, children, suppli, brush, caviti, level, effect*	819	188	631	0.230	k300_122	0.14	k200_13	0.98	E_hygiene	0.57

6.1 Top term beta values

Note that we do not provide labels for topics. Instead, as it can be observed from the plots in Figure S3 below (and as suggested in Nguyen et al. 2020), it is possible to infer the topic content by inspecting the top 5 terms in each topic, which account for the majority of beta weights. In other words, each topic is mostly characterized by the top 5 terms while the contribution of the other terms to the topic is minimal.

Figure S3 – Distribution of beta weights across ks



Note. Each line represents the ordered (decreasing from left to right) beta values for the top 20 terms for one of the k topics (from 100, left, to 300, right).

SM 7 Comparison across 3 ks

As mentioned in the main text (see section 3.6), we extracted three sets of topic distributions based on different k (*i.e.*, number of topics) resolutions, namely $k = 100$, $k = 200$, and $k = 300$. For illustrative purposes (not exhaustive), here, we selected a small set of themes to explore the extent to which similar topics (*i.e.*, topics revolving around the same theme, *e.g.*, Lady Diana's death) correlate, at the document level, on different levels of k resolution. For each theme, we selected a keyword capable of retrieving the associated topic in each k dataset by searching the keyword within the top 15 topic terms. For example, the keyword "mh370" retrieves the topics associated with the Malaysia Airlines Flight 370 that disappeared on 8 March 2014.

The set of keywords (associated with topic) we used was: *diana* (Lady Diana), *sandi* (Sandy Hook school shooting), *tower* (9/11 terroristic attack), *ebola* (epidemic), *mh370* (flight disappeared), *laden* (Osama Bin Laden), *jackson* (Michael Jackson), corona (covid-19 epidemic), *zika* (epidemic), *saddam* (Saddam Hussein), *elvi* (Elvis Presley), *5g* (5G communication technology).

In the following, for each theme, we present a descriptive table followed by a correlation matrix. In the descriptive table, each line describes the topics that revolve around the theme (see description of columns in section SM 6). For example, in Table S4 that contains topics revolving around Lady Diana, the two topics in k200 correlate with each other at r=.17, these are the highest correlations for the topics 107 and 189 within the k200 dataset. Similarly, within the k200 dataset, topics 107 and 189 correlate the most with the categories of *car* and *royalty* (from the Empath dictionary), respectively, suggesting that while topic 107 refers to Lady Diana's death (car accident), the topic 189 refers to the British Royal Family, possibly her life.

Table S4 – Topic description: Lady Diana

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_88	*diana, princ, royal, princess, her, queen, harri, british, william, charl, death, famili, their, london, his*	168	1014	Topic_027	0.17	E_royalty	0.65
k200_107	*car, paul, crash, diana, driver, drive, accid, vehicl, fay, death, polic, pari, who, they, french*	99	293	Topic_189	0.17	E_car	0.36
k200_189	*diana, princ, royal, her, queen, harri, princess, charl, william, british, famili, elizabeth, palac, death, she*	79	727	Topic_107	0.17	E_royalty	0.7
k300_65	*diana, princ, princess, her, harri, royal, charl, death, william, fay, crash, she, british, palac, car*	125	827	Topic_258	0.16	E_royalty	0.55

Table S5 – Topic description: Lady Diana (correlation matrix)

	k100_88	k200_107	k200_189	k300_65
k100_88	1			
k200_107	0.56	1		
k200_189	0.88	0.17	1	
k300_65	0.93	0.63	0.78	1

Table S6 – Topic description: Sandy Hook

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_97	*gun, shoot, polic, school, kill, hook, sandi, violenc, mass, who, shot, victim, year, children, peopl*	282	1143	Topic_043	0.12	E_weapon	0.6
k200_141	*gun, shoot, school, hook, sandi, kill, mass, violenc, lanza, shot, polic, connecticut, newtown, massacr, victim*	243	1037	Topic_096	0.14	E_weapon	0.57
k300_290	*gun, shoot, school, hook, sandi, kill, mass, violenc, lanza, connecticut, newtown, massacr, elementari, victim, firearm*	187	946	Topic_090	0.15	E_weapon	0.55

Table S7 – Topic description: Sandy Hook (correlation matrix)

	k100_97	k200_141	k300_290
k100_97	1		
k200_141	0.96	1	
k300_290	0.94	0.97	1

Table S8 – Topic description: 9/11 terroristic attack

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_79	*attack, new, world, center, tower, build, york, septemb, trade, fire, who, event, peopl, terrorist, explos*	267	691	Topic_061	0.09	E_terrorism	0.22
k200_6	*attack, tower, center, septemb, trade, build, world, hijack, terrorist, york, new, pentagon, collaps, sept, plane*	276	516	Topic_169	0.2	E_terrorism	0.21
k300_72	*attack, tower, center, trade, world, septemb, new, york, pentagon, build, wtc, who, twin, ground, plane*	127	517	Topic_236	0.23	E_terrorism	0.16

Table S9 – Topic description: 9/11 terroristic attack (correlation matrix)

	k100_79	k200_6	k300_72
k100_79	1		
k200_6	0.92	1	
k300_72	0.89	0.92	1

Table S10 – Topic description: Ebola

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_37	*ebola, outbreak, health, who, case, diseas, peopl, west, virus, spread, africa, epidem, countri, emerg, congo*	211	1570	Topic_091	0.1	E_medical_emergency	0.23
k200_159	*ebola, outbreak, health, who, case, diseas, virus, peopl, west, spread, congo, epidem, africa, worker, infect*	197	1471	Topic_115	0.12	E_medical_emergency	0.21
k300_62	*ebola, outbreak, health, virus, diseas, congo, peopl, case, west, spread, who, africa, epidem, worker, liberia*	174	1412	Topic_078	0.24	E_medical_emergency	0.2

Table S11 – Topic description: Ebola (correlation matrix)

	k100_37	k200_159	k300_62
k100_37	1		
k200_159	0.97	1	
k300_62	0.96	0.98	1

Table S12 – Topic description: MH370

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_51	*plane, flight, search, airlin, aircraft, air, pilot, passeng, malaysia, crash, ocean, miss, fli, found, mh370*	144	1074	Topic_091	0.07	E_air_travel	0.63
k200_142	*search, plane, malaysia, flight, ocean, mh370, miss, disappear, airlin, found, malaysian, investig, debri, indian, aircraft*	55	886	Topic_123	0.19	E_air_travel	0.5
k300_272	*search, plane, malaysia, mh370, ocean, flight, miss, malaysian, disappear, found, airlin, investig, indian, debri, area*	48	847	Topic_187	0.21	E_air_travel	0.49

Table S13 – Topic description: MH370 (correlation matrix)

	k100_51	k200_142	k300_272
k100_51	1		
k200_142	0.94	1	
k300_272	0.93	0.99	1

Table S14 – Topic description: Osama Bin Laden

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_61	*bin, laden, al, attack, terror, terrorist, saudi, qaeda, afghanistan, osama, pakistan, kill, islam, muslim, u.*	390	1378	Topic_045	0.11	L_anger	0.27
k200_34	*bin, laden, al, osama, qaeda, pakistan, his, kill, he, leader, attack, pakistani, offici, afghanistan, raid*	179	986	Topic_140	0.27	L_anger	0.19
k300_50	*bin, laden, al, osama, qaeda, afghanistan, pakistan, kill, taliban, attack, leader, his, he, pakistani, raid*	193	1048	Topic_227	0.17	E_war	0.2

Table S15 – Topic description: Osama Bin Laden (correlation matrix)

	k100_61	k200_34	k300_50
k100_61	1		
k200_34	0.89	1	
k300_50	0.91	0.97	1

Table S16 – Topic description: Michael Jackson

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_15	*jackson, his, michael, he, death, murray, not, him, angel, singer, who, los, pop, die, estat*	43	912	Topic_043	0.1	L_male	0.22
k200_45	*jackson, michael, his, he, murray, death, singer, him, pop, angel, propofol, who, los, not, estat*	26	862	Topic_096	0.09	L_male	0.2
k300_18	*jackson, murray, he, his, death, michael, singer, doctor, not, propofol, him, los, dr, angel, aeg*	4	429	Topic_231	0.12	E_occupation	0.2
k300_231	*jackson, michael, his, he, death, pop, estat, music, him, who, fan, king, singer, star, alleg*	23	502	Topic_018	0.12	L_male	0.18

Table S17 – Topic description: Michael Jackson (correlation matrix)

	k100_15	k200_45	k300_18	k300_231
k100_15	1			
k200_45	0.98	1		
k300_18	0.81	0.84	1	
k300_231	0.63	0.61	0.12	1

Table S18 – Topic description: Coronavirus

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_67	*vaccin, trial, develop, covid, test, clinic, phase, dose, coronavirus, research, against, work, effect, candid, treatment*	176	1938	Topic_072	0.14	E_health	0.27
k100_70	*coronavirus, covid, case, pandem, test, health, peopl, death, new, countri, virus, lockdown, confirm, week, report*	524	1864	Topic_099	0.2	L_number	0.16
k100_72	*virus, infect, diseas, test, flu, coronavirus, sar, spread, human, caus, immun, influenza, peopl, infecti, cov*	374	1047	Topic_035	0.19	E_health	0.29
k200_24	*coronavirus, covid, pandem, virus, health, peopl, outbreak, spread, lockdown, diseas, who, wuhan, public, social, infect*	260	441	Topic_028	0.34	E_health	0.15
k200_144	*vaccin, develop, trial, test, covid, dose, phase, candid, coronavirus, work, research, against, clinic, compani, immun*	126	1456	Topic_139	0.11	E_health	0.23
k300_200	*coronavirus, covid, pandem, virus, sar, cov, outbreak, spread, wuhan, peopl, infect, health, novel, diseas, respiratori*	113	230	Topic_119	0.22	E_health	0.18
k300_288	*coronavirus, covid, pandem, peopl, lockdown, virus, infect, reopen, health, week, countri, case, distanc, test, outbreak*	293	1225	Topic_110	0.3	L_relativ	0.15
k300_289	*vaccin, develop, trial, phase, test, dose, candid, covid, compani, clinic, coronavirus, work, human, moderna, research*	106	1254	Topic_036	0.17	E_health	0.2

Table S19 – Topic description: Coronavirus (correlation matrix)

	k100_67	k100_70	k100_72	k200_24	k200_144	k300_200	k300_288	k300_289
k100_67	1							
k100_70	0.03	1						
k100_72	0.14	0.14	1					
k200_24	0.14	0.73	0.31	1				
k200_144	0.92	0.02	0.12	0.08	1			
k300_200	0.22	0.4	0.55	0.71	0.15	1		
k300_288	0.02	0.79	0.05	0.51	0.01	0.07	1	
k300_289	0.88	0.02	0.09	0.08	0.96	0.12	0.01	1

Table S20 – Topic description: Zika Virus

K_Topic	**Top 15 words**	**N C**	**N M**	**Max Topic**	**r**	**Max LF**	**r**
k100_7	*zika, virus, mosquito, infect, case, brazil, diseas, health, travel, spread, pregnant, born, women, microcephali, area*	138	1363	Topic_072	0.04	E_programming	0.27
k200_14	*zika, virus, mosquito, infect, case, spread, travel, health, microcephali, pregnant, diseas, dengu, brazil, born, transmiss*	123	1277	Topic_199	0.06	E_programming	0.27
k300_169	*zika, virus, mosquito, infect, spread, microcephali, case, dengu, brazil, pregnant, travel, born, diseas, area, outbreak*	119	1200	Topic_078	0.12	E_programming	0.26

Table S21 – Topic description: Zika Virus (correlation matrix)

	k100_7	k200_14	k300_169
k100_7	1		
k200_14	0.98	1	
k300_169	0.96	0.98	1

Table S22 – Topic description: Saddam Hussein

K_Topic	**Top 15 words**	**N C**	**N M**	**Max Topic**	**r**	**Max LF**	**r**
k100_45	*iraq, iran, saddam, israel, iraqi, war, hussein, syria, state, isra, against, u., al, arab, regim*	695	1209	Topic_071	0.21	E_war	0.38
k200_89	*iraq, saddam, iraqi, hussein, al, his, baghdad, he, war, regim, forc, arab, invas, kuwait, against*	93	916	Topic_094	0.17	E_war	0.25
k300_134	*iraq, saddam, iraqi, hussein, al, his, baghdad, he, regim, war, arab, invas, forc, dictat, leader*	54	865	Topic_102	0.15	E_war	0.24

Table S23 – Topic description: Saddam Hussein (correlation matrix)

	k100_45	k200_89	k300_134
k100_45	1		
k200_89	0.76	1	
k300_134	0.74	0.98	1

Table S24 – Topic description: Elvis Presley

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_78	*gate, elvi, his, bill, presley, he, year, king, microsoft, world, who, melinda, fan, vega, memphi*	81	1144	Topic_066	0.18	L_home	0.25
k200_46	*elvi, presley, his, he, king, fan, vega, memphi, las, graceland, year, rock, record, roll, music*	6	1079	Topic_131	0.16	E_music	0.3
k300_69	*elvi, presley, his, he, king, fan, memphi, graceland, rock, roll, parker, year, priscilla, record, death*	5	954	Topic_197	0.19	E_music	0.28

Table S25 – Topic description: Elvis Presley (correlation matrix)

	k100_78	k200_46	k300_69
k100_78	1		
k200_46	0.82	1	
k300_69	0.83	0.97	1

Table S26 – Topic description: 5G technology

K_Topic	Top 15 words	N C	N M	Max Topic	r	Max LF	r
k100_16	*5g, technolog, network, system, data, devic, comput, mobil, phone, use, digit, huawei, new, connect, communic*	393	1449	Topic_036	0.11	E_computer	0.51
k200_41	*5g, network, mobil, technolog, huawei, wireless, phone, connect, servic, compani, communic, speed, equip, infrastructur, telecom*	144	901	Topic_061	0.15	E_technology	0.34
k300_243	*5g, network, mobil, technolog, huawei, wireless, connect, servic, phone, compani, communic, telecom, speed, internet, secur*	128	851	Topic_252	0.18	E_technology	0.33

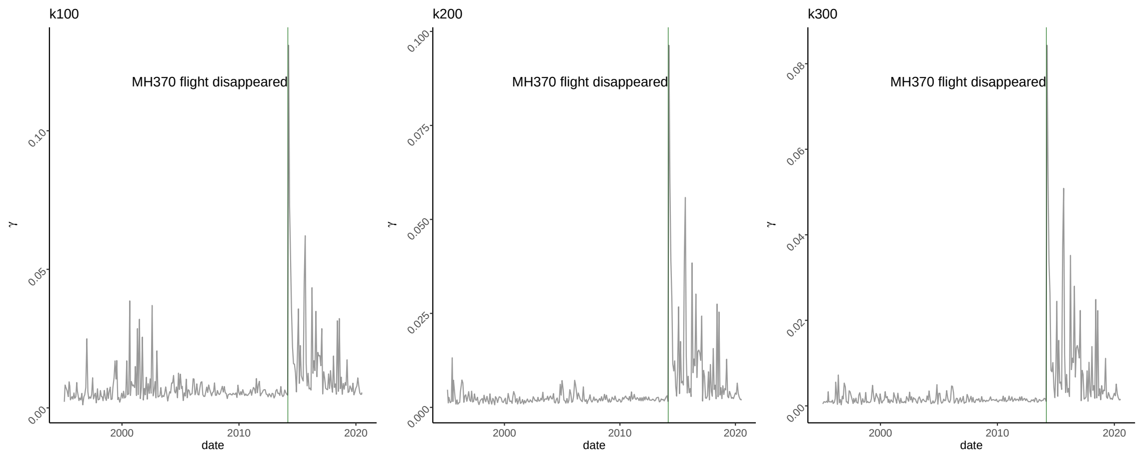
Table S27 – Topic description: 5G technology (correlation matrix)

	k100_16	k200_41	k300_243
k100_16	1		
k200_41	0.84	1	
k300_243	0.83	0.98	1

7.1 Topic specificity across different *k*s

Different ks offer different topic resolutions, where the higher the k (*i.e.*, the highest number of topics within the corpus) the higher the topic resolution (it follows that lower k offers a more general topic description). For example, within the topics around the MH370 flight disappearance, *i.e.*, topics k100_51, k200_142, and k300_272 (see Tables S12 and S13), it is possible to observe how topic specificity increases as a function of k. First, the keyword "mh370" moves from position 15 (in k100) to the 6$^{\text{th}}$ (k200) and 4$^{\text{th}}$ (k300) position indicating higher beta value for such term hence higher weight in the topic, which, from k100 to k300 describes more specifically the MH370 disappearance rather than flights in general. Second, the number of documents where this is the top topic (*i.e.*, higher gamma value across all topics within k dataset) decreases, indicating indeed higher specificity. Lastly, the correlations with the general lexical feature of *air_travel* decrease with the increase of k, suggesting that the topic is less general, focusing less on *air_travel* and more specifically on MH370. An increase in topic specificity as a function of k is visible in the Figure S4 below. For each of the three topics related to the disappearance of the flight MH370, we plotted the gamma values of all documents across a time span ranging from 1995 to 2020. We consider as more specific the topics that have less noise (*i.e.*, lower values) in gamma values in the period that precedes the event on the date 8$^{\text{th}}$ March 2014.

Figure S4 – The same topic over time on different ks



SM 8 Representativeness of conspiracy documents

In the Table S28 and Table S29 below, we show the top/bottom five conspiracy representative documents (first 1,000 characters) ordered by the cosine similarity value.

Table S28 – Top 5 highest representative conspiracy documents

doc_id	**cs**	**text**
C01b90	0.280	The reality of what's truly happening inside the United States is far, far more terrifying than just being screwed. Because quite literally, they are trying to KILL US! "Oh, come on now," I can hear you thinking. "Get real!" Well that's exactly what I'm will be doing in this article. I'm going to get very real. I hope that you'll read all the way to the bottom because you really need to know what's happening and why this isn't just a big government stupidity at work. What I share here is more than deadly serious and it is based upon a lot of research that many others have put together to clarify what's really happening in our country. Be sure to check out some of the videos I've included on this page to confirm what I'm telling you is true. This is hard for us to hear. It is difficult for rational and sane individuals to believe that anything on this scale could actually be happening around the world. That's why its so easy to dismiss this, without bothering to really
C0274e	0.280	Why are they trying to kill us? A deeper look into the mass production of America's toxic food My illustration on the left is really much too tame. I don't mean to imply that you'll just have to live with stupid government stunts, as if the screw they're turning into our collective chests might be only a painful inconvenience. The reality of what's truly happening inside the United States is far, far more terrifying than just being screwed. Because quite literally,they are trying to KILL US! "Oh, come on now," I can hear you thinking. "Get real!" Well that's exactly what I'm will be doing in this article. I'm going to get very real. I hope that you'll read all the way to the bottom because you really need to know what's happening and why this isn't just a big government stupidity at work. What I share here is more than deadly serious and it is based upon a lot of research that many others have put together to clarify what's really happening in our country. Be sure to check out so HT1
C05196	0.277	The great unraveling is gaining momentum by the day, and even now the majority of the masses are fast asleep at the wheel. Our skies look like something from an apocalyptic future, our trees are dying, our oceans are dying, our planet is dying, and still the majority remain unaware. Those in power are utilizing every option at their disposal to keep populations from waking until the last possible moment. Is there still time to change course and keep any part of the ship floating? That remains to be seen, but if there is yet a chance, it will require the concerted, focused, effective, prioritized, and completely dedicated efforts of all those that are already awake. The essay below was penned for geoengineeringwatch.org by a former USAF communications officer, Col. Randall Smith, Ph.D. Though I have no means of conclusively confirming the information presented by Randall, available date does generally confirm his estimations and conclusions. The essay does not address the implosion of t
C00113	0.276	Q: If Donald Trump, the Commander in Chief of the US Military, is Historically, Theologically and Morally Blind and Also Scientifically Illiterate, Should He be Making Life and Death Decisions for the Planet? "We have grasped the mystery of the atom and rejected the Sermon on the Mount. Ours is a world of nuclear giants and ethical infants. We know more about war than we do about peace – more about killing than we do about living." — WWII General Omar Bradley "If (Japan does) not now accept our terms they may expect a rain of ruin from the air, the likes of which has never been seen on this earth." – US President Harry S. Truman (August 6, 1945) "North Korea best not make any more threats to the United States... (If they do) they will be met with fire and fury... the likes of which this world has never seen before." – US President Donald J. Trump (August 8, 2017) On the eve of the anniversary of the United State's nuclear annihilation of the Christian community of Nagasaki on Aug
C036c7	0.276	Still Believe the New World Order Is Just A Conspiracy Theory? people say that they understand what the New World Order and a One World Government is all about, but that it doesn't necessarily have to be a bad thing. Maybe we need a New World Order, where we can live in one global state, in peace with no future wars? If there are no other countries to fight against, there will be no more wars either, right? And perhaps we will feel more united, as One Big Global Family if we erase the borders? Well, that doesn't sound bad, does it? If the intention behind the New World Order was the above, I would agree (although I don't think a centralized power can ever work. What if the Global CEO doesn't have our best interests in mind?). So, is the above utopia what the Globalists have in mind? Is this benevolent World Society what they so eagerly work towards? The best way to answer those questions is to quote from "The horse's mouth" so to speak. What do the Globalists have to say themse

Note. doc_id = LOCO's unique document ID; cs = cosine similarity.

Table S29 – Top 5 lowest representative conspiracy documents

doc_id	**cs**	**text**
C01782	0.002	More science reports have confirmed the engineered DNA of the new Corona virus. Earth and societal changes are accelerating at blinding speed on countless fronts, links in the chain of our current reality are already breaking down. In spite of unfolding biosphere collapse, the US stock market has hit yet another record high. How long can normalcy bias be maintained in first world populations? What happens when the gravity and immediacy of all that is unfolding can no longer be hidden from public consciousness? The latest installment of Global Alert News is below. As the horizon continues to darken, many are choosing to double down on denial and apathy. In contrast, ever increasing numbers are awakening and choosing a path of courage and action toward the greater good. We must never underestimate our collective power if we stand together. All are needed in the critical battle to wake populations to what is coming, we must make every day count. Share credible data from a credible sour
C01234	0.023	Rothschild inherited some very important patents when flight MH-370 disappeared When Malaysia Airlines Flight MH370 went missing, a US tech firm had 20 senior staff on the flight, many of whom were shareholders in the company and some were co-owners of some very important patents for military radar systems and microchips for autonomous driving cars. With the mysterious disappearance of the Boing 777, all of the remaining shares held by the missing travelers for Freescale Semiconductor Ltd were inherited by the only remaining shareholder, Lord Jacob Rothschild, who then became the sole owner of some very important patents. PayPal: Donate in USD PayPal: Donate in EUR PayPal: Donate in GBP
C01607	0.025	Here is the chemtrails patent: "A method is described for reducing atmospheric or world-wide warming resulting from the presence of heat-trapping gases in the environment, *i.e.*, from the greenhouse impact. This kind of gases are relatively transparent to sunshine, but absorb strongly the extended-wavelength infrared radiation released by the earth. The method incudes the phase of seeding the layer of heat-trapping gases in the environment with particles of materials characterized by wavelength-dependent emissivity. Such materials include Welsbach components and the oxides of metals which have higher emissivity (and as a result minimal reflectivities) in the visible and 8-twelve micron infrared wavelength areas."
C02dc1	0.030	As a supposed act of defiance, The Institute of Digital Archaeology, Harvard University, and UNESCO are erecting 43 foot tall 23 foot wide Arches of the Temple of Baal in New York's Times Square and London's Trafalgar Square. The date to unveil the arches falls directly on the celebration of the all important pagan holiday Beltane, and anniversary of the massacre of The Branch Davidians at WACO and The Oklahoma City Bombing, April 19. Related: Will A Gateway Be Opened When The Arch From The Temple Of Baal Is Reconstructed In Times Square? The Infowars Life Lung Cleanse Plus is back in stock at 50% off with double Patriot Points and free shipping!
C00b0a	0.031	WEBINAR: Dead Sea Scrolls show Sanhedrin may have executed Jesus as an ET Disclosure activist – Peter Kling & Alfred Lambremont Webre Dr. Frank Stranges lecture on the Dead Sea scrolls in which he reads from the many open descriptions of UFO, spaceships, other planets and Extraterrestrials, outer space battles, in the Dead Sea Scrolls, and the death threats that Dr. Frank Stranges received from the Israel-based Rabbinical guardians of the Scrolls for making these contents public! Is the Israel the faction behind the national security guardianship of the Dead Sea Scrolls in fear of being exposed by the ETs with the return of Jesus, who is an ET! IMHO – As Stranges illustrates from the guarded Dead Sea Scrolls. Here is the late, great Dr. Stranges bio:

Note. doc_id = LOCO's unique document ID; cs = cosine similarity.

SM 9 Popularity and spread metrics correlations

Besides single URLs shares for each webpage, we also obtained three other metrics for websites shares and compared them. This was done by: a) entering on SharedCount.com (SC) the website domain (*i.e.*, home page, *e.g.*, `www.infowars.com/`); b) computing the sum of all LOCO documents' URLs shares per website; c) computing an estimation of website total shares based on our observed shares in LOCO's documents (see main text for details, sections 3.8.4 and 3.8.6).

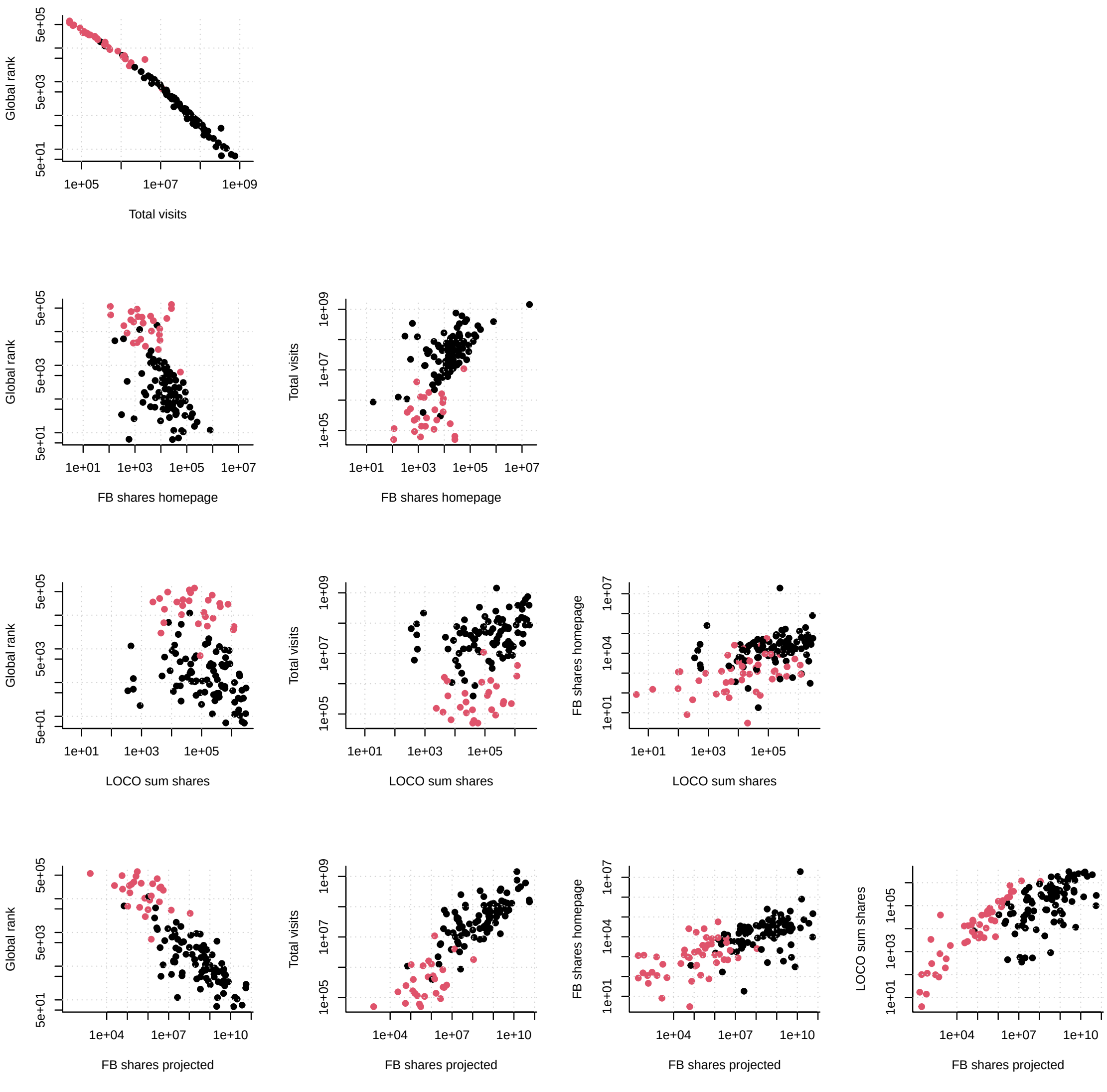
Below, in Table S30 (and in Figure S5) we show the correlation matrix (and scatter plots) for the main spread and popularity measures (all variables are log transformed). After having obtained the website shares from SC (believing it was a measure of overall website shares, see FB_shares_homepage in `website_metadata.json`), we soon realized that, although it correlates with both website rank (r = -.58) and monthly visits (r = .61), this measure does not reflect the total website shares. In fact, when we subtracted the computed sum of all URLs shares to this measure, we obtained a large number of negative values, meaning that the sum of the single URLs was shared more times than the overall website (*i.e.*, homepage). This led us to re-think this measure and in fact interpret it not as a measure of the whole website shares but as a measure of shares of the website's homepage. Although we do not see this measure as useful, we keep it in LOCO to allow users to replicate our analyses or use it for testing hypotheses we have not envisioned so far.

Table S30 – Correlation matrix of spread measures

Variable	1	2	3	4
1. Global rank				
2. Total visits	-.99			
	[-.99, -.99]			
3. FB shares homepage	-.58	.61		
	[-.69, -.45]	[.48, .71]		
4. LOCO sum shares	-.36	.37	.68	
	[-.51, -.20]	[.20, .51]	[.58, .76]	
5. FB shares projected	-.81	.81	.77	.85
	[-.86, -.73]	[.74, .86]	[.70, .83]	[.80, .89]

Note. Global rank: websites' global rank (from SimilarWeb); Total visits: websites' monthly visits; FB shares homepage: number of times the websites' homepage is shared on Facebook; LOCO sum shares: the aggregated number of individual LOCO's webpages shares on Facebook; FB shares projected: the computed number of overall Facebook shares given observed data in LOCO. Values in square brackets indicate the 95% confidence interval for each correlation. All correlations are significant at $p < .01$.

Figure S5 – Scatter plots of spread measures



Note. Red and black dots represent conspiracy and mainstream websites respectively. Global rank: websites' global rank (from SimilarWeb); Total visits: websites' monthly visits; FB shares homepage: number of times the websites' homepage is shared on Facebook; LOCO sum shares: the aggregated number of individual LOCO's webpages shares on Facebook; FB shares projected: the computed number of overall Facebook shares given observed data in LOCO.

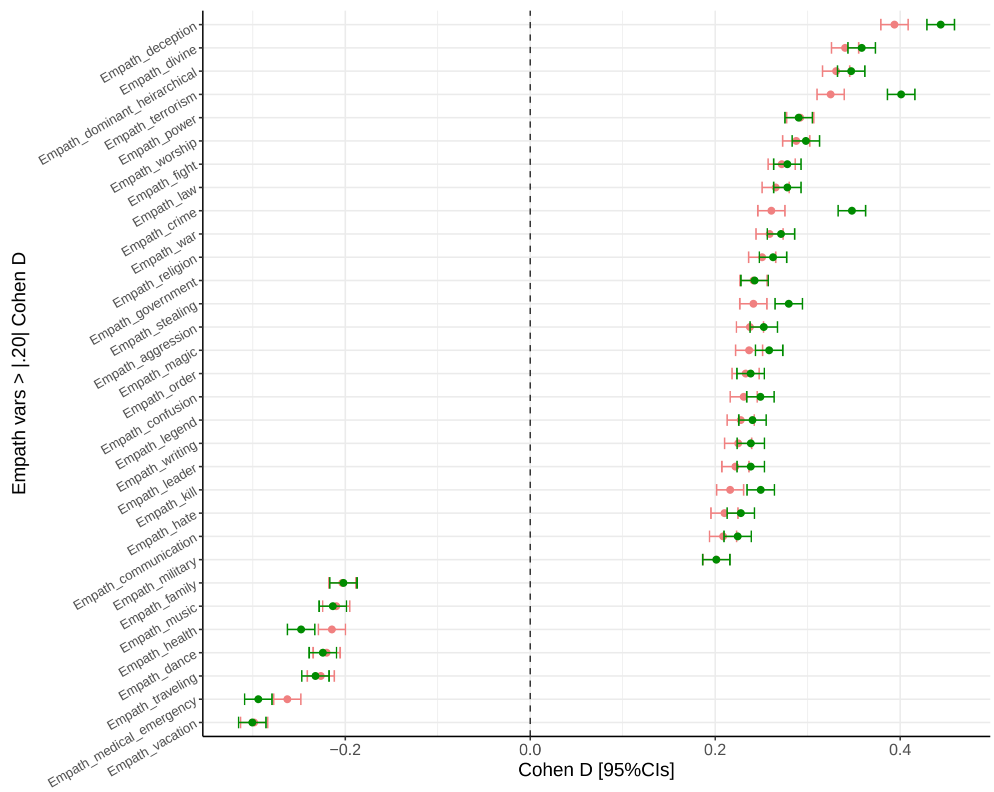
SM 10 Lexical features of mentioning conspiracy

In section 4.2.1 of the main text, we have shown how LOCO's lexical features overlap with those obtained in Klein et al. (2019) from Reddit. Here, with the same protocol and on the same set of Empath categories, we explored the effect of cleaning the mainstream subcorpus from mentions to conspiracy (see section 3.8.1 of the main text). As shown in Figure S6, overall, the absolute effect size is larger when mainstream documents mentioning conspiracy are removed (green error bars) compared to the mainstream subcorpus that includes all documents (red error bars).

We then tested whether changes in effect size from the corpus with all documents (raw data) to the corpus without documents mentioning conspiracy (cleaned data) were significant and so we ran a paired samples t-test on the absolute d values. The t-test showed a significant increase in effect size from raw to clean data, $t_{30} = 5.08$, $p < .001$, see Figure S7 below.

We then took the top four categories that yielded the larger changes in effect sizes and tested whether there was a correlation between the number mentioning conspiracy and those four lexical features. Results with log-transformed variables (plots in Figure S8) showed a positive relationship: crime: $r = .31$; terrorism: $r = .33$; deception: $r = .21$; and stealing: $r = .13$ (Non log-transformed variables: deception: $r = .14$; terrorism: $r = .25$; crime: $r = .20$; stealing: $r = .06$).

Figure S6 – Effect of mentioning conspiracy on lexical features



Note. Differences, in Cohen's d, between conspiracy and all mainstream documents (red error bars, raw data) and between conspiracy and mainstream documents without documents mentioning conspiracy (green error bars, cleaned data). Positive values indicate that the effect is higher in conspiracy.

Figure S7 – Paired t-test on effect size changes from raw to cleaned subcorpora

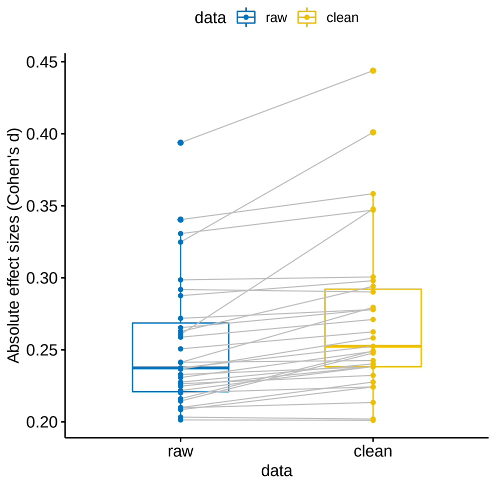
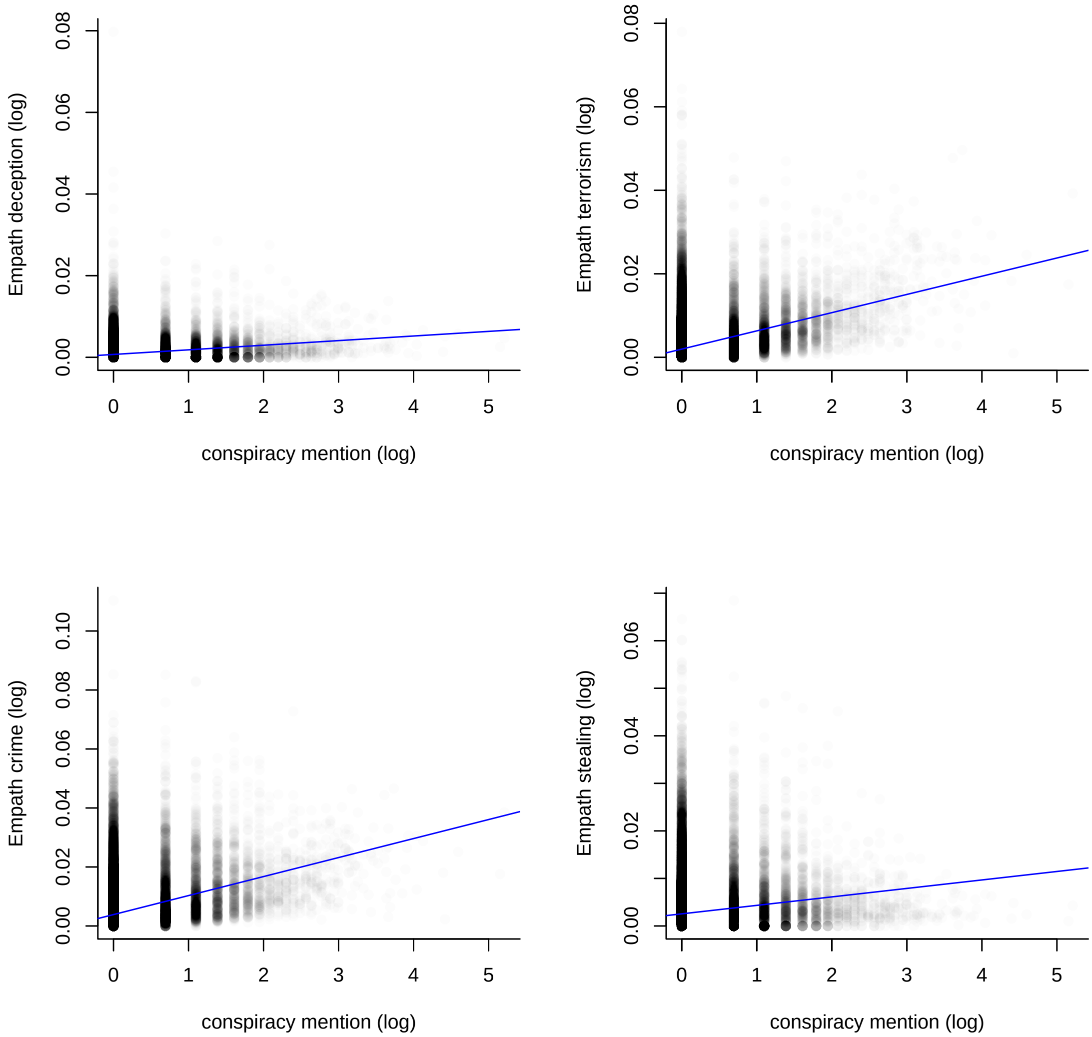


Figure S8 – Scatter plots of number of mentions of conspiracy and lexical features

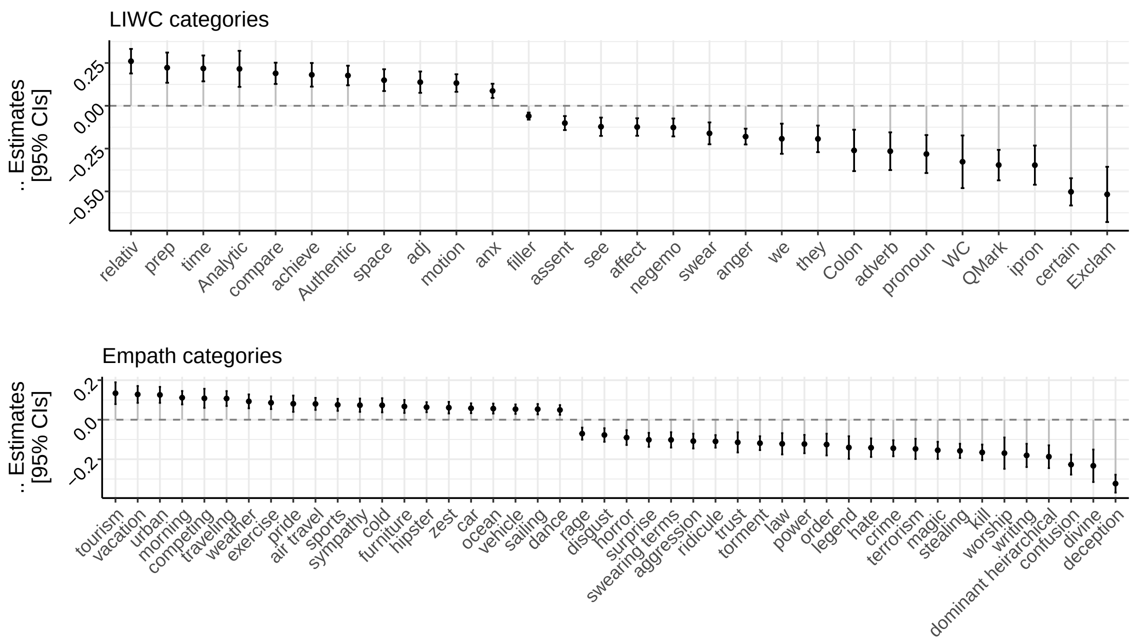


Note. All variables have been log-transformed prior to analyses

SM 11 Properties of representative conspiracy documents

In section 4.2.3 of the main text, we have plotted the Standardized beta estimates of multilevel models in which we specified as random intercept the topics obtained setting k at 100. Here, we replicate the same analyses changing instead the random intercept with topic obtained at $k = 200$. Results are not observably different (compare with Figure 4 in the main text). See Figure S9.

Figure S9 – Differences in lexical features between high and low representative conspiracy documents



References

Benoit, K., Muhr, D., and Watanabe, K. (2020). stopwords: Multilingual Stopword Lists.

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., and Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30):774.

Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., and Deravi, F. (2019). Understanding Conspiracy Theories. *Political Psychology*, 40(S1):3–35.

Fry, E. (2000). *1000 instant words: the most common words for teaching reading, writing and spelling*. Teacher Created Resources.

Klein, C., Clutton, P., and Dunn, A. G. (2019). Pathways to conspiracy: The social and linguistic precursors of involvement in Reddit's conspiracy theory forum. *PLOS ONE*, 14(11):e0225098.

Nguyen, D., Liakata, M., DeDeo, S., Eisenstein, J., Mimno, D., Tromble, R., and Winters, J. (2020). How We Do Things With Words: Analyzing Text as Social and Cultural Data. *Frontiers in Artificial Intelligence*, 3.

Oswald, S. (2016). Conspiracy and bias: argumentative features and persuasiveness of conspiracy theories. *OSSA Conference Archive*, 168:1–16.

Rinker, T. W. (2013). qdapDictionaries: Dictionaries to Accompany the qdap Package.

5.2 Study 2

NEUROSCIENCE

Interconnectedness and (in)coherence as a signature of conspiracy worldviews

Alessandro Miani[1]*, Thomas Hills[2,3], Adrian Bangerter[1]

Conspiracy theories may arise out of an overarching conspiracy worldview that identifies common elements of subterfuge across unrelated or even contradictory explanations, leading to networks of self-reinforcing beliefs. We test this conjecture by analyzing a large natural language database of conspiracy and nonconspiracy texts for the same events, thus linking theory-driven psychological research with data-driven computational approaches. We find that, relative to nonconspiracy texts, conspiracy texts are more interconnected, more topically heterogeneous, and more similar to one another, revealing lower cohesion within texts but higher cohesion between texts and providing strong empirical support for an overarching conspiracy worldview. Our results provide inroads for classification algorithms and further exploration into individual differences in belief structures.

INTRODUCTION

Conspiracy theories (CTs) propose alternative explanations of publicly relevant events [e.g., vaccination, coronavirus disease 2019 (COVID-19), climate change, and Princess Diana's death], evoking secret plots by malevolent and powerful groups who act at the expense of an unwitting population (*1*). CTs are popular. In a nationally representative sample of American adults, half of respondents believed in at least one medical CT, such as that the Food and Drug Administration deliberately prevents access to effective natural cures for cancer and that fluoride is a dangerous by-product that the government allows phosphate mines to dump into the public water supply (*2*). In a cross-national survey in 17 nations, more than half of respondents believed in CTs associated with the 9/11 terrorist attack (*3*). Furthermore, CTs may emerge and propagate during times of crisis (*4*) such as epidemics, deaths of public figures, or even natural disasters, which, in turn, may increase exposure to other CTs (*5*).

Belief in CTs is associated with numerous negative outcomes. Belief in medical CTs correlates with reduced likelihood of influenza vaccination or annual checkups (*2*). Belief in HIV/AIDS CTs reduces intentions to use condoms (*6*). Exposure to CTs reduces intentions to limit carbon footprints (*7*). Belief in CTs is also associated with political extremism and violence (*8*, *9*) and increased intentions to engage in everyday crime (*10*). At the societal level, these phenomena may lead to loss of human lives, waste of public funds, and disruption of social order (*11*–*13*). The ubiquity and negative impact of CTs warrant increasing efforts to understand them.

One pathway to understanding CTs is the observation that people who believe in one CT tend to believe in others, irrespective of how unrelated or contradictory they may seem (*14*–*16*). For example, believing that the AIDS virus was deliberately engineered in a government laboratory is associated with believing that the Federal Bureau of Investigation was involved in Martin Luther King's assassination (*16*). Conspiracy believers tend to identify meaningful relationships among randomly co-occurring events (*17*, *18*), confuse aspects of reality such as believing that prayers have the capacity to heal (*19*), and believe in the paranormal (*20*).

This collection of events, however, can devolve into contradiction. Individuals who felt that it was more likely that Osama Bin Laden was already dead before the U.S. military forces arrived at his compound in Pakistan were also more likely to believe that he is still alive (*15*). As the authors of that work put it, "...the specifics of a conspiracy theory do not matter as much as the fact that it is a conspiracy theory at all" (p. 5). This was further reiterated by Lewandowsky and colleagues (*21*), "The incoherence does not matter to the person rejecting the official account because it is resolved at a higher level of abstraction" (p. 179).

Thus, belief in multiple conspiracies may be self-supporting: The interconnectivity among conspiracy beliefs is supported by a metabelief that resolves the apparent contradictions at the lower level. Hence, CTs may thus constitute a mutually reinforcing network of beliefs, creating self-sustaining evidence for a world dominated by deceptive agents (*15*). These networks would constitute a top-down conspiracy worldview, which coerces unconnected or even contradictory observations into support for a global conspiracy (*21*).

If the criterion for being correct is not about the veridicality or mutual compatibility of individual events but rather their support for the existence of deceptive agents, then a conspiracy worldview can even sustain entirely fictitious beliefs. For example, in one study (*22*), the more participants believed in popular CTs (e.g., about 9/11), the more they perceived as real a set of fictitious CTs created ad hoc for the study. This study suggests that the belief in some CTs increases the chances that an individual will accept evidence for novel CTs.

The combination of incompatible (or, more generally, incoherent) lower-level explanations that get resolved by mutual compatibility with a higher-level belief that authorities are deceptive has recently been claimed to be the hallmark of conspiracy worldviews (*21*). These patterns of beliefs might facilitate the creation and endorsement of the unusual patterns seen in conspiracy narratives (*23*). CTs serve sense-making functions, as they allow one to reduce the complexity of real-world events (*24*, *25*), dismissing them as evidence for a grand conspiracy. Thus, the narrative structure of CTs may be important in supporting a CT worldview. The conjecture that conspiracy worldviews have local incoherence (supporting evidence is drawn from numerous unrelated and sometimes contradictory sources) but global coherence (appealing to an overarching belief in the deceptive

[1]Institute of Work and Organizational Psychology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland. [2]Department of Psychology, University of Warwick, University Road, Coventry CV47AL, UK. [3]The Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, UK.
*Corresponding author. Email: alessandro.miani@unine.ch

nature of authorities) constitutes an important theoretical step forward in the psychology of conspiracy beliefs. However, it remains virtually untested outside of a handful of studies focusing on specific CTs (*15*, *21*, *26*).

In the current study, we exploit the abundant instantiations of CT beliefs in naturally occurring narrative text. We perform a large-scale text analysis to test the conjectures that conspiracy narratives are characterized by a multitude of interconnected ideas (*27*) that manifest in patterns of lower local (within-text) cohesion but higher global (between-text) cohesion (*15*, *21*). If a conspiracy worldview coerces unconnected observations to support for an overarching belief in the deceptive nature of authorities, then a network of potential conspiracy-related topics will be more tightly interconnected in conspiracy documents than in nonconspiracy documents [hypothesis 1 (H1)]. Similarly, if conspiracy narratives focus less on individual topics than nonconspiracy narratives, then topic specificity will be lower for conspiracy than nonconspiracy documents (H2a). Furthermore, text cohesion between paragraphs should be less internally cohesive for conspiracy documents than nonconspiracy documents (H2b). Last, if conspiracy narratives reference similar worldviews, then similarity between documents should be higher for conspiracy documents (as a group) than nonconspiracy documents (H3).

Our analyses rely on the largest corpus of CTs available today, language of conspiracy (LOCO) (*28*), an 88 million–word corpus composed of topic-matched conspiracy (N = 23,937) and nonconspiracy (N = 72,806) text documents (i.e., webpages) harvested from 150 websites. LOCO provides two types of semantic indexes for each document. One is represented by seeds (N = 39), keywords used via a Google search to retrieve documents associated with events that have generated CTs (e.g., 9/11 terroristic attacks, the death of Princess Diana, and COVID-19). The other one is represented by topics, extracted from documents with latent Dirichlet allocation (LDA) (*29*). Topics are expressed as distribution of probabilities, while seeds are either present or not in the document. In terms of content, topics differ from seeds because they represent out-of-domain themes extracted from the corpus a posteriori (versus seeds that were defined, by us, a priori). LDA topics differ from seeds in terms of granularity: We provide three sets of topics that represent the semantic resolution (at 100, 200, and 300 topics) of our corpus. These differences between seeds and topics allow us to evaluate within- and cross-theme similarities as they differ between conspiracy and nonconspiracy documents.

RESULTS

Interconnectedness (H1)

To test H1—conspiracy-related topics will be more tightly interconnected in conspiracy documents than in nonconspiracy documents—we ran network analyses on the co-occurrences of seeds and topics extracted from the conspiracy and nonconspiracy subcorpora. We calculated the average degree of connectedness in the networks by extracting how many edges were connected to each node (either seeds or topics).

Seed interconnectedness

We first tested whether conspiracy documents, on average, contained more different seeds than nonconspiracy documents. That is, we calculated thematic richness by counting how many seeds were present in each document. We fitted a linear mixed-effects regression model predicting the number of seeds contained in documents by subcorpus (either conspiracy or nonconspiracy), adding word count as covariate and nesting documents within websites. On average, conspiracy documents contained more seeds than nonconspiracy documents: $\beta = 0.85$, SE = 0.129, $t_{119.44} = 6.59$, $P < 0.001$, $R^2_{\mathrm{m/c}}$ (marginal/conditional) = 0.093/0.454 (conspiracy: M = 1.293, SD = 0.741, range: 1 to 13; nonconspiracy: M = 1.073, SD = 0.305, range: 1 to 6).

We then tested our main hypothesis (H1), which is how the degree of interconnectedness differed between conspiracy and nonconspiracy networks. We fitted a linear mixed-effects regression model predicting interconnectedness (i.e., number of edges per node) by subcorpus (either conspiracy or nonconspiracy) while nesting nodes within nodes' names (similar to a paired test that tracks differences within a seed, e.g., "lady_diana" between conspiracy and nonconspiracy networks). The conspiracy network was more interconnected compared to the nonconspiracy network: $\beta = 0.97$, SE = 0.121, $t_{38} = 8.03$, $P < 0.001$, $R^2_{\mathrm{m/c}} = 0.235/0.720$. Figure 1 shows the two networks.

In table S1, we report an additional set of analyses that further confirm the above network differences. CT narratives formed a larger giant component than the nonconspiracy network, with fewer distinct subnetworks. Entropy was higher in the conspiracy (versus nonconspiracy) network, suggesting that seeds are, to a larger extent, agglomerated in a random, nonsystematic way. Conspiracy nodes (compared to nonconspiracy nodes) were more similar to each other in terms of connection patterns. The clustering coefficient (the probability that the adjacent vertices of a node are interconnected) was higher in the conspiracy network. The average shortest path length through nodes, i.e., distance, was lower in the conspiracy network, suggesting again more interconnectedness in conspiracy network. Last, density, the ratio of the number of edges to the number of possible edges, was also higher in the conspiracy, compared to the nonconspiracy, network.

Topic interconnectedness

Because seeds represent specific mentions of themes that have generated conspiracies (e.g., 9/11 and Princess Diana's death), we further investigated H1 using a more general pattern of word co-occurrences associated with LDA topics that were extracted in an unsupervised fashion from the corpus. We created topic networks, where nodes are topics and edges are the degree of correlation between topics. Relying on LOCO's three sets of LDA topics (that contain 100, 200, and 300 topics; henceforth, LDA100, LDA200, and LDA300), we then compared the average degree of interconnectedness between conspiracy and nonconspiracy networks. In all three sets, the conspiracy networks were more interconnected than the nonconspiracy networks, LDA100: $\beta = 0.30$, SE = 0.089, $t_{99} = 3.37$, $P = 0.001$, $R^2_{\mathrm{m/c}} = 0.022/0.607$; LDA200: $\beta = 0.301$, SE = 0.068, $t_{199} = 4.42$, $P < 0.001$, $R^2_{\mathrm{m/c}} = 0.023/0.538$; and LDA300: $\beta = 0.33$, SE = 0.06, $t_{299} = 5.55$, $P < 0.001$, $R^2_{\mathrm{m/c}} = 0.027/0.470$. Similar to the results obtained for seed networks, entropy was higher in the conspiracy networks. Conspiracy nodes were more similar to each other in terms of connection patterns. Clustering coefficients were higher in the conspiracy network. Distance was lower in conspiracy networks. In addition, density was higher in conspiracy networks. In table S2, we report the properties of each of the six networks (conspiracy and nonconspiracy networks for the three LDA topic matrices).

Local cohesion (H2)

To test H2a—topic specificity should be lower for conspiracy than nonconspiracy documents—we evaluated the inequality of within-document topic distributions using the Gini coefficient. Because each topic has a probability associated with each document, a more

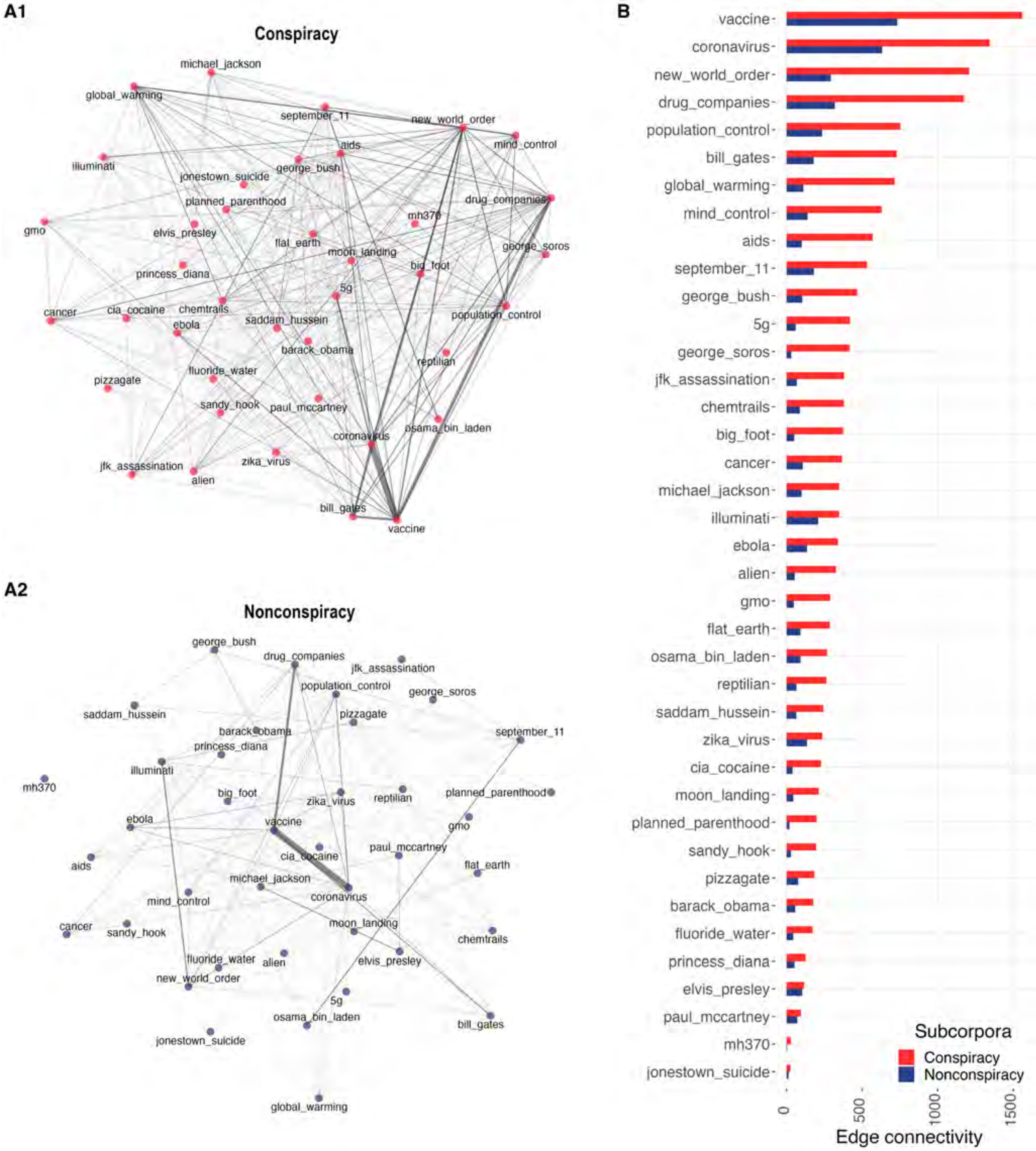


Fig. 1. Network interconnectedness from seeds. Nodes represent seeds (documents' keywords associated with events that have generated CTs), and edges represent the co-occurrence of seeds in documents. Thicker edges indicate higher co-occurrence. Left: Network plots extracted from the conspiracy (**A1**) (red) and nonconspiracy (**A2**) (blue) subcorpora. Right: Numbers of links held by each node (edge connectivity) as a measure of seed interconnectedness (**B**).

unequal topic distribution (i.e., higher Gini coefficient) indicates that the document is focused on fewer topics than a document with a more equal topic distribution (fig. S1). Results of the linear mixed-effects models for predicting topic specificity (controlling for word count and nested within websites) show that conspiracy documents had lower topic specificity compared to nonconspiracy documents: $\beta = -0.29$, SE = 0.063, $t_{145.09} = -4.66$, $P < 0.001$, $R^2_{m/c} = 0.019/0.143$. We then tested H2b—within-document lexical cohesion should be lower for conspiracy documents than nonconspiracy documents—by analyzing how well within-document paragraphs are semantically connected to each other (e.g., lexical cohesion across paragraphs). The measure we use correlates with perceived text coherence (*30*). Conspiracy documents showed lower lexical cohesion than nonconspiracy documents: $\beta = -0.68$, SE = 0.08, $t_{146.54} = -7.98$, $P < 0.001$, $R^2_{\mathrm{m/c}} = 0.080/0.302$. Note that lexical cohesion and topic specificity are two different constructs (they correlate poorly; $r_{96,741} = 0.19$, $P < 0.001$; see table S3): While topic specificity measures how many topics account for a document's content, lexical cohesion measures the lexical overlap of paragraphs.

Global cohesion (H3)

To test H3—similarity between documents should be higher for conspiracy documents than nonconspiracy documents—we reasoned that, at the global level, shared lexical patterns would make conspiracy documents more similar to each other than nonconspiracy documents. We computed a measure of between-document lexical overlap, the cosine similarity (CS) scores between documents within subcorpora (either conspiracy or nonconspiracy), obtaining a value of similarity for each document with all other documents within the subcorpus. Pairwise CS was computed between each document and the remaining documents within the same subcorpus. For each computation, we excluded documents with the same seed and those gathered from the same website to avoid same-topic and same-author lexicons inflating document similarity. Compared to nonconspiracy documents, conspiracy documents were more similar to each other: $\beta = 0.96$, SE = 0.064, $t_{148.14} = 14.87$, $P < 0.001$, $R^2_{\mathrm{m/c}} = 0.422/0.559$. Our results (see table S4 and fig. S2) are robust and replicate across six different subsets of LOCO (in which we artificially created subcorpora perfectly matched for topics and word count; tables S5 to S7 and fig. S3).

DISCUSSION

Popular belief in conspiracies is a riddle. Belief in one CT leads to believe in other CTs (*14–16*). Irrespective of how related events are, an accumulation of CTs reflect a view of a world dominated by deception, resulting in a self-sustaining network of supportive beliefs (*15*). This conjecture was previously tested by measuring the extent to which participants simultaneously endorsed researcher-designed potentially contradictory conspiratorial items. Here, moving beyond individual beliefs and beyond single-case studies of specific CTs, we performed a large-scale text and network analyses on the abundant naturally occurring instantiations of CTs, providing strong empirical support for an overarching conspiracy worldview in conspiracy narratives.

Our results show that conspiracy texts exhibit a pattern of strong interconnectedness with each other, linking multiple ideas that result in a dense and highly interconnected network (H1). Individual conspiracy documents are built from multiple sources and are, on average, less locally (within-document) coherent than corresponding nonconspiracy documents (H2). They nevertheless exhibit higher global (between-document) cohesion, being more lexically similar to each other than nonconspiratorial documents (H3).

Compared to nonconspiracy, conspiracy narratives are more interconnected via a dense and unstructured network of shared themes. These properties emerge not only from themes associated with events that have generated CTs (seeds) but also from out-of-domain themes (LDA topics). Such a high topical interconnectedness mirrors the psychological need to reduce uncertainty and gain control by explaining and finding order in real-world events that might otherwise seem random. Individuals who believe in CTs have an overarching conspiracy mentality that makes them more likely to draw implausible causal connections between random or unrelated events (*17*, *18*).

Qualitative research has also suggested that conspiracy narratives rely on an accumulation of ideas in support of their claims (*27*). Popular figures, organizations, technologies, and even states (e.g., China or the United States) are often cited and interact, giving rise to the unusual narrative patterns emerging in CTs (*23*, *31*). For example, according to one CT, the COVID-19 pandemic is a pretext for distributing harmful vaccines, activated by 5G radiation, which lead to a mass depopulation, all being commanded by George Soros and Bill Gates (*31*). In our study, the co-occurrence of these themes (operationalized by seeds: e.g., covid, 5g, soros, gates, and vaccines) represents the narrative richness of CTs. We have quantified this richness, showing that, on average, conspiracy narratives have higher number of seeds and lower topic specificity (i.e., more topics) in comparison to nonconspiracy narratives. These results confirm previous observations (*27*, *31*). Although displaying high thematic richness (less local cohesion), conspiracy documents are more lexically similar to each other (more global cohesion) compared to nonconspiracy documents. One could think that this conspiratorial (in)coherence is counterintuitive because more topical richness should maximize within-document lexical diversity, hence potentially decreasing between-document lexical similarity. In our sample, the reverse is true, and we stress that this phenomenon is evidence for top-down thematic coercion, where themes fit an overarching conspiracy worldview. In this way, each theme is translated into a conspiracy by adding a set of recurrent lexical patterns involving language of deception, questioning, social identification, and negative emotions (*28*). These lexical patterns can be reused in any conspiratorial context and are shared across conspiracy narratives (*28*).

Despite being less internally cohesive, conspiracy narratives may appear coherent to believers because the believers' mental representations are based on world knowledge that is not explicitly represented in the text. Proneness to hallucinations and delusional ideations, traits that are shared with belief in CTs (*32*), might help connecting the dots (*17*, *18*), so as to reduce uncertainty and gain control. People with delusional ideations (*33*), similar to conspiracy believers (*34*), tend to draw conclusions quicker and tend to be based on less evidence than people without delusions. Quickly connecting poorly related ideas gives rise to a highly chaotic and randomly connected network of ideas. This might explain why the conspiracy networks we built from narratives are denser and more unstructured in comparison to nonconspiracy networks.

Our findings resemble phenomena in thought of individuals on the schizophrenia spectrum, which includes the subclinical and milder form schizotypy. Schizotypy is a trait that correlates with belief in CTs (*32*). Individual with schizotypal personality tend to jump to conclusions quicker and make decisions based on less evidence compared to people without schizotypy (*35*). Schizotypy and schizophrenia overlap substantially and are characterized, yet on different levels, by an impairment in thought and perception that lead to psychotic symptoms (*36*). This impairment is manifested in language production (*37*, *38*). Patients with schizophrenia show disruptions in speech production at the level of causal-motivational and thematic coherence (*39*) and in structural cohesive markers (*40*). Moreover, patients with schizophrenia also show disorganized semantic networks (*41*, *42*). These studies indicate a certain degree of overlap between the schizophrenia spectrum and belief in CTs in regard to semantic processing, suggesting that further research should pay attention to this overlap.

Our findings help advance research on fighting misinformation. Both schizotypal personality and belief in CTs are linked to vulnerability to misinformation (*43*, *44*). The fact that conspiracy narratives are characterized by high global cohesion despite low local cohesion may help develop classification algorithms to detect conspiratorial language either online or offline. This could be achieved by extracting the lexical patterns that are shared across individual conspiracies such as language of deception, questioning, and social identification (e.g., "Are they lying to us?"). Moreover, future computational endeavors in natural language processing could move further, helping detect contradictory statements in texts, hence replicating seminal findings (*15*). This move might benefit not only misinformation and conspiracy research but also cognitive science and clinical research in general.

Our study has some limitations. One limitation is that, despite the number of controls we introduced in our analyses, our results could be, in part, driven by other factors. The variance explained by some of our models (especially topic interconnectedness and topic specificity) was modest, suggesting that other factors might be in play to explain differences in local and global (in)coherence between conspiracy and nonconspiracy narratives. Further research might explore other indicators of cohesion to test the robustness of these findings.

This study leaves some questions open for future research. To what extent is individuals' conspiratorial (in)coherence specifically related to their belief in CTs? Are there any other individual differences that affect such tolerance to incoherence? For example, because worldviews are essential for making sense of the world, disconfirming a worldview would affect the very sense of an individual's reality (*45*). To avoid this, people enable defensive mechanisms such as confirmation bias (*28*, *46*) that allows them to preserve a worldview by seeking confirmation while avoiding challenges. Similar mechanisms could be used to protect any type of worldview. For example, irrespective of how related they are, ideas could be simultaneously endorsed to fit a coherent political or religious worldview. Furthermore, individual characteristics such as education or holistic thinking style might increase the tolerance to accept incoherent ideas. Conversely, analytic thinking style, negatively associated with belief in CTs (*47*), decreases the endorsement of contradictory statements (*48*).

To summarize, our findings contribute to a better understanding of the textual structure of CTs, linking theory-driven psychological research on CT beliefs measured in individuals (*1*) with data-driven, computational approaches to CT narratives measured in texts (*23*, *49*). This move links CT research to a larger body of research on computational approaches to fake news and misinformation (*50*, *51*) and offers inroads to develop classification algorithms and design debunking campaigns and institutional communication to counteract the spread of CTs online.

MATERIALS AND METHODS

Material

Our text material is gathered from the LOCO corpus (*28*). LOCO is a freely available, multilevel, topic-specific, ~88 million–token corpus of documents extracted from ~100,000 webpages. LOCO is composed of both conspiracy (N = 23,937) and nonconspiracy (N = 72,806) documents nested within websites (N = 150).

LDA topic extraction

LOCO's topics were extracted with LDA (*29*). LDA is an unsupervised probabilistic machine learning model capable of identifying co-occurring word patterns and extracting the underlying topic distribution for each text document. By setting a priori the number of topics desired from a given corpus, LDA computes, for each document in a corpus, the probabilities for all topics to be represented in the document. Each word of the corpus has a probability to be part of a topic. That is, a word x has probability β of being part of topic k; a topic k has probability γ of being part of document n. The sum of all the word probabilities within one topic is 1, and the sum of all the topic probabilities within one document is 1.

Before topic extraction, texts were preprocessed: Texts were converted to American Standard Code for Information Interchange (ASCII) characters; lower cased; cleaned by Uniform Resource Locators (URLs), punctuation, numbers, symbols, separators, and stop words [for the full list, see SM5 in the supplementary materials of (*28*)]; and stemmed. We then generated a document term matrix, from which we extracted only the most frequent 10,000 words. Topic extraction was performed with the topicmodels R package (*52*), using Gibbs sampling. The other LDA parameters were set as default. LOCO is provided with three LDA topic matrices, which contain 100, 200, and 300 topics (LDA100, LDA200, and LDA300, respectively).

Seed extraction

In LOCO, seeds are keywords used to retrieve webpages via Google during the corpus construction. A document can be associated with more than one seed. This is because a single webpage can be returned by a Google search using different keywords. For example, if a document relates to Lady Diana's death because of an Illuminati plot, then this document would be returned twice for both "lady_diana_death" and "illuminati" seeds.

Seeds differ from LDA-identified topics. Seeds are a set of keywords (related to straightforwardly identifiable themes such as the Sandy Hook school shooting or AIDS) built a priori to retrieve documents; LDA topics are extracted a posteriori from the given set of documents in an unsupervised fashion. Although they sometimes overlap, they constitute two methodologically different approaches. A webpage is returned by Google if the seed is present in the webpage (but note, not necessarily in the main text) at least once. However, the seed presence in the webpage does not necessarily indicate that the seed reflects the main topic of the document's text because the seed can be contained in boilerplate texts or in the comment section of the webpage.

We tested to what extent seeds reflect text content. We started by searching for the words that compose seeds in each document (e.g., "climate" and "change" for documents associated with the seed "climate_change"). We then we tested the agreement between seeds and text content. For all seeds, the mean of accuracy and precision were 0.909 (SD = 0.126, range: 0.300 to 0.997) and 0.993 (SD = 0.005, range: 0.982 to 0.999), with a sensitivity of 0.911 (SD = 0.132, range: 0.268 to 1.00) and a specificity of 0.788 (SD = 0.141, range: 0.304 to 0.938), respectively. These results show that there is a substantial overlap between seeds and the content of texts; hence, seeds are useful for indexing the semantic content of documents.

During LOCO's construction, some seeds were entered with synonyms to accommodate different spellings [e.g., "new_world_order" and

"NWO" as well as "climate_change" and "global_warming"; see table 4 in (*28*)]. Before our analyses, here, we aggregated the synonym seeds, reducing the seed pool from 47 to 39. A list of the 39 seeds is visible in Fig. 1 and fig. S2.

Networks from LDA topics and seeds

Interconnectedness—how documents are connected to each other via seeds or LDA topics—was tested on the networks resulting from the co-occurrences of seeds and LDA topics. We tested interconnectedness as edge connectivity, which is the number of nodes interconnected to each other.

Extracting networks from topics

To extract co-occurrences of LDA topics, we created network objects from the three LDA gamma values matrices provided with LOCO [LDA100, LDA200, and LDA300, whose dimensions are N documents (rows) by k topic (columns)], which contain 100, 200, and 300 topics, respectively. This was done by creating correlation matrices from the LDA topics matrices and then converting those matrices into graph objects, i.e., networks, using the igraph R package (*53*). In these networks, nodes are represented by topics, while edges are represented by their co-occurrences. We assessed interconnectedness by computing the edge connectivity, which is the number of edges associated with each node.

We started by computing, for each LDA matrix, the between-topic correlation matrix, extracting the Pearson r coefficient from the correlation between topic i and topic j within each matrix. To convert these correlation matrices into co-occurrence matrices, we needed a threshold of correlation values above which we consider a co-occurrence of topics. This is because if no threshold is provided, then all topics co-occur with each other (i.e., all topics whose $|r| > 0$; highest degree of connectivity that is equal to k, the number of topics). Conversely, if the threshold is too high, then no topic will co-occur (i.e., the degree of connectivity is equal to zero).

We explored how different $|r|$ thresholds would return a degree of connectivity different from zero and different from k (i.e., all topics co-occurring). To this purpose, we created a vector of $|r|$ values for each of the three correlation matrices (ranging from 0 to max r). For each value in the $|r|$ vector, we created a network object and extracted the degree of connectivity. As a threshold, for each of the three sets of networks, we selected the mean of all absolute correlation values from both conspiracy and nonconspiracy networks (see fig. S4). For LDA100, the mean of the absolute correlation values, above which we considered a topic co-occurrence, was $r = 0.034$ (range: 0 to 0.313). For LDA200, the mean was $r = 0.023$ (range: 0 to 0.400). For LDA300, the mean was $r = 0.018$ (range: 0 to 0.430).

Extracting networks from seeds

To create the networks of seeds for each subcorpus (both conspiracy and nonconspiracy), we created two co-occurrence matrices using the *fcm* function from the quanteda R package (*54*). We then created the graph networks using the R package igraph (*53*). The nodes of this network represent seeds, and the edges represent the co-occurrences of seeds within each matrix.

Topic specificity

Topic specificity measures the extent to which documents contain more or less topics. Here, we computed topic specificity by extracting documents' topics using LDA and by computing the inequality of topic distribution within each document using the Gini coefficient. Topic specificity can be thought in terms of inequality: The more unequal a distribution of topic is, the more a document is well represented by the highest value. As a measure of inequality, we used the unbiased Gini coefficient, which ranges from 0 to 1, where lower values indicate equal distribution. Thus, documents with higher Gini coefficients are better represented by a single LDA topic, whereas documents with lower Gini coefficients are more equally represented by a large number of topics. To extract the Gini coefficient, we used the function *Gini* from the R package DescTools that relies on the following equation

$$Gini = \frac{2\sum_{i=1}^{n} i\, y_i}{n\sum_{i=1}^{n} y_i} - \frac{n+1}{n}$$

The Gini coefficient was computed (for each document) on the top 10 topics with the highest gamma value. The choice of top 10 topics was justified for two main reasons. First, most of a document's content can be summarized within a handful of topics. Second, we visually explored the distributions of the documents where the Gini coefficient was either the highest or the lowest value, and we assessed that, by visual inspection, most of the variation in topic distributions occurs within the top 10 highest gamma values. To put differently, the top 10 highest gamma topics account for most of the document's content. This is visible in fig. S1, where we show the gamma values (black lines) and their cumulative sum (red lines) for the top 500 documents with the highest Gini (left) and top 500 documents with the lowest Gini (right) coefficient. The figure shows that in high-Gini documents, the cumulative topic probability (i.e., gamma, on the y axis) reaches around 0.70 to 0.80 proportion of all topic distributions with less than five documents. Differently, in low-Gini documents, the cumulative topic probability reaches around 0.70 to 0.80 proportion of all topic distributions with about 25 documents. This shows that in documents with high Gini coefficient, fewer topics are needed to account for a large part of the document's semantic content, suggesting therefore higher topic specificity.

Before testing our hypotheses on topic specificity, we removed all documents shorter than six paragraphs to provide a sufficient amount of text to evaluate topic distribution. Note that results do not change in a meaningful way when all documents (including those having less than six paragraphs) are included (see table S4).

Lexical cohesion

Texts provide the means to objectively measure lexical cohesion. Cohesive devices in text include word substitution, pronominal reference, conjunctions, and lexical repetition. Here, we computed lexical cohesion features using the Tool for the Automatic Analysis of Cohesion (TAACO) (*30*), a freely available standalone application that allows batch processing of text files. In TAACO, cohesion measures are extracted using several methods such as computing type/token ratios, as well as lexical and semantic overlaps for different part-of-speech categories. For our purpose, investigating semantic cohesion (i.e., different topics within a text across paragraphs), we used measures of semantic overlap. This is computed in TAACO with three computational models: latent semantic analysis (LSA) (*55*), LDA (*29*), and Word2vec (*56*). Different from other word-counting tools such as the Linguistic Inquiry and Word Count (LIWC) (*57*), these probabilistic models are capable of extracting the underlying semantic relations in texts. These models are based on unsupervised machine learning algorithms, meaning that human biases (e.g., associating a word to a category) are minimized. Last, TAACO outperforms similar model-based tools such as Coh-Metrix (*58*) because its semantic space is built on a larger corpus (~219 million words), and correlations with human rating of coherence were stronger than those with Coh-Metrix (*30*). TAACO uses these models to provide measures of lexical cohesion for segments within a text, i.e., adjacent paragraphs. For the LSA and the Word2vec models, TAACO computes similarity scores by the CS between segments (ranging between 0 and 1). LDA scores are computed using the Jensen-Shannon divergence between the normalized summed vector weights for the words in each segment (ranging between 0 and 1).

Extracting cohesion from documents

To obtain an output from TAACO, we first needed to feed it with a batch of documents. For this purpose, we first exported all documents as text files. Note that we did not perform any text preprocessing before this step (e.g., removing stop words or stemming) because TAACO performs analysis on the parsed text that needs to be syntactically valid. From the TAACO output, we extracted the three sets of measures computed with LDA, LSA, and Word2vec models. Specifically, we extracted the measures that computed dis/similarity between all adjacent paragraphs. While LSA and Word2vec outputs are in the form of similarity (LSA CS and Word2vec similarity scores, respectively), LDA output was computed as divergence; hence, LDA scores were reversed (i.e., by subtracting them from one). To obtain a single score of similarity for each document, we aggregated all three measures by computing the mean. Before testing our hypotheses on lexical cohesion, we removed all webpage documents whose length was less than six paragraphs, for reasons described above (note that, using all documents, results do not change in a meaningful way). Cohesion scores are assigned to each document and range from 0 (low cohesion) to 1 (high cohesion).

Testing cohesion metrics

To test whether TAACO measures the within-document lexical cohesion, we generated a sample of documents whose internal cohesion was artificially lowered. To this aim, we created a sample of "synthetic" documents composed of scrambled paragraphs randomly obtained from LOCO and tested whether cohesion was lower compared to a sample of "natural" documents. Our reasoning is that we cannot simply use TAACO to test differences between conspiracy and nonconspiracy documents because we do not know (i) whether TAACO is capable of detecting cohesion differences and (ii) whether there are real differences in lexical cohesion between conspiracy and nonconspiracy documents. Therefore, we first tested TAACO on documents that we know a priori are different, namely, by creating two groups of documents in which there are true differences in cohesion. It follows that if we find differences between these two groups (synthetic versus natural), then TAACO is capable of detecting cohesion differences.

To build our two test corpora, we first selected a random sample of 1000 documents from LOCO (both nonconspiracy and conspiracy) and created a bag of paragraphs (N = 19,528). We then selected a random sample of 500 nonconspiracy documents from LOCO and kept those that had at least six paragraphs, obtaining 385 nonconspiracy documents, whose length was 18.32 paragraphs on average (SD = 14.56). This set of high-cohesion, natural documents is used as a control group for the low-cohesion scrambled, i.e., synthetic, group. To build low-cohesion documents, we extracted the exact number of paragraphs from each of the high-cohesion document and generated a matched-by-length scrambled version. This resulted in a set of scrambled documents as large as the set of natural documents with the exact number of paragraphs (hence, two groups of N = 385 documents). For example, if a high-cohesion document is composed of 18 paragraphs, then we create its low-cohesion version by taking 18 paragraphs randomly selected from the bag of paragraphs. The two sets of documents did not differ in word count: t_{753} = 0.176, P = 0.86, d = 0.01. Using TAACO, we extracted between-paragraphs cohesion metrics and aggregated them in a unique score for each document. A t test between the two groups showed that the synthetic documents had lower cohesion than the natural ones: $t_{651.697}$ = 42.05, P < 0.001, d = 3.03 (synthetic: M = 0.363, SD = 0.028; natural: M = 0.477, SD = 0.045). We conclude that the cohesion metrics (and their aggregated value) are reliable in capturing differences in text cohesion.

Document similarity

To test the between-document similarity, we used the documents' pairwise CS with other documents in the same subcorpora (i.e., conspiracy and nonconspiracy). We used CS instead of other measures, for example, Jaccard similarity, because while the latter relies on unique word overlaps, CS is more sensitive to repetitions [CS and Jaccard similarity are highly correlated (in LOCO: r_{96741} = 0.82, P < 0.001)].

One could argue that a similarity score that is tested on all documents within a subcorpus might be inflated if documents rely on the same topic, simply because of word overlap. In addition, similarity between documents extracted from the same website might also be inflated by authors that might copy and paste pieces of narratives across webpages. We control for these confounds by computing the pairwise similarity of each document with the remaining documents in the subcorpus that (i) were not extracted from the same website and (ii) did not have the same seed. Texts were preprocessed following the same method used to extract LDA topics. Documents' similarity scores were computed using the *textstat_simil* function from the R package quanteda (*54*). Values range from 0 to 1, indicating either no overlap (0) or a perfect overlap (1) of terms. The returned output length of the CS for each document was a vector whose length was equal to the number of documents against which the similarity was tested. We therefore averaged this vector, obtaining a unique value for each document.

Statistical analyses (testing H1, H2, and H3)

To test H1, i.e., conspiracy networks have a higher interconnectedness than nonconspiracy networks, we extracted the degree of interconnectedness by counting the number of edges for each node in the network. To statistically test whether interconnectedness was higher in conspiracy compared to nonconspiracy networks, we ran linear mixed-effects models using the lme4 and the lmerTest R packages (*59*, *60*). In each model, we predicted the number of edges by the subcorpus (i.e., conspiracy and nonconspiracy), clustering observations within nodes. Note that this is similar to running a paired t test (β coefficients from these models are equal to the Cohen's d obtained from t tests). We preferred to rely on these multilevel models—instead of t tests—for consistency, so results are expressed in the same format throughout the paper.

For each network, we also provide descriptive statistics. In particular, we measured (i) entropy [with the function *graph.entropy* from the R package statGraph (*61*)], related to the extent to which nodes in a network are interconnected in a random, nonsystematic way; (ii) similarity [with the function *similarity* from the R package igraph (*53*)], which measures of how similar are connection patterns within a network; (iii) clustering (with the function *transitivity* from the package igraph), which calculates the probability that adjacent vertices of a node are interconnected; (iv) distance (with the function *mean_distance* from the package igraph), which extracts the average shortest path length through nodes; and (v) density (with the function *edge_density* from the package igraph), which computes the ratio of the number of edges to the number of possible edges.

To test H2 and H3, for each dependent variable (topic specificity, lexical cohesion, and CS), we ran a series of linear mixed-effects models using the lme4 and the lmerTest R packages. In each model, we specified as fixed effects the dichotomous subcorpus variable (i.e., whether the document is conspiracy or nonconspiracy) and added document word count as a covariate. As random intercept, we specified the websites from which documents were extracted. Theoretically, it is reasonable to assume that longer documents have space to accommodate more topics than shorter documents, which, consequently, decreases topic specificity and lexical cohesion. Likewise, larger documents, with a potential larger vocabulary, have higher chances to resemble the whole subcorpus vocabulary, hence resulting in high CS scores. Conspiracy documents are longer than nonconspiracy documents in word count: $t_{32,452} = 47.11$, $P < 0.001$, $d = 0.35$ (conspiracy: $M = 1236$, SD = 1307; nonconspiracy: $M = 806$, SD = 939). Second, word count correlates with our dependent variables (see tables S3 and S7). Thus, we include word count as a covariate in our analyses.

For multilevel models, as measures of effect sizes, we report the standardized regression coefficients beta (β) for predictors of interest and measures of fit such as R^2 (*62*). We report both marginal and conditional R^2 ($R^2_{m/c}$) associated with the variance explained by the fixed effects (marginal) and the variance explained by the entire model that includes both fixed and random effects (conditional), respectively. As a measure of effect size for *t* tests, we use Cohen's *d*. Note that because of the large samples we used in our main analyses, most *P* values are significant at $P < 0.001$. Although we report all *P* values from our analyses following the APA (American Psychological Association) style, we suggest readers focus more on the effect sizes and variance explained by the models, which vary greatly between analyses.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abq3668

REFERENCES AND NOTES

1. K. M. Douglas, J. E. Uscinski, R. M. Sutton, A. Cichocka, T. Nefes, C. S. Ang, F. Deravi, Understanding conspiracy theories. *Polit. Psychol.* **40**, 3–35 (2019).
2. J. E. Oliver, T. Wood, Medical conspiracy theories and health behaviors in the United States. *JAMA Intern. Med.* **174**, 817–818 (2014).
3. J. Allen, No consensus on who was behind Sept 11: Global poll (Reuters, 2008); www.reuters.com/article/us-sept11-qaeda-poll-idUSN1035876620080910.
4. J.-W. van Prooijen, K. M. Douglas, Conspiracy theories as part of history: The role of societal crisis situations. *Mem. Stud.* **10**, 323–333 (2017).
5. K. M. Douglas, COVID-19 conspiracy theories. *Group Process. Intergroup Relat.* **24**, 270–275 (2021).
6. L. M. Bogart, S. Thorburn, Are HIV/AIDS conspiracy beliefs a barrier to HIV prevention among African Americans? *J. Acquir. Immune Defic. Syndr.* **38**, 213–218 (2005).
7. D. Jolley, K. M. Douglas, The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint. *Br. J. Psychol.* **105**, 35–56 (2014).
8. R. Imhoff, L. Dieterle, P. Lamberty, Resolving the puzzle of conspiracy worldview and political activism: Belief in secret plots decreases normative but increases nonnormative political engagement. *Soc. Psychol. Personal. Sci.* **12**, 71–79 (2021).
9. D. Jolley, J. L. Paterson, Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *Br. J. Soc. Psychol.* **59**, 628–640 (2020).
10. D. Jolley, K. M. Douglas, A. C. Leite, T. Schrader, Belief in conspiracy theories and intentions to engage in everyday crime. *Br. J. Soc. Psychol.* **58**, 534–549 (2019).
11. P. Davies, Antivaccination activists on the world wide web. *Arch. Dis. Child.* **87**, 22–25 (2002).
12. J. Tollefson, Tracking QAnon: How Trump turned conspiracy-theory research upside down. *Nature* **590**, 192–193 (2021).
13. P. Ball, Anti-vaccine movement could undermine efforts to end coronavirus pandemic, researchers warn. *Nature* **581**, 251–251 (2020).
14. V. Swami, T. Chamorro-Premuzic, A. Furnham, Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs. *Appl. Cogn. Psychol.* **24**, 749–761 (2010).
15. M. J. Wood, K. M. Douglas, R. M. Sutton, Dead and alive. *Soc. Psychol. Personal. Sci.* **3**, 767–773 (2012).
16. T. Goertzel, Belief in conspiracy theories. *Polit. Psychol.* **15**, 731 (1994).
17. R. C. van der Wal, R. M. Sutton, J. Lange, J. P. N. Braga, Suspicious binds: Conspiracy thinking and tenuous perceptions of causal connections between co-occurring and spuriously correlated events. *Eur. J. Soc. Psychol.* **48**, 970–989 (2018).
18. J.-W. van Prooijen, K. M. Douglas, C. De Inocencio, Connecting the dots: Illusory pattern perception predicts belief in conspiracies and the supernatural. *Eur. J. Soc. Psychol.* **48**, 320–335 (2018).
19. E. Lobato, J. Mendoza, V. Sims, M. Chin, Examining the relationship between conspiracy theories, paranormal beliefs, and pseudoscience acceptance among a university population. *Appl. Cogn. Psychol.* **28**, 617–625 (2014).
20. H. Darwin, N. Neave, J. Holmes, Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Pers. Individ. Differ.* **50**, 1289–1293 (2011).
21. S. Lewandowsky, J. Cook, E. Lloyd, The 'Alice in Wonderland' mechanics of the rejection of (climate) science: Simulating coherence by conspiracism. *Synthese* **195**, 175–196 (2018).
22. V. Swami, R. Coles, S. Stieger, J. Pietschnig, A. Furnham, S. Rehim, M. Voracek, Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories. *Br. J. Psychol.* **102**, 443–463 (2011).
23. T. R. Tangherlini, S. Shahsavari, B. Shahbazi, E. Ebrahimzadeh, V. Roychowdhury, An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, Pizzagate and storytelling on the web. *PLOS ONE* **15**, e0233879 (2020).
24. B. Franks, A. Bangerter, M. W. Bauer, Conspiracy theories as quasi-religious mentality: An integrated account from cognitive science, social representations theory, and frame theory. *Front. Psychol.* **4**, 424 (2013).
25. A. Bangerter, P. Wagner-Egger, S. Delouvée, How conspiracy theories spread, in *Routledge Handbook of Conspiracy Theories*, M. Butter, P. Knight, Eds. (Routledge, 2020), pp. 206–218.
26. P. Lukić, I. Žeželj, B. Stanković, How (ir)rational is it to believe in contradictory conspiracy theories? *Eur. J. Psychol.* **15**, 94–107 (2019).
27. S. Oswald, Conspiracy and bias: Argumentative features and persuasiveness of conspiracy theories. *OSSA Conf. Arch.* **168**, 1–16 (2016).
28. A. Miani, T. Hills, A. Bangerter, LOCO: The 88-million-word language of conspiracy corpus. *Behav. Res. Methods* **54**, 1794–1817 (2022).
29. D. M. Blei, A. Y. Ng, M. I. Jordan, Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022 (2003).
30. S. A. Crossley, K. Kyle, M. Dascalu, The Tool for the Automatic Analysis of Cohesion 2.0: Integrating semantic similarity and text overlap. *Behav. Res. Methods.* **51**, 14–27 (2019).
31. A. Bruns, S. Harrington, E. Hurcombe, 'Corona? 5G? or both?': The dynamics of COVID-19/5G conspiracy theories on Facebook. *Media Int. Aust.* **177**, 12–29 (2020).
32. N. Dagnall, K. Drinkwater, A. Parker, A. Denovan, M. Parton, Conspiracy theory and cognitive style: A worldview. *Front. Psychol.* **6**, 206 (2015).
33. M. Rodier, M. Prévost, L. Renoult, C. Lionnet, Y. Kwann, E. Dionne-Dostie, I. Chapleau, J. B. Debruille, Healthy people with delusional ideation change their mind with conviction. *Psychiatry Res.* **189**, 433–439 (2011).
34. R. Moulding, S. Nix-Carnell, A. Schnabel, M. Nedeljkovic, E. E. Burnside, A. F. Lentini, N. Mehzabin, Better the devil you know than a world you don't? Intolerance of uncertainty and worldview explanations for belief in conspiracy theories. *Pers. Individ. Differ.* **98**, 345–354 (2016).
35. V. Juárez-Ramos, J. L. Rubio, C. Delpero, G. Mioni, F. Stablum, E. Gómez-Milán, Jumping to conclusions bias, BADE and feedback sensitivity in schizophrenia and schizotypy. *Conscious. Cogn.* **26**, 133–144 (2014).
36. U. Ettinger, I. Meyhöfer, M. Steffens, M. Wagner, N. Koutsouleris, Genetics, cognition, and neurobiology of schizotypal personality: A review of the overlap with schizophrenia. *Front. Psychiatry.* **5**, 18 (2014).

37. N. Rezaii, E. Walker, P. Wolff, A machine learning approach to predicting psychosis using semantic density and latent content analysis. *NPJ Schizophr.* **5**, 9 (2019).
38. B. Elvevåg, P. W. Foltz, D. R. Weinberger, T. E. Goldberg, Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophr. Res.* **93**, 304–316 (2007).
39. M. C. Allé, J. Potheegadoo, C. Köber, P. Schneider, R. Coutelle, T. Habermas, J.-M. Danion, F. Berna, Impaired coherence of life narratives of patients with schizophrenia. *Sci. Rep.* **5**, 12934 (2015).
40. J. A. Willits, T. Rubin, M. N. Jones, K. S. Minor, P. H. Lysaker, Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. *Schizophr. Res.* **197**, 365–369 (2018).
41. J. S. Paulsen, R. Romero, A. Chan, A. V. Davis, R. K. Heaton, D. V. Jeste, Impairment of the semantic network in schizophrenia. *Psychiatry Res.* **63**, 109–121 (1996).
42. M. S. Aloia, M. L. Gourovitch, D. R. Weinberger, T. E. Goldberg, An investigation of semantic space in patients with schizophrenia. *J. Int. Neuropsychol. Soc.* **2**, 267–273 (1996).
43. M. V. Bronstein, G. Pennycook, A. Bear, D. G. Rand, T. D. Cannon, Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *J. Appl. Res. Mem. Cogn.* **8**, 108–117 (2019).
44. A. Anthony, R. Moulding, Breaking the news: Belief in fake news and conspiracist beliefs. *Aust. J. Psychol.* **71**, 154–162 (2019).
45. M. E. Koltko-Rivera, The psychology of worldviews. *Rev. Gen. Psychol.* **8**, 3–58 (2004).
46. E. Brugnoli, M. Cinelli, W. Quattrociocchi, A. Scala, Recursive patterns in online echo chambers. *Sci. Rep.* **9**, 20118 (2019).
47. V. Swami, M. Voracek, S. Stieger, U. S. Tran, A. Furnham, Analytic thinking reduces belief in conspiracy theories. *Cognition* **133**, 572–585 (2014).
48. D. Santos, B. Requero, M. Martín-Fernández, Individual differences in thinking style and dealing with contradiction: The mediating role of mixed emotions. *PLOS ONE* **16**, e0257864 (2021).
49. S. Phadke, M. Samory, T. Mitra, What makes people join conspiracy communities?: Role of social factors in conspiracy engagement. *Proc. ACM Hum. Comput. Interact.* **4**, 223 (2020).
50. D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, J. L. Zittrain, The science of fake news. *Science* **359**, 1094–1096 (2018).
51. S. Lewandowsky, U. K. H. Ecker, J. Cook, Beyond misinformation: Understanding and coping with the "post-truth" era. *J. Appl. Res. Mem. Cogn.* **6**, 353–369 (2017).
52. B. Grün, K. Hornik, topicmodels: An R package for fitting topic models. *J. Stat. Softw.* **40**, 1–30 (2011).
53. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal Complex Syst.* **1695**, 1–9 (2006).
54. K. Benoit, K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, A. Matsuo, quanteda: An R package for the quantitative analysis of textual data. *J. Open Source Softw.* **3**, 774 (2018).
55. T. K. Landauer, P. W. Foltz, D. Laham, An introduction to latent semantic analysis. *Discourse Process.* **25**, 259–284 (1998).
56. T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality. arXiv:1310.4546 [cs.CL] (16 October 2013).
57. Y. R. Tausczik, J. W. Pennebaker, The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* **29**, 24–54 (2010).
58. A. C. Graesser, D. S. McNamara, M. M. Louwerse, Z. Cai, Coh-Metrix: Analysis of text on cohesion and language. *Behav. Res. Methods Instrum. Comput.* **36**, 193–202 (2004).
59. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
60. A. Kuznetsova, P. B. Brockhoff, R. H. B. Christensen, lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).
61. D. R. da Costa, T. C. Ramos, G. E. C. Guzman, S. S. Santos, E. S. Lira, A. Fujita, statGraph: Statistical methods for graphs (2021); https://CRAN.R-project.org/package=statGraph.
62. J. Lorah, Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large Scale Assess. Educ.* **6**, 8 (2018).
63. D. Nguyen, M. Liakata, S. DeDeo, J. Eisenstein, D. Mimno, R. Tromble, J. Winters, How we do things with words: Analyzing text as social and cultural data. *Front. Artif. Intell.* **3**, 62 (2020).
64. A. Colin, J. Murdock, in *Dynamics Of Science: Computational Frontiers in History and Philosophy of Science*, G. Ramsey, A. de Block, Eds. (Pittsburgh Univ. Press, 2020).
65. M. Nikita, ldatuning: Tuning of the latent Dirichlet allocation models parameters (2020); https://CRAN.R-project.org/package=ldatuning.
66. B. Han, D. Duong, J. H. Sul, P. I. W. de Bakker, E. Eskin, S. Raychaudhuri, A general framework for meta-analyzing dependent studies with overlapping subjects in association mapping. *Hum. Mol. Genet.* **25**, 1857–1866 (2016).
67. W. Viechtbauer, Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2010).

Acknowledgments

Funding: We acknowledge that we received no funding in support for this research. **Author contributions:** Conceptualization: A.M., T.H., and A.B. Methodology: A.M. Investigation: A.M., T.H., and A.B. Visualization: A.M. Funding acquisition: None. Project administration: A.B. Supervision: A.B. and T.H. Writing—original draft: A.M. Writing—review and editing: A.M., T.H., and A.B. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Scripts for replication are available at https://osf.io/aqnmb.

Submitted 4 April 2022
Accepted 6 September 2022
Published 26 October 2022
10.1126/sciadv.abq3668

Science Advances

Interconnectedness and (in)coherence as a signature of conspiracy worldviews

Alessandro MianiThomas HillsAdrian Bangerter

Sci. Adv., 8 (43), eabq3668.

View the article online
https://www.science.org/doi/10.1126/sciadv.abq3668
Permissions
https://www.science.org/help/reprints-and-permissions

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Science Advances
AAAS

Supplementary Materials for

Interconnectedness and (in)coherence as a signature of conspiracy worldviews

Alessandro Miani, Thomas Hills, Adrian Bangerter

Corresponding author: Alessandro Miani, alessandro.miani@unine.ch

Sci. Adv. **8**, eabq3668 (2022)
DOI: 10.1126/sciadv.abq3668

This PDF file includes:

Materials and Methods for replication
Figs. S1 to S4
Tables S1 to S7
References

Materials and Methods for replication

At the time of writing, LOCO is the largest corpus of CTs available. However, the features of this corpus represent potential confounds. For example, conspiracy documents are on average longer (in word count) than non-conspiracy documents. The two conspiracy and non-conspiracy subcorpora are also unequally sized, representing, respectively, one fourth and three fourths of LOCO's documents. Furthermore, topics and seeds are not evenly distributed across documents (e.g., the seed "chemtrails" is over-represented in the conspiracy subcorpus). Although some of these limitations were considered while building our statistical models (e.g., adding word count as a covariate and nesting documents within websites), in the following analyses we attempt to remove these confounds by creating a series of different subsets of LOCO with particular features such as documents matched by topic and documents matched by length. These corpora allow us to replicate the analyses we previously performed on LOCO, excluding the possibility that our results were due to imbalanced features of LOCO (namely, word count between conspiracy and non-conspiracy documents, and different subcorpora group sizes).

Building corpora matched for topics

In this set of corpora, we aimed to obtain a sample of conspiracy and non-conspiracy documents matched for topics. That is, within each corpus, each topic has approximately the same ratio (around .50) of conspiracy and non-conspiracy documents.

For each document, we first extracted the topic (among the initial 200 topics per document in LOCO) with the highest gamma value (i.e., the probability a topic accounts for the document's content) to obtain a proxy of the most prevalent topic per document. This resulted in a table (see e.g., Table 4 in LOCO) whose rows indicated a topic while columns were associated to a subcorpus; each cell contained the number of documents on a specific topic for conspiracy and non-conspiracy subcorpora. For each topic, we computed the proportion of conspiracy and non-conspiracy documents. We then extracted only the topics whose ratio of conspiracy and non-conspiracy documents was within a window around .50 (that is, approximately equal proportion of conspiracy and non-conspiracy documents for a given topic). This procedure excluded topics appearing asymmetrically in either conspiracy or non-conspiracy documents. For example, if a topic is present in 100 documents, 10 of which are conspiracy (hence 90 non-conspiracy documents), then this is an unbalanced ratio of .10 (conspiracy) and .90 (non-conspiracy), meaning that such a topic was disproportionally present in non-conspiracy documents. Differently, if a topic is present in 300 documents, whose 155 of them are conspiracy, then, with a ratio of .52 (conspiracy) and .48 (non-conspiracy), this topic is more evenly distributed across the two subcorpora.

We applied this procedure with three windows around a ratio of .50, representing more or less balanced numbers of documents per topic for each subcorpora. In Table S5, we report sample size and proportion of non-conspiracy documents. These windows are: .50 ± .20 (i.e., within .30 and .70, corpus name *C1*, N = 18,512, $N_{conspiracy}$ = 7,819; $N_{non\text{-}conspiracy}$ = 10,693), .50 ± .10 (within .40 and .60, corpus name *C2*, N = 4,186, $N_{conspiracy}$ = 2,090; $N_{non\text{-}conspiracy}$ = 2,096), and .50 ± .05 (within .45 and .55, corpus name *C3*, N = 1,490, $N_{conspiracy}$ = 745; $N_{non\text{-}conspiracy}$ = 745).

Building corpora matched for word count

In this set of corpora, we aimed to obtain a sample of conspiracy and non-conspiracy documents matched for word count. We chose three different document lengths: medium, long, and short documents.

In the first corpus (corpus name *C4*), we selected medium-length documents whose length was within a window of ± 0.2 standard deviations (SD) around the mean (i.e., 912 words) of LOCO's documents length (i.e., word count between 700 and 1,124 words). We arbitrarily selected 0.2 SD (compared to e.g., 0.5 and 1 SDs) because this value (after having tried other ones) returned documents whose length was overall not statistically different between conspiracy and non-conspiracy documents (p = .709). This corpus is composed of 19,944 documents ($N_{conspiracy}$ = 5,077; $N_{non\text{-}conspiracy}$ = 14,867).

In the second corpus (corpus name *C5*), we selected long-length documents with long texts. We extracted documents whose length was within a window of ± 0.1 SDs around 1SD above the mean (i.e., 1,971 words) of LOCO's documents length (i.e., word count between 1,865 and 2,077 words), which resulted in documents was length was not statistically significant (p = .501). This corpus is composed of 1,497 documents ($N_{conspiracy}$ = 675; $N_{non\text{-}conspiracy}$ = 822).

In the third corpus (corpus name *C6*), we selected short-length documents with short texts. We extracted documents whose length was within a window of ± 0.1 SDs around 1SD below the mean (i.e., 383 words) of LOCO's documents length (i.e., word count between 277 and 489 words). Because this resulted in a highly unbalanced dataset ($N_{conspiracy}$ = 3,624; $N_{non\text{-}conspiracy}$ = 17,120), we selected a random sample of 1,000 documents in each subcorpus, resulting in a total of 1,843 unique documents ($N_{conspiracy}$ = 865; $N_{non\text{-}conspiracy}$ = 978) matched for length (p = .879).

Measures for replication

As for network analyses, for each corpus, we replicated the procedure to create networks from both seeds and LDA topics as described above.

The topic specificity measure was obtained by replicating our previous analyses. However, because LDA topic extraction relies on a specific corpus structure, for each of our corpora C1-C6 we re-extracted topics. First, the number of topics for each corpus needed to be estimated. In LDA, the "right" number of topics is determined by the goal of the task more than the data itself (*63*). LDA topics can be thought of as being analogous to the resolution of a microscope (*63*): if a fine-grained resolution is required, then a large number of topics is better; if the number of topics is small, these topics become more general (*64*). In the current analysis, we did not have any preference for topic resolution, but our main concern was to apply the same criterion for all corpora. To this aim, we relied on the R package *ldatuning* (*65*), which provides a function to estimate the number of topics within a corpus based on four different metrics obtained via an unsupervised fashion. For a given corpus, this estimation works by setting a priori a vector containing the number of topics (k) to be evaluated, by fitting an LDA model for each k, and by computing the goodness of fit for each of the four metrics. The function provides the four-output metrics on different scales, therefore, for each metric we rescaled the values from 0 to 1 (worst and best k, respectively, within metrics) and then we averaged the four metrics into a single score ranging from 0 to 1: the higher the value, the better k is.

To find the appropriate number of topics for each corpus, we proceeded in two steps: first, we created a vector v of ks, where v = {5, 10, 15, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200}. Once the best number of topics was returned, we reran the same process narrowing v to values around

the best k found in the first step. For example, if $k = 120$, then, in the second step, we set v to $v = \{105, 110, 115, 125, 130, 135\}$. If the chosen k was at the end of the vector v (*i.e.*, $v = 200$), we then expanded v to $k = 250$ and, if needed, to $k = 300$ in the third and last step. The final number of topics k for each corpus (C1-C6 and LOCO) is reported in Table S5.
As for lexical cohesion and document similarity, for each corpus, we replicated the procedure as described above.

Statistical analyses for replication
Models (one per measure) were fit with multilevel linear regressions adding word count as covariate and nesting documents within websites. Note that for connectedness, analyses relied on seeds and LDA topics (network nodes), instead of documents. Therefore, we could not add word count as covariate, and we could not nest observations within websites. Instead, we nested observations within seeds and topics.
We computed the effect sizes from each regression and aggregated by dependent variables. We relied on the function *esc_B* from the *esc* package to compute Hedges' *g* effect size from the regressions' coefficients. From the regression output, the function takes as input the unstandardized coefficient as well as the sample size for each group and the standard deviation of the dependent variable and returns Hedges' *g*.
Because corpora presented a certain degree of overlap of documents with each other (see Table S6), we accounted for overlapping documents. This was done by transforming the covariance structure of the data to inflate the standard error of each corpus as a function of shared documents (*66*). Once effect sizes and standard errors (along with N and N_{shared} documents) were obtained, we computed an aggregated effect size for each dependent variable relying on the R package *metafor* (*67*). In Table S7, we show the corpora (C1-C6) main variables correlations. Results are summarized in Fig. S3.

Table S1.
Statistics for conspiracy and non-conspiracy networks obtained from seeds.
Note. Connectedness is expressed as the average of edges per node (and SD in parenthesis).

	Conspiracy	***non-Conspiracy***
connectedness	429 (149)	128 (149)
entropy	3.649	3.104
similarity	0.803	0.396
clustering	0.926	0.631
distance	1.101	1.468
density	11.289	3.36

Table S2.
Statistics for the conspiracy and non-conspiracy networks obtained from the three LDA topic matrices (LDA100, LDA200, and LDA300).
Note. C: conspiracy network; nC: non-conspiracy network. Connectedness is expressed as the average of edges per node (and SD in parenthesis).

	LDA100		**LDA200**		**LDA300**	
	C	**nC**	**C**	**nC**	**C**	**nC**
connectedness	41.22 (14.74)	36.60 (15.85)	77.23 (29.72)	68.02 (30.81)	116.04 (43.30)	101.31 (44.66)
entropy	0.722	0.685	0.667	0.616	0.652	0.59
similarity	0.292	0.268	0.271	0.242	0.269	0.237
clustering	0.516	0.481	0.501	0.478	0.497	0.474
distance	1.584	1.631	1.614	1.66	1.612	1.662
density	0.416	0.37	0.388	0.342	0.388	0.339

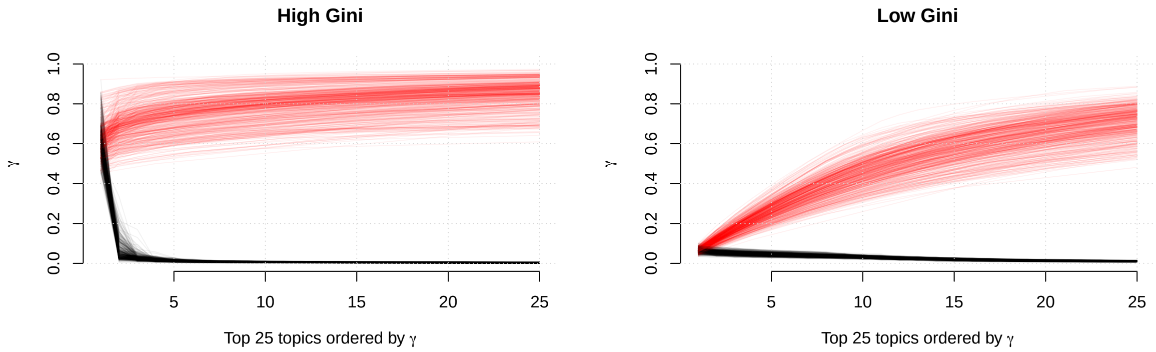


Fig. S1.
Gamma values distribution by documents with high (left) and low (right) Gini coefficient. Each line (N = 500 per plot) represents a document's gamma value (black) and its cumulative sum (red). In high-Gini documents (left), the cumulative topic probability (i.e., ɣ, on the Y axis) reaches around .70-.80 proportion of all topic distributions with less than five documents. Differently, in low-Gini documents (right), the cumulative topic probability reaches around .70-.80 proportion of all topic distributions with about 25 documents.

Table S3.
Correlation matrix of variables used in LOCO analyses.
All correlations are significant at *p* < .01 [95% confidence intervals].

Variable	1	2	3
1. Word count			
2. Topic specificity	-.05 [-.05, -.04]		
3. Lexical cohesion	.13 [.13, .14]	.19 [.18, .19]	
4. Cosine similarity	.52 [.51, .52]	-.29 [-.30, -.29]	-.03 [-.03, -.02]

Table S4.
Regression models for study 1.
Observations in within-document analyses (A) are fewer than between-document analysis (B) because for Topic specificity and Lexical cohesion we included documents with at least six paragraphs. Note that results do not change in a meaningful way when all documents (including those having less than six paragraphs) are included: Topic specificity: β = -0.302, *SE* = 0.06, t = -5.027, $p < .001$; Lexical cohesion: β = -0.418, *SE* = 0.076, t = -5.488, $p < .001$.

A) Within Documents

	Topic specificity					Lexical cohesion				
Predictors	*β*	*SE*	*95% CI*	*t*	*p*	*β*	*SE*	*95% CI*	*t*	*p*
Subcorpus [conspiracy]	-0.29	0.06	-0.42,-0.17	-4.66	**<0.001**	-0.68	0.08	-0.84,-0.51	-7.97	**<0.001**
Word count	-0.04	0.00	-0.05,-0.04	-11.26	**<0.001**	0.03	0.00	0.03,0.04	9.77	**<0.001**
Random Effects										
σ^2	0.89					0.76				
τ_{00}	0.13 websites					0.24 websites				
ICC	0.13					0.24				
N	148 websites					148 websites				
Observations	79,491					79,491				
Marginal R^2/ Conditional R^2	0.019 / 0.143					0.080 / 0.302				

B) Between Document

	Cosine similarity				
Predictors	*β*	*SE*	*95% CI*	*t*	*p*
Subcorpus [conspiracy]	0.96	0.06	0.83,1.08	14.87	**<0.001**
Word count	0.45	0.00	0.44,0.45	185.56	**<0.001**
Random Effects					
σ^2	0.46				
τ_{00} websites	0.14				
ICC	0.24				
N websites	150				
Observations	96743				
Marginal R^2 / Conditional R^2	0.422 / 0.559				

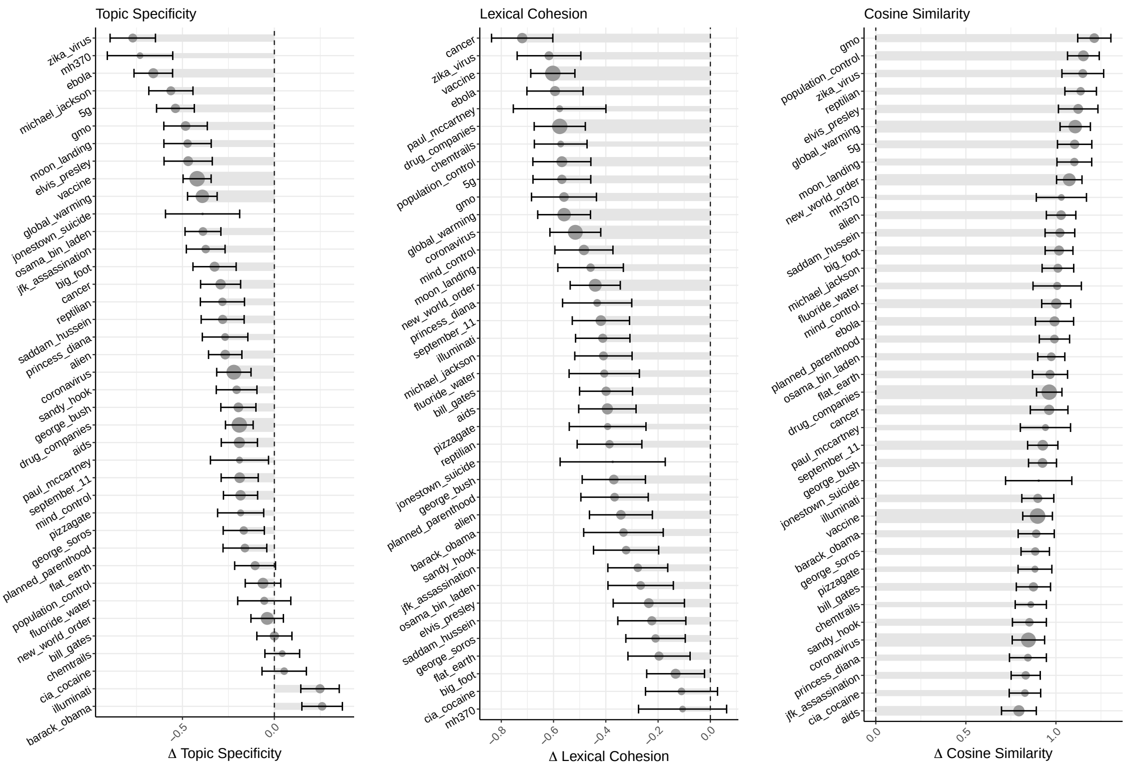


Fig. S2.
Estimated differences between conspiracy and non-conspiracy subcorpora in LOCO for each seed on within- (left and center) and between- (right) document measures.
Positive estimates indicate that the measures are higher in the conspiracy subcorpus. Each estimate was obtained by fitting linear mixed effects models (with wordcount as covariate and nesting documents within websites) for documents associated with the given seed. The size of the gray dots and shadows represent the sample size of documents associated with the given seed.

Table S5.
C1-C6 and LOCO corpora description
Docs: number of documents in the corpus; Tokens: number of words in the corpus; Types: number of unique words in corpus; d: Cohen's d from differences in word count between conspiracy and non-conspiracy subcorpora; * word count difference significant at $p < .001$; WC: mean (M) and standard deviation (SD) of word count for documents in the corpus; Ratio: ratio of conspiracy in the corpus; k: number of corpus' topics extracted with LDA. Note that LOCO's token size is different from the original paper (98 M vs 88 M, (*1*)); this is because here we counted words with the *quanteda* package that also allows to extract unique terms (for vocabulary size), while in LOCO we relied on TAACO.

Corpus	Docs	Tokens	Types	*d*	WC M (SD)		Ratio	*k*	notes
LOCO	96,743	98,151,459	574,113	.41*	912	(1,059)	0.25	200	LOCO
C1	18,512	20,754,678	207,301	.35*	1,008	(1,145)	0.42	300	Topic matched (window ± .20)
C2	4,186	5,053,969	92,516	.41*	1,083	(1,271)	0.50	140	Topic matched (window ± .10)
C3	1,490	2,081,375	56,197	.60*	1,258	(1,380)	0.50	110	Topic matched (window ± .5)
C4	19,944	19,479,347	199,794	.01	884	(119)	0.25	280	Length matched (medium)
C5	1,497	3,277,181	80,364	.04	1,968	(61)	0.45	90	Length matched (long)
C6	1,843	775,599	38,131	.01	382	(59)	0.47	70	Length matched (short)

Table S6.
Number of documents overlap between C1-C6 corpora and LOCO.
The number of documents shared with LOCO (N = 96,743) from the subset corpora (C1-C6) extracted from LOCO represents the sample size of these corpora. Also, the documents of topic-matched corpora follow the same principle since a restricted window (e.g., C3) is in fact a subset of larger windows (both C2 and C1). We retained these numbers in the matrix for consistency and highlighted them in **bold**.

	LOCO	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**
LOCO	-						
C1	**18,512**	-					
C2	**4,186**	**4,186**	-				
C3	**1,490**	**1,490**	**1,490**	-			
C4	**19,944**	4,044	852	320	-		
C5	**1,497**	315	79	35	0	-	
C6	**1,843**	417	83	16	0	0	-

Table S7.
Correlations between main variables in subcorpora C1-C6.
Values expressed as Pearson coefficient *r*. WC: word count; TS: topic specificity; CH: cohesion; CS: cosine similarity. The gray shadowed cells highlight our four main dependent variables with the word count variable that we used as a covariate in analyses.

	WC TS	WC CH	WC CS	TS CH	TS CS	CH CS
C1	.04	.11	.56	.17	-.27	.01
C2	.17	.15	.57	.17	-.13	.09
C3	.23	.15	.62	.12	-.09	.15
C4	-.04	.01	.13	.19	-.40	-.42
C5	.03	.01	.03	.32	-.59	-.50
C6	.05	.08	.25	.25	-.44	-.24

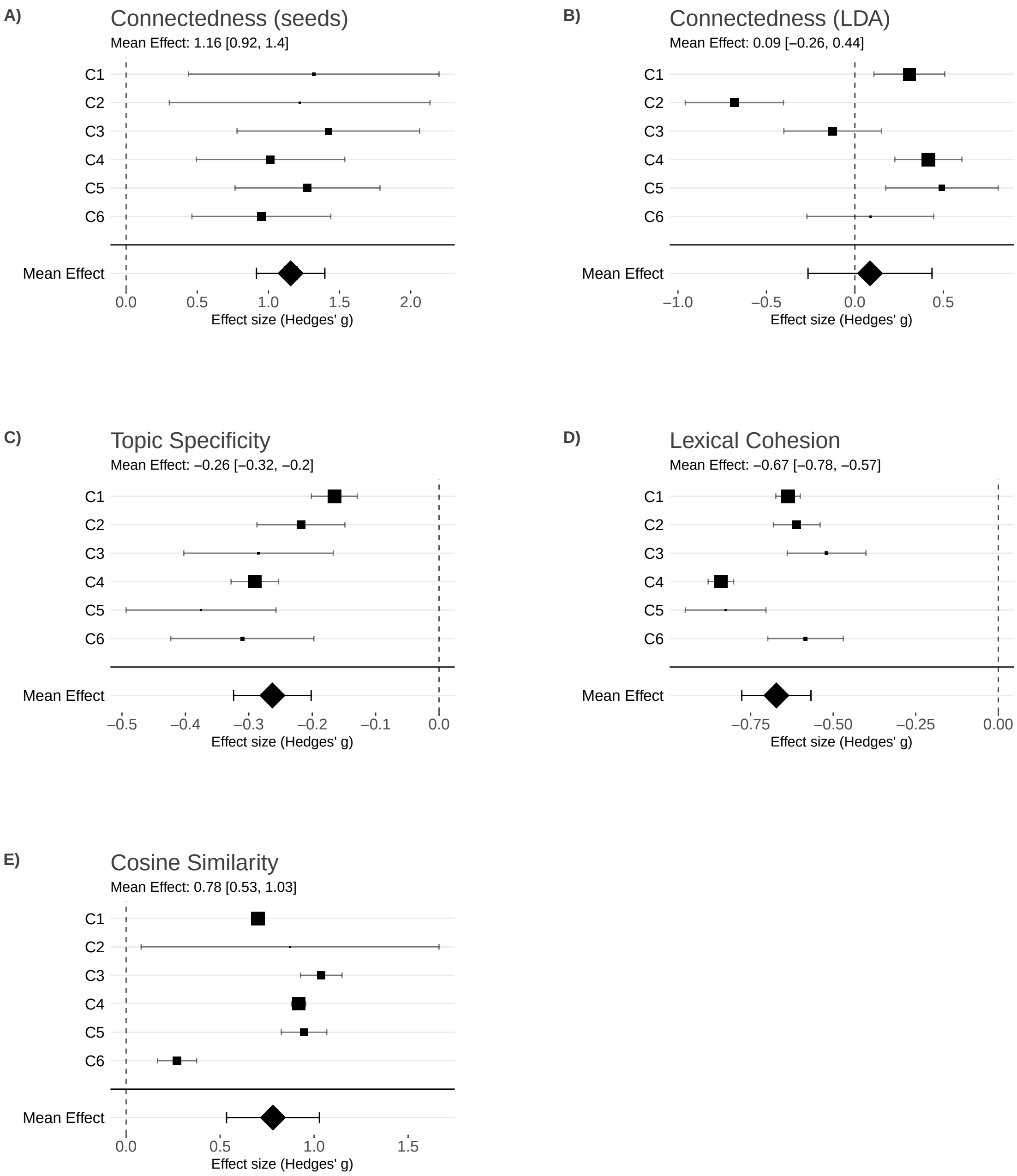


Fig. S3.
Individual and aggregated effect sizes (95% CIs) for C1-C6 models.
Positive effects indicate that the measures are higher in the conspiracy subcorpus. Mean effect size (Hedges' *g*) computed accounting for documents overlap between corpora. Models for Topic specificity, Lexical cohesion, and Cosine similarity were fit with multilevel linear regressions adding word count as covariate and nesting documents within websites. Models for Connectedness of seeds and LDA topics were fit with multilevel linear regressions nesting observations within nodes (i.e., seeds or LDA topics).

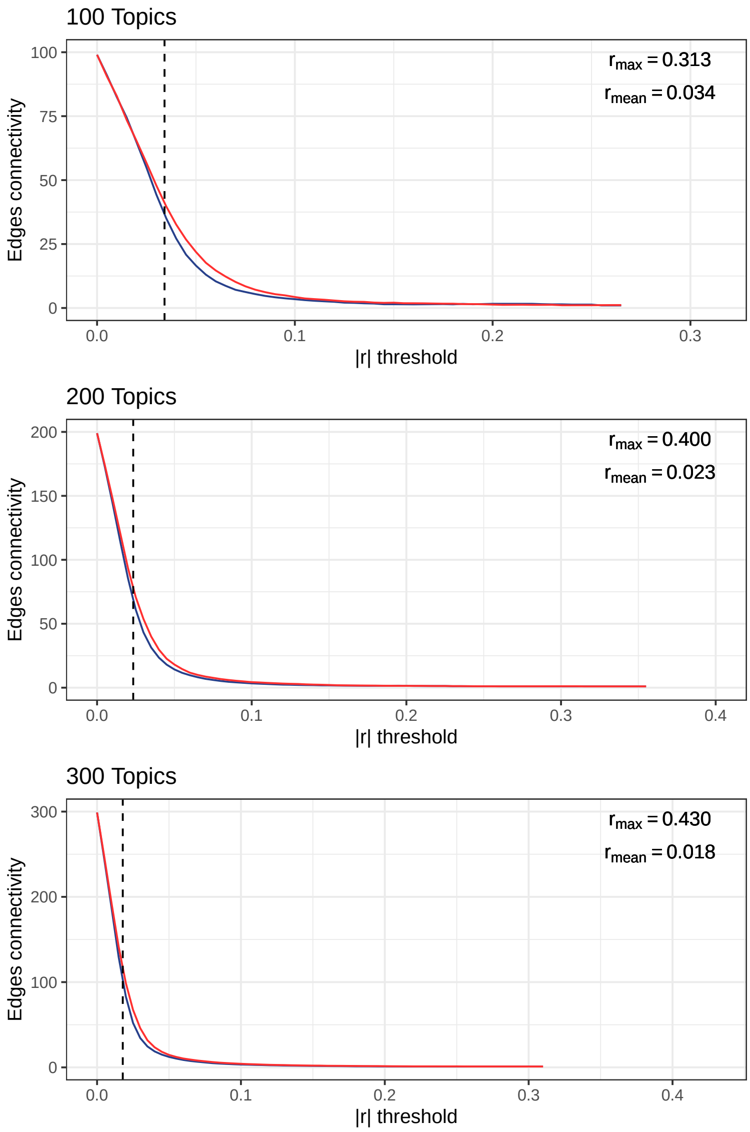


Fig. S4.
Edges connectivity as a function of |r| threshold for the three LDA matrices.
Edges connectivity (Y axis) by $|r|$ thresholds (X axis) for the three LDA correlation matrices. Lines refer to conspiracy (red) and non-conspiracy (blue) networks. r_{max} and r_{mean} refer to the maximum $|r|$ and the mean of $|r|$ coefficients in each correlation matrix. If the threshold is too low ($|r| = 0$ on the X axis), all topics are connected to each other, hence the mean average connectivity is equal to k, the number of all topics. If the threshold is too high, topic co-occurrences are not detected, hence nodes in the network are disconnected to each other resulting in an average degree connectivity of 0. The vertical dashed line represents the mean of $|r|$ coefficients in each correlation matrix, the threshold we have chosen above which we considered a topic co-occurrence.

REFERENCES AND NOTES

1. K. M. Douglas, J. E. Uscinski, R. M. Sutton, A. Cichocka, T. Nefes, C. S. Ang, F. Deravi, Understanding conspiracy theories. *Polit. Psychol.* **40**, 3–35 (2019).

2. J. E. Oliver, T. Wood, Medical conspiracy theories and health behaviors in the United States. *JAMA Intern. Med.* **174**, 817–818 (2014).

3. J. Allen, No consensus on who was behind Sept 11: Global poll (Reuters, 2008); www.reuters.com/article/us-sept11-qaeda-poll-idUSN1035876620080910.

4. J.-W. van Prooijen, K. M. Douglas, Conspiracy theories as part of history: The role of societal crisis situations. *Mem. Stud.* **10**, 323–333 (2017).

5. K. M. Douglas, COVID-19 conspiracy theories. *Group Process. Intergroup Relat.* **24**, 270–275 (2021).

6. L. M. Bogart, S. Thorburn, Are HIV/AIDS conspiracy beliefs a barrier to HIV prevention among African Americans? *J. Acquir. Immune Defic. Syndr.* **38**, 213–218 (2005).

7. D. Jolley, K. M. Douglas, The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint. *Br. J. Psychol.* **105**, 35–56 (2014).

8. R. Imhoff, L. Dieterle, P. Lamberty, Resolving the puzzle of conspiracy worldview and political activism: Belief in secret plots decreases normative but increases nonnormative political engagement. *Soc. Psychol. Personal. Sci.* **12**, 71–79 (2021).

9. D. Jolley, J. L. Paterson, Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *Br. J. Soc. Psychol.* **59**, 628–640 (2020).

10. D. Jolley, K. M. Douglas, A. C. Leite, T. Schrader, Belief in conspiracy theories and intentions to engage in everyday crime. *Br. J. Soc. Psychol.* **58**, 534–549 (2019).

11. P. Davies, Antivaccination activists on the world wide web. *Arch. Dis. Child.* **87**, 22–25 (2002).

12. J. Tollefson, Tracking QAnon: How Trump turned conspiracy-theory research upside down. *Nature* ***590***, 192–193 (2021).

13. P. Ball, Anti-vaccine movement could undermine efforts to end coronavirus pandemic, researchers warn. *Nature* **581**, 251–251 (2020).

14. V. Swami, T. Chamorro-Premuzic, A. Furnham, Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs. *Appl. Cogn. Psychol.* **24**, 749–761 (2010).

15. M. J. Wood, K. M. Douglas, R. M. Sutton, Dead and alive. *Soc. Psychol. Personal. Sci.* **3**, 767–773 (2012).

16. T. Goertzel, Belief in conspiracy theories. *Polit. Psychol.* **15**, 731 (1994).

17. R. C. van der Wal, R. M. Sutton, J. Lange, J. P. N. Braga, Suspicious binds: Conspiracy thinking and tenuous perceptions of causal connections between co-occurring and spuriously correlated events. *Eur. J. Soc. Psychol.* **48**, 970–989 (2018).

18. J.-W. van Prooijen, K. M. Douglas, C. De Inocencio, Connecting the dots: Illusory pattern perception predicts belief in conspiracies and the supernatural. *Eur. J. Soc. Psychol.* **48**, 320–335 (2018).

19. E. Lobato, J. Mendoza, V. Sims, M. Chin, Examining the relationship between conspiracy theories, paranormal beliefs, and pseudoscience acceptance among a university population. *Appl. Cogn. Psychol.* **28**, 617–625 (2014).

20. H. Darwin, N. Neave, J. Holmes, Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Pers. Individ. Differ.* **50**, 1289–1293 (2011).

21. S. Lewandowsky, J. Cook, E. Lloyd, The ‘Alice in Wonderland’ mechanics of the rejection of (climate) science: Simulating coherence by conspiracism. *Synthese* **195**, 175–196 (2018).

22. V. Swami, R. Coles, S. Stieger, J. Pietschnig, A. Furnham, S. Rehim, M. Voracek, Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories. *Br. J. Psychol.* **102**, 443–463 (2011).

23. T. R. Tangherlini, S. Shahsavari, B. Shahbazi, E. Ebrahimzadeh, V. Roychowdhury, An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, Pizzagate and storytelling on the web. *PLOS ONE* **15**, e0233879 (2020).

24. B. Franks, A. Bangerter, M. W. Bauer, Conspiracy theories as quasi-religious mentality: An integrated account from cognitive science, social representations theory, and frame theory. *Front. Psychol.* **4**, 424 (2013).

25. A. Bangerter, P. Wagner-Egger, S. Delouvée, How conspiracy theories spread, in *Routledge Handbook of Conspiracy Theories*, M. Butter, P. Knight, Eds. (Routledge, 2020), pp. 206–218.

26. P. Lukić, I. Žeželj, B. Stanković, How (ir)rational is it to believe in contradictory conspiracy theories? *Eur. J. Psychol.* **15**, 94–107 (2019).

27. S. Oswald, Conspiracy and bias: Argumentative features and persuasiveness of conspiracy theories. *OSSA Conf. Arch.* **168**, 1–16 (2016).

28. A. Miani, T. Hills, A. Bangerter, LOCO: The 88-million-word language of conspiracy corpus. *Behav. Res. Methods* **54**, 1794–1817 (2022).

29. D. M. Blei, A. Y. Ng, M. I. Jordan, Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022 (2003).

30. S. A. Crossley, K. Kyle, M. Dascalu, The Tool for the Automatic Analysis of Cohesion 2.0: Integrating semantic similarity and text overlap. *Behav. Res. Methods.* **51**, 14–27 (2019).

31. A. Bruns, S. Harrington, E. Hurcombe, ‘Corona? 5G? or both?’: The dynamics of COVID-19/5G conspiracy theories on Facebook. *Media Int. Aust.* **177**, 12–29 (2020).

32. N. Dagnall, K. Drinkwater, A. Parker, A. Denovan, M. Parton, Conspiracy theory and cognitive style: A worldview. *Front. Psychol.* **6**, 206 (2015).

33. M. Rodier, M. Prévost, L. Renoult, C. Lionnet, Y. Kwann, E. Dionne-Dostie, I. Chapleau, J. B. Debruille, Healthy people with delusional ideation change their mind with conviction. *Psychiatry Res.* **189**, 433–439 (2011).

34. R. Moulding, S. Nix-Carnell, A. Schnabel, M. Nedeljkovic, E. E. Burnside, A. F. Lentini, N. Mehzabin, Better the devil you know than a world you don't? Intolerance of uncertainty and worldview explanations for belief in conspiracy theories. *Pers. Individ. Differ.* **98**, 345–354 (2016).

35. V. Juárez-Ramos, J. L. Rubio, C. Delpero, G. Mioni, F. Stablum, E. Gómez-Milán, Jumping to conclusions bias, BADE and feedback sensitivity in schizophrenia and schizotypy. *Conscious. Cogn.* **26**, 133–144 (2014).

36. U. Ettinger, I. Meyhöfer, M. Steffens, M. Wagner, N. Koutsouleris, Genetics, cognition, and neurobiology of schizotypal personality: A review of the overlap with schizophrenia. *Front. Psychiatry.* **5**, 18 (2014).

37. N. Rezaii, E. Walker, P. Wolff, A machine learning approach to predicting psychosis using semantic density and latent content analysis. *NPJ Schizophr.* **5**, 9 (2019).

38. B. Elvevåg, P. W. Foltz, D. R. Weinberger, T. E. Goldberg, Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophr. Res.* **93**, 304–316 (2007).

39. M. C. Allé, J. Potheegadoo, C. Köber, P. Schneider, R. Coutelle, T. Habermas, J.-M. Danion, F. Berna, Impaired coherence of life narratives of patients with schizophrenia. *Sci. Rep.* **5**, 12934 (2015).

40. J. A. Willits, T. Rubin, M. N. Jones, K. S. Minor, P. H. Lysaker, Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. *Schizophr. Res.* **197**, 365–369 (2018).

41. J. S. Paulsen, R. Romero, A. Chan, A. V. Davis, R. K. Heaton, D. V. Jeste, Impairment of the semantic network in schizophrenia. *Psychiatry Res.* **63**, 109–121 (1996).

42. M. S. Aloia, M. L. Gourovitch, D. R. Weinberger, T. E. Goldberg, An investigation of semantic space in patients with schizophrenia. *J. Int. Neuropsychol. Soc.* **2**, 267–273 (1996).

43. M. V. Bronstein, G. Pennycook, A. Bear, D. G. Rand, T. D. Cannon, Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *J. Appl. Res. Mem. Cogn.* **8**, 108–117 (2019).

44. A. Anthony, R. Moulding, Breaking the news: Belief in fake news and conspiracist beliefs. *Aust. J. Psychol.* **71**, 154–162 (2019).

45. M. E. Koltko-Rivera, The psychology of worldviews. *Rev. Gen. Psychol.* **8**, 3–58 (2004).

46. E. Brugnoli, M. Cinelli, W. Quattrociocchi, A. Scala, Recursive patterns in online echo chambers. *Sci. Rep.* **9**, 20118 (2019).

47. V. Swami, M. Voracek, S. Stieger, U. S. Tran, A. Furnham, Analytic thinking reduces belief in conspiracy theories. *Cognition* **133**, 572–585 (2014).

48. D. Santos, B. Requero, M. Martín-Fernández, Individual differences in thinking style and dealing with contradiction: The mediating role of mixed emotions. *PLOS ONE* **16**, e0257864 (2021).

49. S. Phadke, M. Samory, T. Mitra, What makes people join conspiracy communities?: Role of social factors in conspiracy engagement. *Proc. ACM Hum. Comput. Interact.* **4**, 223 (2020).

50. D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, J. L. Zittrain, The science of fake news. *Science* **359**, 1094–1096 (2018).

51. S. Lewandowsky, U. K. H. Ecker, J. Cook, Beyond misinformation: Understanding and coping with the “post-truth” era. *J. Appl. Res. Mem. Cogn.* **6**, 353–369 (2017).

52. B. Grün, K. Hornik, topicmodels: An R package for fitting topic models. *J. Stat. Softw.* **40**, 1–30 (2011).

53. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal Complex Syst.* **1695**, 1–9 (2006).

54. K. Benoit, K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, A. Matsuo, quanteda: An R package for the quantitative analysis of textual data. *J. Open Source Softw.* **3**, 774 (2018).

55. T. K. Landauer, P. W. Foltz, D. Laham, An introduction to latent semantic analysis. *Discourse Process.* **25**, 259–284 (1998).

56. T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality. arXiv:1310.4546 [cs.CL] (16 October 2013).

57. Y. R. Tausczik, J. W. Pennebaker, The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* **29**, 24–54 (2010).

58. A. C. Graesser, D. S. McNamara, M. M. Louwerse, Z. Cai, Coh-Metrix: Analysis of text on cohesion and language. *Behav. Res. Methods Instrum. Comput.* **36**, 193–202 (2004).

59. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

60. A. Kuznetsova, P. B. Brockhoff, R. H. B. Christensen, lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).

61. D. R. da Costa, T. C. Ramos, G. E. C. Guzman, S. S. Santos, E. S. Lira, A. Fujita, statGraph: Statistical methods for graphs (2021); https://CRAN.R-project.org/package=statGraph.

62. J. Lorah, Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large Scale Assess. Educ.* **6**, 8 (2018).

63. D. Nguyen, M. Liakata, S. DeDeo, J. Eisenstein, D. Mimno, R. Tromble, J. Winters, How we do things with words: Analyzing text as social and cultural data. *Front. Artif. Intell.* **3**, 62 (2020).

64. A. Colin, J. Murdock, in *Dynamics Of Science: Computational Frontiers in History and Philosophy of Science*, G. Ramsey, A. de Block, Eds. (Pittsburgh Univ. Press, 2020).

65. M. Nikita, ldatuning: Tuning of the latent Dirichlet allocation models parameters (2020); https://CRAN.R-project.org/package=ldatuning.

66. B. Han, D. Duong, J. H. Sul, P. I. W. de Bakker, E. Eskin, S. Raychaudhuri, A general framework for meta-analyzing dependent studies with overlapping subjects in association mapping. *Hum. Mol. Genet.* **25**, 1857–1866 (2016).

67. W. Viechtbauer, Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2010).

5.3 Study 3

Overinclusive thinking in conspiracy texts

Alessandro Miani[1*], Lonneke van der Plas[2], and Adrian Bangerter[1]

[1] Institute of Work and Organizational Psychology, University of Neuchâtel, Switzerland
[2] Idiap Research Institute, Martigny, Switzerland

Word count (with abstract): 2,969 [limit 3,000]

All data and scripts for replication are available at the project's Open Science Framework page (http://osf.io/WORK_ON_THIS).

Conflicts of interest: none.

Funding: none.

* Correspondence concerning this article should be addressed to Alessandro Miani, Institute of Work and Organizational Psychology, University of Neuchâtel, Rue Emile-Argand 11, 2000, Neuchâtel, Switzerland. Email: alessandro.miani@unine.ch

Abstract

Conspiracy theories (CTs) are spectacular narratives, widely spread, that pose societal threats. We test whether CTs might be linguistically creative products, which would facilitate their transmission. We analyzed nominal compounds (e.g., *chocolate cake*) from a large corpus of conspiracy and non-conspiracy texts matched by topic. Compounds in conspiracy texts were more original, divergent, and sophisticated, but less appropriate to their context and variable compared to those in non-conspiracy texts. Results suggest CT texts involve ideas generated by divergent thinking but are not filtered out by convergent thinking. Resonating with studies showing overinclusive thinking in conspiracy theorizing, an imbalance between divergent and convergent thinking might explain the accumulation of CTs within people's belief systems.

Keywords: conspiracy theories, creativity, divergent thinking, persuasion, natural language processing, nominal compounds

1 Introduction

Conspiracy theories (CTs) are accounts of significant social and political events alleging secret plots by powerful actors (Douglas et al., 2019). Diverging from mainstream accounts, CTs are often spectacular, counter-intuitive, and entertaining (van Prooijen, 2018; van Prooijen et al., 2022). CTs are widely endorsed (Allen, 2008; Oliver & Wood, 2014) and capable of mobilizing people (Imhoff et al., 2021; Jolley & Paterson, 2020) causing detrimental societal effects (Ball, 2020; Jolley & Douglas, 2014). They compete for attention with mainstream news, but being freed from constraints of veridicality they can exploit cognitive biases for appealing information (Acerbi, 2019; Hills, 2019; van Prooijen et al., 2022). Thus, false information is more likely to be transmitted online, eliciting responses of surprise (Vosoughi et al., 2018). This transmission advantage might stem from content features such as novelty, divergence from the norm, and sophistication that, orchestrated perhaps strategically, make narratives more attention-grabbing (Oswald, 2016; Vosoughi et al., 2017, 2018).

If standard definitions of creativity involve novelty and divergence from the norm combined with effectiveness (Runco, 2010; Runco & Jaeger, 2012), then CTs may fulfill some of the criteria of creative products (Bonetto & Arciszewski, 2021). More precisely, CTs may reflect the "dark side" of creativity (D. H. Cropley et al., 2008), especially in cases where transmission is motivated strategically (Allcott & Gentzkow, 2017; Douglas et al., 2019).

Like creative products, CTs are effective, as measured by their successful cultural transmission. CTs are based on a rhetoric of persuasiveness that emulates formal features of academic discourse using a multitude of arguments (Miani et al., 2022; Oswald, 2016). Malicious (vs non-malicious) rumors feature high levels of linguistic sophistication, perhaps to make the message look more legitimate and believable (Vosoughi et al., 2017). CTs combine intuitive with counterintuitive concepts involving omniscient and omnipotent agents (Franks et al., 2013), making them more memorable and more easily transmitted (Norenzayan et al., 2006). Need for uniqueness might explain people's consumption and transmission of CTs (Biddlestone et al., 2022) to enhance personal and social identity (Sternisko et al., 2020) by feeling special (Tian et al., 2001). For example, people find CTs interesting, exciting, and attention-grabbing, especially those high in sensation-seeking (van Prooijen et al., 2022), who value novelty. Uniqueness-motivated individuals show characteristics such as anti-conformity, inventiveness and innovation (Skinner, 1996), and need for uniqueness is associated with creativity of written stories (Dollinger, 2003) and infrequent or novel responses in word association tests (Snyder & Fromkin, 1977).

Some scholars, however, suggest that mere novel and unusual ideas without reference to reality cannot be considered creative. Creativity is in fact a process where convergent thinking acts upon ideas generated from divergent thinking (A. Cropley, 2006; Guilford, 1950, 1967; Runco & Jaeger, 2012). Convergent thinking helps filter out uncommon and merely random ideas generated from ignorance or delusion (F. Barron, 1955), which are often the product of "nonconformity, lack of discipline, and blind rejection of what already exists […] regardless of accuracy" (A. Cropley, 2006, p. 392).

Divergence not controlled by convergence leads to overinclusive thinking, a cognitive processing style characterized by lack of conceptual boundaries and over-responsiveness to irrelevant stimuli (Wang et al., 2018). It is related to psychoticism and the schizophrenia spectrum (F. Barron, 1993; Eysenck, 2003; Runco, 2010), which is linked with both belief in CTs (D. Barron et al., 2014) and creativity (Batey & Furnham, 2008). Compared to healthy controls, patients with schizophrenia show lowered semantic cohesion (Willits et al., 2018) and alterations in semantic networks (Paulsen et al., 1996) that are also found in creative individuals (Kenett et al., 2014), displaying enhanced connectivity. Similarly, semantic networks built from conspiracy narratives are hyper-connected (Miani et al., 2022). CT believers, attempting to make sense of events (Bangerter et al., 2020; Franks et al., 2013), are moved to search for hidden motives in search for truth, appearing to consider possible alternative scenarios and explore their environment more carefully (Mayer & Mussweiler, 2011). In doing so, believers tend to identify meaningful relationships among randomly co-occurring events (van der Wal et al., 2018), confuse aspects of reality such as believing that prayers have the capacity to heal (Lobato et al., 2014), and believe in the paranormal (Darwin et al., 2011).

To date, besides a theoretical contribution (Bonetto & Arciszewski, 2021), no empirical work has addressed the link between CTs and creativity and, specifically, overinclusive thinking.

2 This study

Here, we explore whether CTs are products of creativity or overinclusive thinking (i.e., an imbalance between convergent and divergent thinking). In creativity research, measures of individual creativity include self-report scales, producing metaphors, drawings, stories, essays, or listing uncommon uses for common objects (Dollinger, 2003; Dollinger et al., 2004; Runco, 2010, 2010; Silvia et al., 2012). A drawback of these measures is that they require costly scoring by human raters and are also subject to measurement error from raters' individual differences. Thus, algorithms are increasingly used to automatically and objectively assess creativity from people's responses, generating scores that correlate with human ratings (Dumas et al., 2021).

A problem with lab studies, however, is that generalized tests do not reliably predict domain-specific creativity (e.g., predicting music creativity from how many uses a person can list for a brick; Simonton, 1999). A more ecologically valid approach, then, would require shifting the focus towards the fields in which creativity is employed for a specific goal. Conspiracy narratives found on websites represent a suitable source for investigating creativity because are specialized sources created for the purpose of developing, collecting, and spreading CTs: to be more credible and memorable hence increasing persuasion, these narratives could rely on creativity. A recent large-scale text analysis on conspiracy webpages found that, compared to non-conspiracy narratives, conspiracy narratives, are internally less coherent but more interconnected in a dense and unstructured network of shared topics (Miani et al., 2022). These patterns suggest that associative networks of conspiracy texts are different from those of non-conspiracy texts, converging with existing research on creative individuals and schizophrenic populations (Elvevåg et al., 2007; Kenett et al., 2014).

Leveraging methods from creativity research, here we move from the macro level of textual coherence and topical analysis to the micro level of lexical choice (e.g., Olson et al., 2021). A very productive way in which new words are formed in English is by means of compounding. Whenever there is need for a speech community to name a new concept, new words are coined and this often happens by means of combining existing words to create a new one. For this purpose, nominal compounds represent a suitable material for modeling linguistic creativity (Dhar & van der Plas, 2019; Körtvélyessy et al., 2021), because they provide access to the structure and content of the mental lexicon and creative processing (Gagné, 2002). Nominal compounds are concatenations of two or more nouns (e.g., *chocolate cake*) functioning as a nominal multiword expression, encapsulating manifold semantic relations between their constituents (Aitchison, 1987). They represent the communicator's creative choices, *the ideas* (Runco, 2010), of compressing a sentence into a combination of words attempting to minimize information loss (Körtvélyessy et al., 2021).

Here, we extract from nominal compounds a series of measures that are widely employed in creativity research such as *fluency* (number of ideas generated), *flexibility* (number of different ideas generated), *originality* (uniqueness and newness of ideas), *divergence* (conceptual distance of ideas), *appropriateness* (contextual fitness of ideas), and *complexity* (sophistication of ideas). High scores in all measures would indicate creativity, whereas high divergence and low appropriateness, would indicate the presence of overinclusive thinking.

3 Method

3.1 Material

Our text material is from the language of conspiracy (LOCO) corpus (Miani et al., 2021). LOCO is a freely available, multilevel and topic-matched, ~88-million-token corpus composed of both conspiracy (N = 23,937) and non-conspiracy (N = 72,806) documents gathered from 150 websites. For each document, compound extraction was done in a semi-automatic way from the parsed text (see Supplemental Material, SM). We extrapolated nominal compounds whose maximal length was fixed to five constituents. Our final dataset includes 1,713,568 compounds, of which 436,376 are unique.

3.2 Measures

Creative individuals generate high number of ideas. In creativity research, fluency measures quantity of ideas generated. Here, we compute fluency as the rate of compounds for each text, normalized by word count so to account for texts' lengths.

Flexibility relates to the number of unique ideas generated. We compute flexibility via two indexes: lexical flexibility and topical flexibility. The first one relies on lexical diversity, that is, the type-token ratio (TTR), which is associated with verbal creativity (Fradis et al., 1992). We compute TTR by dividing the number of unique compounds by the total number of compounds in each text: higher values indicate higher compound variability. The second index is a count of the semantic dimensions in which ideas are generated (Acar & Runco, 2015). We created this index via topic modelling (Blei et al., 2003), counting the number of different topics needed to explain each document's semantic content. A similar approach was previously used to compute topic specificity, showing that conspiracy narratives, vs non-conspiracy, are composed of a higher number of topics (Miani et al., 2022), hence higher flexibility. Here, we test flexibility relying on topics extracted from compounds, instead of words. Higher values indicate higher topical variability, hence flexibility.

Originality is a crucial component of creativity (Runco, 2010; Runco & Jaeger, 2012; Thys et al., 2014). We operationalize compound originality via two measures: rarity and novelty. We surveyed the Google Ngram corpus (Michel et al., 2011) from 1800 to 2019: for each compound, we extracted statistical frequency and tracked the first usage so to extract the relative age of the compound for each document. Frequency and age of compounds were then sign reversed so to obtain measures rarity and novelty. High scores indicate high originality.

Measuring divergence of participants' responses is one of the most-often used algorithms in creativity research (Beaty & Johnson, 2021; Olson et al., 2021): the higher the conceptual distance between participants' responses (e.g., *doctor* and *bread* vs *doctor* and *nurse*), the higher the creativity. We operationalized divergence via two indices extracted from compounds' constituents: semantic distance and metaphoricity. Semantic distance is based on semantic models (GloVe, Pennington et al., 2014) that correlate with other behavioral measures of creativity (Beaty & Johnson, 2021; Olson et al., 2021). We trained our model on the Wikipedia dump of December 2021, so as to include recent compounds (e.g., about COVID-19). Note that semantic distance is based on frequency of word-pair co-occurrences. Therefore, we developed another measure, metaphoricity, which is unrelated to frequency. Metaphoricity is based on metaphorical relations between constituents, which is an example of real-world creativity (Benczes, 2005; Runco, 2014) often used in creativity research (Primi, 2014; Silvia & Beaty, 2012). Basing our reasoning on the fact that metaphorical mappings are often asymmetrical in terms of concreteness (Casasanto & Boroditsky, 2008), we computed differences in concreteness (Brysbaert et al., 2014) between the head of the compound (i.e., the last word, the semantic root) and other compounds' constituents. The higher the (semantic and metaphorical) divergence between the word pairs, the higher the creativity.

Complexity is another marker of creativity (F. Barron, 1955, 1995; Taylor, 1975). For our purposes, we define complexity as linguistic sophistication. We measure compound sophistication relying on two measures: lexical and structural sophistication. Lexical sophistication was extracted by computing from each compound the mean age-of-acquisition (Kuperman et al., 2012), which is a measure of lexical sophistication (Kim et al., 2018). Structural sophistication was defined as the number of constituents of each compound. Long compounds are persuasive (Bhatia, 1992; Rogers, 2000) and help drawing attention to the message with the help of a creative word formation increasing saliency of the message (Nerlich & Koteyko, 2009).

We create a measure of appropriateness as a proxy for convergent thinking. Appropriateness reflects the degree to which compounds are key to their context (the documents). We did so by computing, for each compound within its document, the term frequency-inverse document frequency, a measure of how relevant a word is to a document in a corpus. The higher the score, the higher the appropriateness of the compound.compound

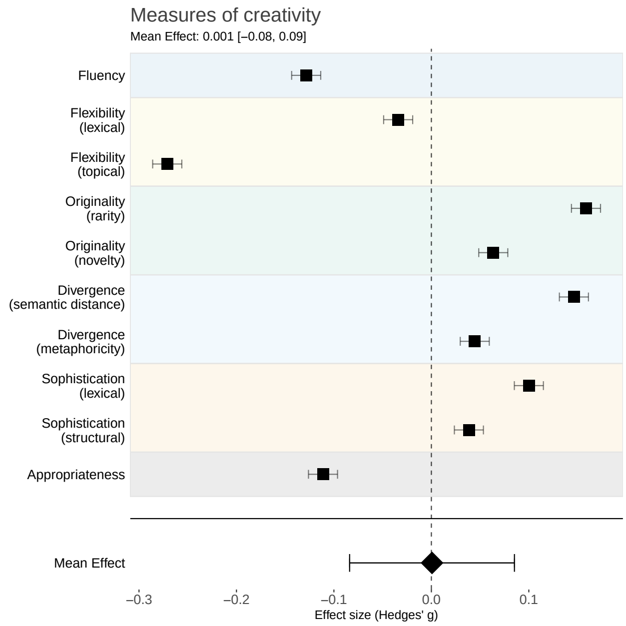
4 Results

To test differences in creativity measures between conspiracy and non-conspiracy texts, we fitted a series of linear mixed-effects models predicting our measures of creativity from conspiracy and non-conspiracy texts and clustering documents within topics. For each model, we report the standardized beta (β) coefficients (all $p < 0.001$, Bonferroni corrected for 10 tests; models' specifications and fits are reported in SM).

Fluency, the rate of compounds per document, was lower in conspiracy texts compared to non-conspiracy texts ($\beta = -.128$). Flexibility was also lower in conspiracy texts, as shown by lower lexical ($\beta = -.034$) and topical ($\beta = -.214$) diversity. This means that conspiracy texts use less compounds, and that the compound vocabulary is also smaller and more limited in semantic range. Low idea generation and low diversity of ideas suggest low creativity in CTs. However, intra- compound analyses show a different pattern. Compounds in conspiracy texts show higher divergence: constituents were more semantically distant ($\beta = .146$) and their relationship was more metaphorical ($\beta = .045$) compared to non-conspiracy texts. Compounds in conspiracy texts were also rarer ($\beta = .163$) and more novel ($\beta = .070$), indicating higher originality. Compounds in conspiracy texts were also more lexically ($\beta = .100$) and structurally ($\beta = .038$) sophisticated than in non-conspiracy texts, suggesting higher complexity. When we analyzed the relationship between compounds and their context, however, we found that conspiracy compounds were less appropriate, i.e., less relevant to the text, compared to mainstream compounds ($\beta = -.111$).

To test whether there was an overall difference in creativity measures between conspiracy and non-conspiracy texts, we converted the regressions' β coefficients into Hedges' g and tested whether the aggregated effect size was higher in CTs texts. Conspiracy and non-conspiracy texts did not differ in creativity ($g = 0.0007$, $SE = 0.043$, $z = 0.0157$, $p = .988$).

Figure 1 — Individual (squares) and aggregated (diamond) effect sizes (Hedges' *g* with 95% CIs) for creativity measures



Note. Positive effects indicate that the measures are higher in the conspiracy subcorpus.

5 Discussion

Our goal was to test whether conspiracy theories are products of creativity (Bonetto & Arciszewski, 2021) or overinclusive thinking. We found no differences between conspiracy and non-conspiracy texts in creativity. However, compounds in conspiracy texts were more original, divergent, and sophisticated, but less appropriate and variable compared to those in non-conspiracy texts. These patterns suggest that CT narratives may emerge from overinclusive thinking. Our results help understand the accumulation of CTs, even contradictory and fictitious ones (Swami et al., 2011; Wood et al., 2012), and the connection of randomly co-occurring events, within one's belief system. Overinclusive thinking might lead to accumulate CTs, which reinforce each other providing evidence for a global conspiracy (Miani et al., 2022; Wood et al., 2012), and potentially increasing persuasive power by making each theory more believable (Pennycook et al., 2018).

Lower fluency and flexibility reflect the use in conspiracy texts of fewer compounds, repeated over texts, revolving around fewer topics. This is in contrast with evidence showing higher topic variability in conspiracy texts (Miani et al., 2022). We interpret this pattern as an effect of reduced vocabulary in conspiracy writers, as suggested by low scores in measures of verbal intelligence (Fiagbenu, 2022) and low education (van Prooijen, 2017) in conspiracy believers. Sophisticated compound constructions might be borrowed from specialistic texts (Oswald, 2016) and/or, as our results from divergence and originality show, generated to elicit surprise. Once a congenial lexical solution is found (a surprisingly complex compound), then it is repeated over across the document (low flexibility), spanning multiple topics (low appropriateness). While this remains a speculative interpretation, future studies could help understand the poietic process in the production of conspiracy texts (e.g., Raab et al., 2013).

Our results extend beyond creativity and shed light on the transmission advantage of CTs and misinformation. People instructed to convince others shift their language spontaneously towards persuasive rhetoric (Rocklage et al., 2018). A shift towards a rhetoric of persuasion might be prompted by the need to persuade that characterizes conspiracy believers and the deliberate transmitters (Marie & Petersen, 2022; Wood & Douglas, 2013). Linguistic complexity, for example, is a persuasive device (Renouf, 2007; Vosoughi et al., 2017) used for flattering the audience (Bruthiaux, 2000) exploiting the readership's need for uniqueness to enhance personal and social identity. Also metaphors, especially when original, are used in persuasive communication (Pogacar et al., 2018). Originality and divergence generate surprise, which attracts attention (Dillard, 2001; Schomaker & Meeter, 2015) and encourages information transmission by conveying social status on people who believe have access to unique information (Berger & Milkman, 2012), tapping into need for uniqueness and sensation-seeking.

We investigated linguistic creativity from texts that are likely produced by specialist writers, therefore our results cannot generalize to a broader context. Future studies could implement our pipeline (freely available at http://www.osf/WORK_ON_THIS) on texts obtained from individuals' speech or social media posts or by devising specific tasks to elicit compound formation. We also acknowledge that other measures could have been used to assess creativity, e.g., lexical innovation such as the internet slang (Del Tredici & Fernández, 2018). We nevertheless believe that compounds are best suited to investigate linguistic creativity in CTs because of their variability in transparency that offers several layers of interpretation (Körtvélyessy et al., 2021). Interpretation is key to conspiracy theorizing (Mayer & Mussweiler, 2011), where opaque semantic relations between ideas may be filled with pre-existing knowledge to a larger extent (Zwaan, 2022).

6 References

Acar, S., & Runco, M. A. (2015). Thinking in multiple directions: Hyperspace categories in divergent thinking. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(1), 41–53. https://doi.org/10.1037/a0038501

Acerbi, A. (2019). Cognitive attraction and online misinformation. *Palgrave Communications*, *5*(1), 15. https://doi.org/10.1057/s41599-019-0224-y

Aitchison, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon*. John Wiley & Sons.

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

Allen, J. (2008, September 10). No consensus on who was behind Sept 11: Global poll. *Reuters*. https://www.reuters.com/article/us-sept11-qaeda-poll-idUSN1035876620080910

Ball, P. (2020). Anti-vaccine movement could undermine efforts to end coronavirus pandemic, researchers warn. *Nature*, *581*(7808), 251–251. https://doi.org/10.1038/d41586-020-01423-4

Bangerter, A., Wagner-Egger, P., & Delouvée, S. (2020). How Conspiracy Theories Spread. In M. Butter & P. Knight (Eds.), *Routledge Handbook of Conspiracy Theories* (1st ed., pp. 206–218). Routledge. https://doi.org/10.4324/9780429452734-2_5

Barron, D., Morgan, K., Towell, T., Altemeyer, B., & Swami, V. (2014). Associations between schizotypy and belief in conspiracist ideation. *Personality and Individual Differences*, *70*, 156–159. https://doi.org/10.1016/j.paid.2014.06.040

Barron, F. (1955). The disposition toward originality. *The Journal of Abnormal and Social Psychology*, *51*(3), 478–485. https://doi.org/10.1037/h0048073

Barron, F. (1993). Controllable Oddness as a Resource in Creativity. *Psychological Inquiry*, *4*(3), 182–184. https://doi.org/10.1207/s15327965pli0403_3

Barron, F. (1995). *No Rootless Flower: An Ecology of Creativity*. Hampton Pr.

Batey, M., & Furnham, A. (2008). The relationship between measures of creativity and schizotypy. *Personality and Individual Differences*, *45*(8), 816–821. https://doi.org/10.1016/j.paid.2008.08.014

Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, *53*(2), 757–780. https://doi.org/10.3758/s13428-020-01453-w

Benczes, R. (2005). Creative noun-noun compounds. *Review of Cognitive Linguistics. Published under the Auspices of the Spanish Cognitive Linguistics Association*, *3*, 250–268.

Berger, J., & Milkman, K. L. (2012). What Makes Online Content Viral? *Journal of Marketing Research*, *49*(2), 192–205. https://doi.org/10.1509/jmr.10.0353

Bhatia, V. K. (1992). Pragmatics of the Use of Nominals in Academic and Professional Genres. *Pragmatics and Language Learning*, *3*, 217–230.

Biddlestone, M., Green, R., Cichocka, A., Douglas, K., & Sutton, R. (2022). *A systematic review and meta-analytic synthesis of the motives associated with conspiracy beliefs*. PsyArXiv. https://doi.org/10.31234/osf.io/rxjqc

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3*(null), 993–1022.

Bonetto, E., & Arciszewski, T. (2021). The Creativity of Conspiracy Theories. *The Journal of Creative Behavior*, *55*(4), 916–924. https://doi.org/10.1002/jocb.497

Bruthiaux, P. (2000). In a nutshell: Persuasion in the spatially constrained language of advertising. *Language & Communication*, *20*(4), 297–310. https://doi.org/10.1016/S0271-5309(00)00003-3

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593. https://doi.org/10.1016/j.cognition.2007.03.004

Cropley, A. (2006). In Praise of Convergent Thinking. *Creativity Research Journal*, *18*(3), 391–404. https://doi.org/10.1207/s15326934crj1803_13

Cropley, D. H., Kaufman, J. C., & Cropley, A. J. (2008). Malevolent Creativity: A Functional Model of Creativity in Terrorism and Crime. *Creativity Research Journal*, *20*(2), 105–115. https://doi.org/10.1080/10400410802059424

Darwin, H., Neave, N., & Holmes, J. (2011). Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Personality and Individual Differences*, *50*(8), 1289–1293. https://doi.org/10.1016/j.paid.2011.02.027

Del Tredici, M., & Fernández, R. (2018). The Road to Success: Assessing the Fate of Linguistic Innovations in Online Communities. *Proceedings of the 27th International Conference on Computational Linguistics*, 1591–1603. https://aclanthology.org/C18-1135

Dhar, P., & van der Plas, L. (2019). Learning to Predict Novel Noun-Noun Compounds. *Proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)*, 30–39. https://doi.org/10.18653/v1/W19-5105

Dillard, J. (2001). Persuasion and the structure of affect. Dual systems and discrete emotions as complementary models. *Human Communication Research*, *27*(1), 38–68. https://doi.org/10.1093/hcr/27.1.38

Dollinger, S. J. (2003). Need for Uniqueness, Need for Cognition, and Creativity. *The Journal of Creative Behavior*, *37*(2), 99–116. https://doi.org/10.1002/j.2162-6057.2003.tb00828.x

Dollinger, S. J., Urban, K. K., & James, T. A. (2004). Creativity and Openness: Further Validation of Two Creative Product Measures. *Creativity Research Journal*, *16*(1), 35–47. https://doi.org/10.1207/s15326934crj1601_4

Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding Conspiracy Theories. *Political Psychology*, *40*(S1), 3–35. https://doi.org/10.1111/pops.12568

Dumas, D., Organisciak, P., & Doherty, M. (2021). Measuring divergent thinking originality with human raters and text-mining models: A psychometric comparison of methods. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(4), 645–663. https://doi.org/10.1037/aca0000319

Elvevåg, B., Foltz, P. W., Weinberger, D. R., & Goldberg, T. E. (2007). Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophrenia Research*, *93*(1–3), 304–316. https://doi.org/10.1016/j.schres.2007.03.001

Eysenck, H. J. (2003). Creativity, personality and the convergent-divergent continuum. In *Critical creative processes* (pp. 95–114). Hampton Press.

Fiagbenu, M. E. (2022). The stock market is rigged? Conspiracy beliefs and distrust predict lower stock market participation. *Applied Cognitive Psychology*, acp.3981. https://doi.org/10.1002/acp.3981

Fradis, A., Mihailescu, L., & Jipescu, I. (1992). The distribution of major grammatical classes in the vocabulary of Romanian aphasic patients. *Aphasiology*, *6*(5), 477–489. https://doi.org/10.1080/02687039208249485

Franks, B., Bangerter, A., & Bauer, M. W. (2013). Conspiracy theories as quasi-religious mentality: An integrated account from cognitive science, social representations theory, and frame theory. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00424

Gagné, C. L. (2002). Lexical and Relational Influences on the Processing of Novel Compounds. *Brain and Language*, *81*(1–3), 723–735. https://doi.org/10.1006/brln.2001.2559

Guilford, J. P. (1950). *Fundamental statistics in psychology and education, 2nd ed* (pp. xiii, 633). McGraw-Hill.

Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.

Hills, T. T. (2019). The Dark Side of Information Proliferation. *Perspectives on Psychological Science*, *14*(3), 323–330. https://doi.org/10.1177/1745691618803647

Imhoff, R., Dieterle, L., & Lamberty, P. (2021). Resolving the Puzzle of Conspiracy Worldview and Political Activism: Belief in Secret Plots Decreases Normative but Increases Nonnormative Political Engagement. *Social Psychological and Personality Science*, *12*(1), 71–79. https://doi.org/10.1177/1948550619896491

Jolley, D., & Douglas, K. M. (2014). The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one’s carbon footprint. *British Journal of Psychology*, *105*(1), 35–56. https://doi.org/10.1111/bjop.12018

Jolley, D., & Paterson, J. L. (2020). Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *British Journal of Social Psychology*, *59*(3), 628–640. https://doi.org/10.1111/bjso.12394

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00407

Kim, M., Crossley, S. A., & Kyle, K. (2018). Lexical Sophistication as a Multidimensional Phenomenon: Relations to Second Language Lexical Proficiency, Development, and Writing Quality. *The Modern Language Journal*, *102*(1), 120–141. https://doi.org/10.1111/modl.12447

Körtvélyessy, L., Štekauer, P., & Kačmár, P. (2021). On the role of creativity in the formation of new complex words. *Linguistics*, *59*(4), 1017–1055. https://doi.org/10.1515/ling-2020-0003

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lobato, E., Mendoza, J., Sims, V., & Chin, M. (2014). Examining the Relationship Between Conspiracy Theories, Paranormal Beliefs, and Pseudoscience Acceptance Among a University Population: Relationship between unwarranted beliefs. *Applied Cognitive Psychology*, *28*(5), 617–625. https://doi.org/10.1002/acp.3042

Marie, A., & Petersen, M. B. (2022). Political conspiracy theories as tools for mobilization and signaling. *Current Opinion in Psychology*, 101440. https://doi.org/10.1016/j.copsyc.2022.101440

Mayer, J., & Mussweiler, T. (2011). Suspicious spirits, flexible minds: When distrust enhances creativity. *Journal of Personality and Social Psychology*, *101*(6), 1262–1277. https://doi.org/10.1037/a0024407

Miani, A., Hills, T., & Bangerter, A. (2021). LOCO: The 88-million-word language of conspiracy corpus. *Behavior Research Methods*, *54*(4), 1794–1817. https://doi.org/10.3758/s13428-021-01698-z

Miani, A., Hills, T., & Bangerter, A. (2022). Interconnectedness and (in)coherence as a signature of conspiracy worldviews. *Science Advances*, *8*(43). https://doi.org/10.1126/sciadv.abq3668

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, *331*(6014), 176–182. https://doi.org/10.1126/science.1199644

Nerlich, B., & Koteyko, N. (2009). Compounds, creativity and complexity in climate change communication: The case of ‘carbon indulgences.’ *Global Environmental Change*, *19*(3), 345–353. https://doi.org/10.1016/j.gloenvcha.2009.03.001

Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and Mystery: The Cultural Selection of Minimally Counterintuitive Narratives. *Cognitive Science*, *30*(3), 531–553. https://doi.org/10.1207/s15516709cog0000_68

Oliver, J. E., & Wood, T. (2014). Medical Conspiracy Theories and Health Behaviors in the United States. *JAMA Internal Medicine*, *174*(5), 817. https://doi.org/10.1001/jamainternmed.2014.190

Olson, J. A., Nahas, J., Chmoulevitch, D., Cropper, S. J., & Webb, M. E. (2021). Naming unrelated words predicts creativity. *Proceedings of the National Academy of Sciences*, *118*(25). https://doi.org/10.1073/pnas.2022340118

Oswald, S. (2016). Conspiracy and bias: Argumentative features and persuasiveness of conspiracy theories. *OSSA Conference Archive*, *168*, 1–16.

Paulsen, J. S., Romero, R., Chan, A., Davis, A. V., Heaton, R. K., & Jeste, D. V. (1996). Impairment of the semantic network in schizophrenia. *Psychiatry Research*, *63*(2–3), 109–121. https://doi.org/10.1016/0165-1781(96)02901-0

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. https://doi.org/10.3115/v1/D14-1162

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865–1880. https://doi.org/10.1037/xge0000465

Pogacar, R., Shrum, L. J., & Lowrey, T. M. (2018). The Effects of Linguistic Devices on Consumer Information Processing and Persuasion: A Language Complexity × Processing Mode Framework. *Journal of Consumer Psychology*, *28*(4), 689–711. https://doi.org/10.1002/jcpy.1052

Primi, R. (2014). Divergent productions of metaphors: Combining many-facet Rasch measurement and cognitive psychology in the assessment of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(4), 461–474. https://doi.org/10.1037/a0038055

Raab, M. H., Auer, N., Ortlieb, S. A., & Carbon, C.-C. (2013). The Sarrazin effect: The presence of absurd statements in conspiracy theories makes canonical information less plausible. *Frontiers in Psychology*, *4*(JUL), 1–8. https://doi.org/10.3389/fpsyg.2013.00453

Renouf, A. (2007). Tracing lexical productivity and creativity in the British media: The Chavs and the Chav-Nots. In J. Munat (Ed.), *Studies in Functional and Structural*

Linguistics (Vol. 58, pp. 61–89). John Benjamins Publishing Company. https://doi.org/10.1075/sfsl.58.12ren

Rocklage, M. D., Rucker, D. D., & Nordgren, L. F. (2018). Persuasion, Emotion, and Language: The Intent to Persuade Transforms Language via Emotionality. *Psychological Science*, *29*(5), 749–760. https://doi.org/10.1177/0956797617744797

Rogers, M. (2000). Genre and Terminology. In A. Trosborg (Ed.), *Pragmatics & Beyond New Series* (Vol. 74, p. 3). John Benjamins Publishing Company. https://doi.org/10.1075/pbns.74.03rog

Runco, M. A. (2010). Divergent Thinking, Creativity, and Ideation. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (1st ed., pp. 413–446). Cambridge University Press. https://doi.org/10.1017/CBO9780511763205.026

Runco, M. A. (2014). *Creativity: Theories and themes: Research, development, and practice, 2nd ed* (pp. xiii, 500). Elsevier Academic Press.

Runco, M. A., & Jaeger, G. J. (2012). The Standard Definition of Creativity. *Creativity Research Journal*, *24*(1), 92–96. https://doi.org/10.1080/10400419.2012.650092

Schomaker, J., & Meeter, M. (2015). Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neuroscience & Biobehavioral Reviews*, *55*, 268–279. https://doi.org/10.1016/j.neubiorev.2015.05.002

Silvia, P. J., & Beaty, R. E. (2012). Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence*, *40*(4), 343–351. https://doi.org/10.1016/j.intell.2012.02.005

Silvia, P. J., Wigert, B., Reiter-Palmon, R., & Kaufman, J. C. (2012). Assessing creativity with self-report scales: A review and empirical evaluation. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 19–34. https://doi.org/10.1037/a0024071

Simonton, D. K. (1999). Creativity as blind variation and selective retention: Is the creative process Darwinian? *Psychological Inquiry*, *10*(4), 309–328. https://doi.org/10.1207/S15327965PLI1004_4

Skinner, N. F. (1996). Behavioral implications of adaption-innovation: II. Adaption-innovation and motivation for uniqueness. *Social Behavior and Personality: An International Journal*, *24*(3), 231–234. https://doi.org/10.2224/sbp.1996.24.3.231

Snyder, C. R., & Fromkin, H. L. (1977). Abnormality as a positive characteristic: The development and validation of a scale measuring need for uniqueness. *Journal of Abnormal Psychology*, *86*(5), 518–527. https://doi.org/10.1037/0021-843X.86.5.518

Sternisko, A., Cichocka, A., & Bavel, J. J. V. (2020). The dark side of social movements: Social identity, non-conformity, and the lure of conspiracy theories. *Current Opinion in Psychology*, *35*, 1–6. https://doi.org/10.1016/j.copsyc.2020.02.007

Swami, V., Coles, R., Stieger, S., Pietschnig, J., Furnham, A., Rehim, S., & Voracek, M. (2011). Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories: Conspiracist ideation. *British Journal of Psychology*, *102*(3), 443–463. https://doi.org/10.1111/j.2044-8295.2010.02004.x

Taylor, I. A. (Ed.). (1975). An Emerging View of Creative Actions. In *Perspectives in Creativity*. Routledge.

Thys, E., Sabbe, B., & Hert, M. D. (2014). The assessment of creativity in creativity/psychopathology research – a systematic review. *Cognitive Neuropsychiatry*, *19*(4), 359–377. https://doi.org/10.1080/13546805.2013.877384

Tian, K. T., Bearden, W. O., & Hunter, G. L. (2001). Consumers' Need for Uniqueness: Scale Development and Validation. *Journal of Consumer Research*, *28*(1), 50–66. https://doi.org/10.1086/321947 van der Wal, R. C., Sutton, R. M., Lange, J., & Braga, J. P. N. (2018). Suspicious binds: Conspiracy thinking and tenuous perceptions of causal connections between co-occurring and spuriously correlated events. *European Journal of Social Psychology*, *48*(7), 970–989. https://doi.org/10.1002/ejsp.2507

van Prooijen, J.-W. (2017). Why Education Predicts Decreased Belief in Conspiracy Theories: Education and Conspiracy Beliefs. *Applied Cognitive Psychology*, *31*(1), 50–58. https://doi.org/10.1002/acp.3301

van Prooijen, J.-W. (2018). *The Psychology of Conspiracy Theories*. Routledge.

van Prooijen, J.-W., Ligthart, J., Rosema, S., & Xu, Y. (2022). The entertainment value of conspiracy theories. *British Journal of Psychology*, *113*(1), 25–48. https://doi.org/10.1111/bjop.12522

Vosoughi, S., Mohsenvand, M. 'Neo,' & Roy, D. (2017). Rumor Gauge: Predicting the Veracity of Rumors on Twitter. *ACM Transactions on Knowledge Discovery from Data*, *11*(4), 1–36. https://doi.org/10.1145/3070644

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

Wang, L., Long, H., Plucker, J. A., Wang, Q., Xu, X., & Pang, W. (2018). High Schizotypal Individuals Are More Creative? The Mediation Roles of Overinclusive Thinking and Cognitive Inhibition. *Frontiers in Psychology*, *9*, 1766. https://doi.org/10.3389/fpsyg.2018.01766

Willits, J. A., Rubin, T., Jones, M. N., Minor, K. S., & Lysaker, P. H. (2018). Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. *Schizophrenia Research*, *197*, 365–369. https://doi.org/10.1016/j.schres.2017.11.014

Wood, M. J., & Douglas, K. M. (2013). What about building 7?" A social psychological study of online discussion of 9/11 conspiracy theories. *Frontiers in Psychology*, *4*(JUL), 1–9. https://doi.org/10.3389/fpsyg.2013.00409

Wood, M. J., Douglas, K. M., & Sutton, R. M. (2012). Dead and Alive: Beliefs in Contradictory Conspiracy Theories. *Social Psychological and Personality Science*, *3*(6), 767–773. https://doi.org/10.1177/1948550611434786

Zwaan, R. A. (2022). Conspiracy thinking as situation model construction. *Current Opinion in Psychology*, *47*, 101413. https://doi.org/10.1016/j.copsyc.2022.101413

1 Method

1.1 Compounds extraction

Before extracting compounds, we parsed texts via the R package *spacyr* (Benoit & Matsuo, 2020), a wrapper to the *spaCy* library in Python (Montani et al., 2021). Note that we did not perform any text pre-processing prior to parsing the text (e.g., removing stopwords or stemming) because *spacyr* needs a syntactically valid text in order to apply the parser. We used the trained model `en_core_web_trf` (438 Mb) because it provided the highest accuracy (i.e., ACC = .98 vs ACC = .97 of other models) in extracting the part-of-speech (POS) tags and labelling dependencies (ACC = .94-.95 vs ACC = .90-.92 of other models).[1] SpaCy returns the parsed output of each document as a data frame, where each row corresponds to a token from the document. The data frame contains the fully parsed syntactical tree with sequential numbers for each word and their dependencies along with universal dependency POS tagset (e.g., NOUN, VERB, ADJ [adjective]) and the detailed part-of-speech tags (e.g., VBZ: verb, 3rd person sing. present; NNS: noun, plural). Crucial for our task, spaCy provides the universal dependency relations (de Marneffe et al., 2014),[2] from which we extracted compounds.

Compound extraction was done by selecting tokens where the universal dependency relation was defined as compound (compound's modifier) and adding to the selection the subsequent tokens for the whole length of the compound (so to include also the compound's head), including hyphens (See Table 1). Once all compounds were extracted, we started cleaning the list of compounds: 1) we removed numeric tokens; 2) we kept only compounds where the tag of the components was either a noun or a hyphen; 3) we removed compounds starting or ending with an hyphen; 4) we removed non-compound single words; 5) we removed compounds whose length was higher than five constituents —hyphen excluded; and 6) we kept only tokens that were not parsed as entity (e.g., proper nouns, organizations, and locations).

We checked whether there were spelling mistakes among the most infrequent compounds. We started by creating a word-frequency table for all tokens in LOCO and

---

[1] see https://spacy.io/models/en

[2] see: https://universaldependencies.org/u/dep/index.html

extracted the frequency associated with the most infrequent constituent. For example, if the tokens *microcode* and *climate* occur respectively 1 and 50,859 times in LOCO, then the compound *climate microcode* will have a value of 1. In doing so, we searched for a threshold of frequency below which we could exclude compounds based on infrequent constituent, possibly spelling errors. We therefore started visually exploring compounds whose at least one of the constituents had a frequency of 1 (e.g., *climate microcode*) but found no evidence for spelling mistakes. Instead, these infrequent compounds were mostly technical terms (e.g., *orrery model*) or infrequent Latin or other languages phrases (e.g., *manu militari*, *concert etude*), while misspellings were extremely rare (e.g., *socio-eco-nomic revival*, occurring twice). Also, the number of infrequent components was unimportant: there were 90 compounds whose constituents had a min frequency of 1, there were 113 compounds whose minimum frequency of constituents was 2, and so on with similar figures. Up to a minimum frequency of 10, there was a total of 1,917 unique compounds (N = 2,251 in total), accounting for 0.44% of all unique compounds (and 0.13% of all compounds). We therefore decided not to remove any compound based on frequency threshold.

Our final dataset includes 1,713,568 compounds, whose 436,376 are unique compounds. Two-word compounds account for 89.30% of all compounds (N = 1,530,299), followed by three- (N = 170,354), four- (N = 12,101), and five-word (N = 814) compounds. Compounds' length was limited to five constituents because this is the limit of the Google Ngram Corpus (Michel et al., 2011) so to exclude the artifact that longer compounds had lower originality scores (i.e., not present in Google) due to their length. Moreover, before cleaning the dataset, longer compounds accounted for less than 0.5% of the whole dataset and visual inspection suggested that most were misclassified compounds (e.g., a 40-word compound). The most popular compounds for the whole corpus as well as for conspiracy and non-conspiracy subcorpora are listed in Table 2.

Table 1 ~ Example of parsed texts by spaCy

sentence_id	token_id	token	lemma	pos	tag	head_token_id	dep_rel
1	1	Exercise	exercise	NOUN	NN	2	nsubj
1	2	lowers	lower	VERB	VBZ	2	ROOT
1	**3**	**blood**	**blood**	**NOUN**	**NN**	**4**	**compound**
1	**4**	**sugar**	**sugar**	**NOUN**	**NN**	**5**	**compound**
1	**5**	**levels**	**level**	**NOUN**	**NNS**	**2**	**dobj**
1	6	without	without	ADP	IN	2	prep
1	7	medication	medication	NOUN	NN	6	pobj
1	8	.	.	PUNCT	.	2	punct

2	1	During	during	ADP	IN	11	prep
2	2	the	the	DET	DT	6	det
2	**3**	**dinner**	**dinner**	**NOUN**	**NN**	**5**	**compound**
2	**4**	**-**	**-**	**PUNCT**	**HYPH**	**5**	**punct**
2	**5**	**dance**	**dance**	**NOUN**	**NN**	**6**	**compound**
2	**6**	**cruise**	**cruise**	**NOUN**	**NN**	**1**	**pobj**
2	7	with	with	ADP	IN	6	prep
2	8	orchestra	orchestra	NOUN	NN	7	pobj
2	9	,	,	PUNCT	,	11	punct
2	10	John	John	PROPN	NNP	11	nsubj
2	11	played	play	VERB	VBD	11	ROOT
2	12	guitar	guitar	NOUN	NN	11	dobj
2	13	.	.	PUNCT	.	11	punct

Note. Bolded words represent compounds in each sentence: “Exercise lowers blood sugar levels without medication.” (compound: *blood sugar levels*), and “During the dinner-dance cruise with orchestra, John played guitar.” (compound: *dinner-dance cruise*).

Table 2 ~ Most popular compounds for the whole dataset as well as the conspiracy and non-conspiracy datasets

ALL		**Conspiracy**		**Mainstream**	
compound	*Freq*	*compound*	*Freq*	*compound*	*Freq*
climate change	17,823	climate change	2,973	climate change	14,850
conspiracy theories	5,169	mainstream media	1,583	conspiracy theories	4,137
side effects	4,688	mind control	1,474	side effects	3,473
drug companies	3,821	world order	1,409	health care	2,835
health care	3,585	pharmaceutical industry	1,302	drug companies	2,576
pharmaceutical industry	3,465	drug companies	1,245	health officials	2,374
health officials	2,975	side effects	1,215	pharmaceutical industry	2,163
world order	2,896	conspiracy theories	1,032	breast cancer	2,155
conspiracy theory	2,801	conspiracy theory	977	vice president	2,076
breast cancer	2,656	world government	956	carbon dioxide	2,003
carbon dioxide	2,574	drinking water	798	family members	1,909
vice president	2,506	intelligence agencies	771	conspiracy theory	1,824
family members	2,471	cancer cells	769	greenhouse gases	1,615
cancer cells	2,225	health care	750	blood pressure	1,557
drinking water	2,103	government officials	741	drug prices	1,552
conspiracy theorists	2,036	intelligence services	740	death toll	1,503
mind control	1,994	law enforcement	712	world order	1,487
blood pressure	1,993	police state	711	greenhouse gas emissions	1,473
press conference	1,915	mainstream news	672	cancer cells	1,456
death toll	1,906	population control	619	coronavirus cases	1,439

1.2 Measures

1.2.1 Fluency

We compute fluency as the number of compounds in each text divided by word count.

1.2.2 Flexibility: lexical flexibility

Lexical flexibility was operationalized via lexical diversity, that is, the type-token ratio (TTR), that was previously associated with verbal creativity (Fradis et al., 1992). For each document, we compute TTR by dividing the number of unique compounds by the total number of compounds in each text: higher values indicate higher compound variability.

1.2.3 Flexibility: topical flexibility

Topical flexibility is the count of semantic dimensions in which ideas are generated (Acar & Runco, 2015). We created the topical flexibility index using the Latent Dirichlet Allocation, LDA (Blei et al., 2003). LDA is an algorithm that allows to identify co-occurring word patterns aiming at extracting the underlying topic distribution for each document in a corpus. By setting *a priori* a value for number of topics desired, LDA computes, for each document in the corpus, the probabilities for all topics to be represented in the document. The algorithm returns a matrix with topic-in-document probabilities where a topic k has probability γ of being part of document n. The sum of all the topic probabilities within one document is 1.

We created a corpus composed of compounds, meaning that features were compounds instead of words (e.g., *climate_change* instead of *climate* and *change*). From the whole corpus, we extracted 50 topics using the *seededlda* R package (Watanabe & Xuan-Hieu, 2022). Extracting 50 topics means that each document has 50 gamma values that sum to 1. In order to compute topical flexibility, i.e., the number of different semantic dimensions for each document, we counted the number of different topics needed to explain each document. Specifically, because all topics sum to 1, being each topic a probability of representing document’s semantic content, we counted the number of topics needed to explain 50% of the document. We relied on the following pipeline. For each document, 1) we extracted the gamma values for each topic (N = 50) and ordered in descending way; 2) from the gamma values vector, we computed the cumulative sum (i.e., the first values represents the highest gamma value and the last value is equal to 1); 3) we counted how many gamma values (i.e., topics) were needed to reach a value of .50 by selecting from the vector only values below .50 and counting the length of the vector (i.e., number of values below .50 extracted).

Note that we have also tried other settings: we extracted also 100 topics and explored other thresholds such as .25 and .75 so to see if results change with other parameters. We therefore extracted six different topical flexibility measures by crossing 2 topics k = {50, 100} x 3 thresholds t = {.25, .50, .75}. These six measures correlate to each other with a mean of r = .84 (range: .58 - .98). To test whether the different measures impacted the difference between conspiracy and non-conspiracy documents, we ran a linear regression for each measure. The six regressions yielded coefficients between β = -0.261 and β = -0.225, meaning that results were not substantially different.

1.2.4 Originality: rarity and novelty

To extract measures of originality, that is rarity and novelty, we checked whether our compounds were present in the Google Ngram corpus (Michel et al., 2011) between the year 1800 and 2019. Data was retrieved (between October and December 2021) via the R package *ngramr* (Carmody, 2022). We used the corpus of English books (Google identifier: `googlebooks-eng-20200217`) published in any country until 2019. We have chosen this corpus because it is the most recent one (as of 2021, when we started data collection) and it includes books written in both British and US English.[3] The search was set as case insensitive. We used the non-lemmatized form including hyphens so to include different forms of the same lemma. In doing so, *pharmaceuticals_industry*, *pharmaceutical_industry*, *pharmaceutical_industries*, and *pharmaceutical_-_industry* returned different values, providing more granular results.

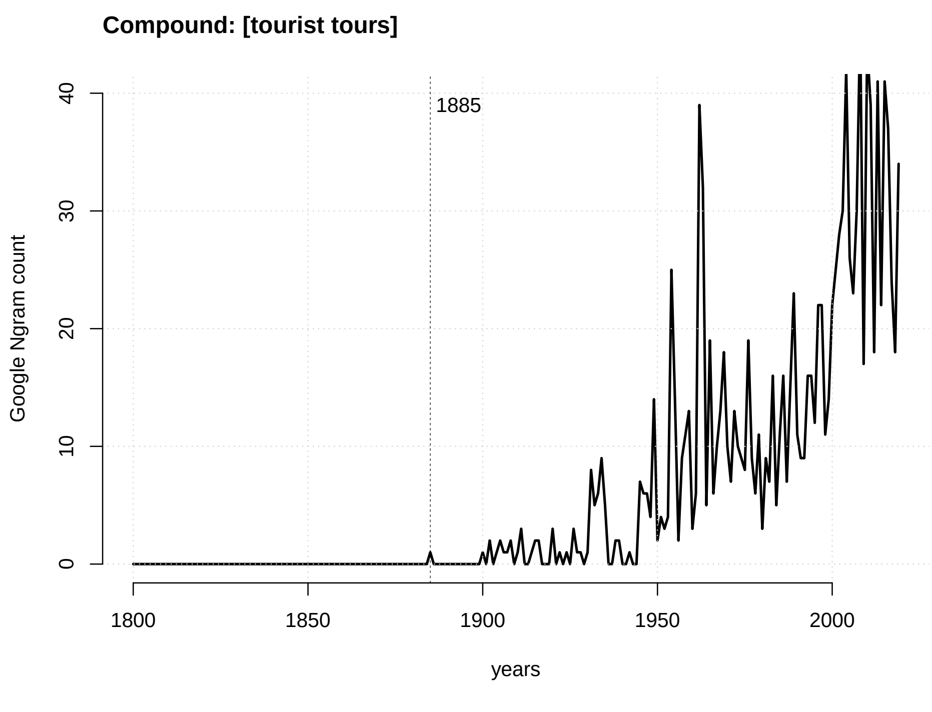
Our measure of rarity was obtained by summing up all occurrences of each compound. For example, the compound *world war*, with 52,581,927 instances, is more popular than *century phobia* (N = 44), which, in turn, is more popular than *mkultra agenda* (N = 0), since the latter does not appear in the Google corpus. Instances of each compound were then log transformed (to approach a normal distribution) and sign-reverted so to obtain a measure of rarity where lower values are associated with more popular compounds and higher values associated with rarity. The highest possible score for rarity is zero (i.e., not present in Google), followed by negative values where lower values indicate high popularity.

[3] see: https://books.google.com/ngrams/info

Our measure of novelty was created by extracting the age of each compound relative to the document in use. To do so, we extracted the first year in which each compound appeared in the Google corpus by extracting the first non-zero value in the time series of compound frequency between 1800 and 2019. For example, the compound *tourist tours* appeared on the Google corpus for the first time in 1885 (see Figure 1) while the compound *torture chamber* appeared in 1800. Then, the absolute age of compound was subtracted from the age of the document in which the compound appeared. For example, the compound *autism occurrence* appeared for the first time in the Google corpus in 2015. When this compound was included in a document dated 2016, then the relative compound age was 1; when it was contained in a document dated 2002, then its relative age was -13 years. We reasoned that the younger the relative age, the more creative the compound is.

Compounds measures of rarity and novelty were averaged so to obtain values of originality and novelty for each document.

Figure 1 — Google Ngram count for the compound [tourist tour] from 1800 to 2019



1.2.5 Divergence: semantic distance

We rely on a measure of semantic distance that has been previously shown to correlate with other behavioral measures of creativity (Olson et al., 2021): the higher the distance between the pair words, the more creativity associated with the compound. Two words are semantically related if they often co-occur together and share the same context, while they are unrelated when each word is used in different contexts.

In order to compute semantic distance, we first needed to create a semantic space, which is a word-to-word matrix of semantic relationships. We did so via an algorithm called GloVe (Pennington et al., 2014), an unsupervised learning algorithm that extracts vector representations for words within a corpus. We trained our semantic space using the latest (when data collection occurred) Wikipedia dump of the 20th of December 2021 (about 19 GB of text)[4] so to include words related to recent events that were not included in previous versions of GloVe (2014) but were included in LOCO (2020). We converted the xml files and extracted the useful text with the Wikipedia dump file to text converter, `wp2txt`, a Ruby/Unix script that works similar to a boilerplate splitter that returns a human readable text (Hasebe, 2006).[5] Once texts were extracted, we removed titles and lists so to obtain cleaned chunks of texts. From these texts, because the file was too large to be handled by R (maximum length of vector allowed is 2^31), we built a vocabulary from words contained in LOCO (N = 53,569 cleaned by stopwords [stopwords obtained by (Benoit et al., 2021)] and single letters of the alphabet) accounting for a total of 1.7B tokens. Note that a large vocabulary leads to intense use of computational power with no particular gain (Pereira et al., 2016). We then tokenized (removing punctuation, URLs, numbers, symbols, and splitting hyphens), and stemmed each word in the character vector using the R package *quanteda* (Benoit et al., 2018). Stemming was done so to further reduce vocabulary by aggregating together words with similar roots (e.g., “frequenc” for *frequency* and *frequencies*). This removed noise from different words with similar meaning (we do not believe that, for example, plural nouns are different from singular nouns to our purposes).

To create the GloVe semantic space from the Wikipedia dump, we relied on the R package *txt2vec* (Selivanov et al., 2022). The output, i.e., the trained semantic space, is a matrix of N (words) vectors per D (dimensions, in our case set to 50). Semantic distance was then computed from the cosine distance between vectors, i.e., constituents of compounds (see Figure 2A). The Python script to compute semantic distance was obtained from (Olson et al., 2021).[6] Semantic distance theoretically ranges from 0 to 200; in our sample it ranged from 2.86 to 163.35 (mean = 62.34). For example, in Wikipedia, the constituents of the compound

[4] see: https://dumps.wikimedia.org/enwiki/20220101/

[5] see: https://github.com/yohasebe/wp2txt

[6] see: https://github.com/jayolson/divergent-association-task

coca cola (6.30) co-occur together and share context more often than the constituents of the compound *aspartame death* (142.01). Prior to entering the function, words were stemmed. Therefore, *pharmaceutical industry* and *pharmaceutical industries* have the same score (37.44). Note that order did not affect the scores, therefore, semantic distance scores for *island reunion* and *reunion island* are the same (146.94). Note that we were capable of extracting semantic information also for recent events such as COVID-19 (e.g., *coronavirus deceit*, 121.28; *coronavirus pandemic*, 18.04).

1.2.6 Divergence: metaphoricity

We developed metaphoricity basing our reasoning on the fact that metaphorical mappings are often asymmetrical in terms of concreteness (Casasanto & Boroditsky, 2008; Lakoff & Johnson, 2008): for example, in the *time flow* compound, an abstract entity (*time*) is conceptualized via a more concrete domain (*flow*). It follows that differences in concreteness between the head of the compound and its constituents should be a valuable indicator for the metaphoricity of a compound, hence its creativity.

Concreteness norms were collected by (Brysbaert et al., 2014) from four thousands Mturk workers who rated about 40,000 words on their concreteness degree. Concreteness was defined as the "*degree to which the concept denoted by a word refers to a perceptible entity*" (p. 904). For example, while *essentialness* shows low degree of concreteness (1.04), *apple* shows a higher degree of concreteness (5.00). Then, the metaphoricity measure was obtained by computing the absolute difference in concreteness scores between the head of the compound (the last word) and the average scores of concreteness of the other constituents (see Figure 2B). For example, in the compound *computer radio frequency propagation program*, from the concreteness value of *program* (3.43), we subtracted the mean concreteness scores of *computer* (4.93) *radio* (4.74) *frequency* (2.69) *propagation* (2.08), obtaining a total metaphoricity score of |-0.18|. Note that our measure of metaphoricity is not a measure of concreteness. In fact, both *denial phase* and *soy cheese*, respectively abstract and concrete compounds, score 0.00, while *information highway*, *earth belief* and *belief engine* score respectively 1.85, 3.61, and 3.67, meaning that the head and constituents of the compounds are asymmetric in terms of concreteness.

We validated our measure of metaphoricity, by testing whether differences in concreteness between head and constituents would be higher in metaphors. From previous research (Arzouan et al., 2007; Gold et al., 2010), we obtained pairs of words labelled as either conventional metaphors (N = 14, e.g., *juicy gossip*) or literal expressions (N = 15, e.g., *soft blanket*) and tested whether metaphors score higher than literal expressions in our measure of metaphoricity. Conventional metaphors scored higher in metaphoricity than literal expressions, $t_{24.71} = 2.338$, $p = .0278$, Cohen's $d = 0.87$ (conventional metaphors: $M = 1.263$, $SD = 0.724$, range: 0 - 2.35; literal expression: $M = 0.644$, $SD = 0.668$, range: 0 - 2.55, see Figure 3).

Figure 2 — Computation of semantic distance (A) and metaphoricity (B)

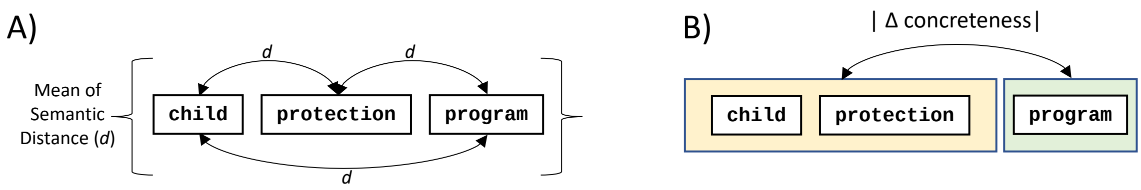
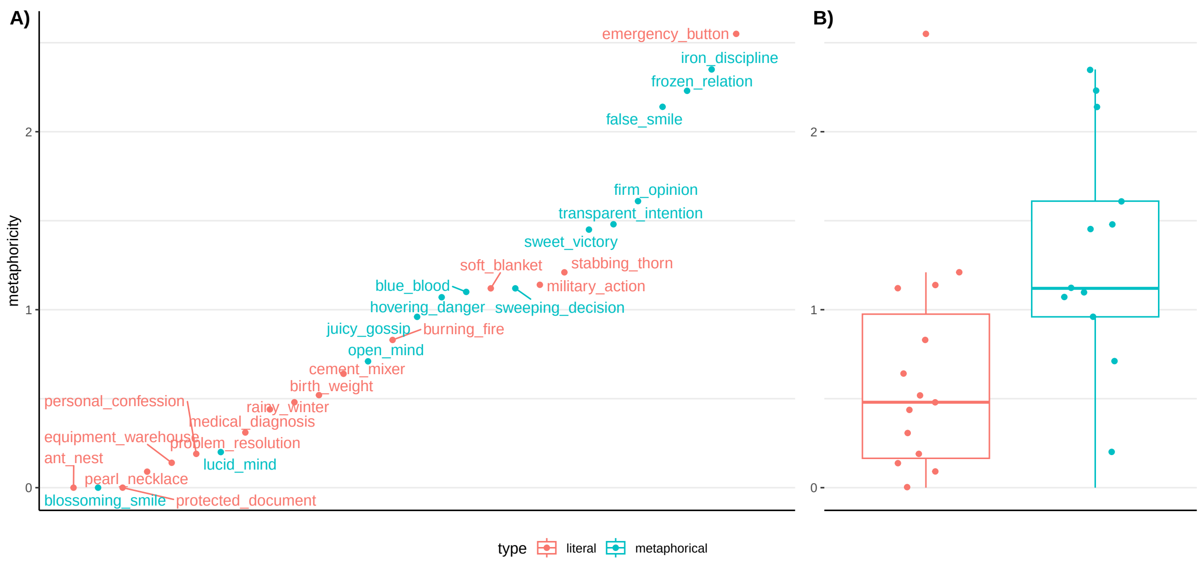


Figure 3 — Testing metaphoricity measure



1.2.7 Complexity: lexical sophistication

Compound sophistication was measured using the norms of age of acquisition (AoA), obtained from (Kuperman et al., 2012) for ~30,000 English words. AoA norms are based on human judgments of the age at which a particular word was learned. They predict lexical sophistication in second language acquisition (Kim et al., 2018; Kyle & Crossley, 2015). For each compound, we averaged AoA values so to obtain a single score of sophistication per compound. For instance, the compound *mama dada* has an average score of 1.89, meaning that the two constituents are learned on average at an age of about 2 years, while the constituents of the compound *neurotransmitter serotonin* are learned at an average of 16.49 years of age.

1.2.8 Complexity: structural sophistication

Structural sophistication was computed as the number of constituents for each compound (e.g., *mama dada*: N = 2; *child protection program*: N = 3; *dinner-dance cruise*: N = 3).

1.2.9 Appropriateness

A measure of appropriateness was created by assessing the degree to which compounds are key to their document. We relied on the term frequency-inverse document frequency (TF-IDF), a technique that assesses the relevance of a word to a document in a corpus. For each word in a document, TF-IDF is computed by counting how many times a word appears in a document divided by the inverse document frequency of the word in the corpus. In this way, high- (e.g., stopwords) and low-frequency (e.g., rare) words are offset.

TF-IDF was computed using the function `dfm_tfidf` from the R package *quanteda* (Benoit et al., 2018). The function takes as input the document-term matrix (DTM) obtained from a corpus. DTM is a bag of words for the whole corpus that stores the occurrences of each word (columns) for each document (rows). The function returns a weighted DTM, where each cell contains the TF-IDF value for a word in a document. Before computing the DTM, we preprocessed the corpus with the following pipeline: 1) lower casing; 2) removing punctuation, symbols, URLs, numbers, and split hyphenated words; 3) expanding contractions (e.g., from *I'm* to *I am*) so to remove stopwords; removing stopwords (using the R package *stopwords* (Benoit et al., 2021)); 4) tokenizing; and 5) lemmatizing. After preprocessing, we constructed the DTM (using the function `dfm` from the package *quanteda*), and weighted the DTM by the TF-IDF (using the function `dfm_tfidf` from the package *quanteda*).

Compounds' appropriateness was obtained by averaging the TF-IDF weights for the constituents of compounds in each document. For example, given the values associated with the words *covid*, *vaccine*, and *deception* (respectively associated with 7.00, 14.00, 3.00), the resulting compounds *covid_vaccine* and *covid_deception* have different values (respectively 10.5 and 5.00).

Note that TF-IDF values are associated with a word in a document, hence a word has different values depending on the document in which it appears. In this way, the compound *covid_vaccine* has different values of appropriateness depending on the context: a high value in a document about covid-19 (where the compound is highly appropriate), but a low value in a document about the death of Lady Diana.

1.3 Statistical analyses

For each measure of creativity, we fitted a series of linear mixed-effects models using the *lme4* (Bates et al., 2015) and the *lmerTest* (Kuznetsova et al., 2017) R packages. To test differences between conspiracy and non-conspiracy subcorpora, for each model, we specified as dependent variable our measures of creativity (one per model), and as fixed effects the dichotomous subcorpus category of being or not a conspiracy document. We added the following covariates. For measures that rely on temporal changes, namely those extracted from the Google corpus such as originality (novelty and rarity), we added as covariate the year in which the document was written or published online. For measures focused on the relationship between compounds' constituents, such as divergence (semantic distance and metaphoricity), we added as covariate the number of constituents. For measures that are built from algorithms that rely on number of words and compounds in the document and their relations such as TF-IDF and topic modelling, namely appropriateness and topical flexibility, we added as covariate compound count and word count (per document) as covariate. For all models, as random intercept, we clustered documents within topics so to aggregate together documents with similar vocabulary. This choice was motivated by the fact that topics might differ in terms of compound use. For example, a topic revolving around technology might use more novel and therefore rare compounds (e.g., *computer chip*) than topics about family (e.g., *mama dada*). Therefore, clustering documents within topics ensures that comparison is made on a comparable set of compounds.

All continuous variables have been converted to *z* scores (mean = 0 and SD = 1) prior to entering the models, so to provide standardized beta (β) coefficients. We report both marginal and conditional R squared (R^2), respectively associated with the variance explained by the fixed effects (marginal) and the variance explained by the entire model that includes both fixed and random effects (conditional). Models fits are reported in section 2.2, with standardized beta coefficient (*β* and their 95% confidence intervals), with standard error (*SE*), as well as *t* and *p* value.

In order to obtain an aggregated measure of creativity, we extracted the Hedges' g effect size from each regression's coefficient and aggregated following a meta-analytic procedure. Note that we first attempted to aggregate these measures as a unique creativity factor, using the *lavaan* package (Rosseel, 2012), but resulted in a bad fit (χ^2 = 83202, $p <$ .001, CFI = .518, RMSEA = .157). We relied on the function `esc_B` from the *esc* package (Lüdecke, 2018) to convert regressions' coefficients into Hedges' g. From the regression output, the function takes as input the unstandardized coefficient as well as the sample size for each group and the standard deviation of the dependent variable and returns the desired effect size (in our case *g*). Once effect sizes and standard errors were obtained from each regression, we computed the aggregated effect size relying on the R package metafor (Viechtbauer, 2010).

2 Results

2.1 Correlation matrix

Variable	1	2	3	4	5	6	7	8	9	10	11
1. Fluency											
2. Flexibility (lexical)	-.37**										
3. Flexibility (topical)	-.41**	.49**									
4. Originality (rarity)	-.05**	.20**	.18**								
5. Originality (novelty)	.04**	.10**	.11**	.59**							
6. Divergence (semantic distance)	-.02**	.04**	.06**	.45**	.30**						
7. Divergence (metaphoricity)	.00	-.02**	-.02**	-.06**	-.01*	.07**					
8. Sophistication (lexical)	.11**	-.03**	-.11**	.11**	.23**	.02**	-.02**				
9. Sophistication (structural)	.17**	-.01*	-.04**	.35**	.25**	.02**	-.01**	.04**			
10. Appropriateness	.26**	-.51**	-.49**	.01	.04**	.07**	.01	.08**	.05**		
11. word count	-.05**	-.22**	-.61**	-.00	-.03**	.03**	.02**	.03**	-.00	.46**	
12. document year	.04**	-.06**	-.11**	.06**	-.17**	-.03**	-.04**	.08**	.02**	.03**	.02**

Note. * indicates $p < .05$. ** indicates $p < .01$.

2.2 Models fits for individual measures

	Fluency			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	0.01 (-0.07, 0.08)	0.04	0.16	0.874
subcorpus [conspiracy]	-0.13 (-0.14, -0.11)	0.01	-17.24	**<0.001**
Random Effects				
σ^2	0.71			
$\tau_{00\ topic}$	0.29			
ICC	0.29			
N_{topic}	200			
Observations	96743			
Marginal R^2 / Conditional R^2	0.003 / 0.295			

	Flexibility (lexical)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	0.03 (-0.02, 0.07)	0.02	1.12	0.264
subcorpus [conspiracy]	-0.03 (-0.05, -0.02)	0.01	-4.09	**<0.001**
Random Effects				
σ^2	0.88			
$\tau_{00\ topic}$	0.11			
ICC	0.11			
N_{topic}	200			
Observations	95105			
Marginal R^2 / Conditional R^2	0.000 / 0.110			

	Flexibility (topical)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	0.13 (0.08, 0.17)	0.02	5.79	**<0.001**

subcorpus [conspiracy]	-0.27 (-0.28, -0.25)	0.01	-35.31	**<0.001**
n compounds norm words z	-0.39 (-0.39, -0.38)	0.00	-115.46	**<0.001**
Random Effects				
σ^2	0.74			
$\tau_{00\ topic}$	0.09			
ICC	0.11			
N_{topic}	200			
Observations	95105			
Marginal R^2 / Conditional R^2	0.152 / 0.245			

	Originality (rarity)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	0.00 (-0.04, 0.05)	0.02	0.06	0.948
subcorpus [conspiracy]	0.16 (0.14, 0.18)	0.01	15.09	**<0.001**
date	0.04 (0.03, 0.05)	0.00	9.82	**<0.001**
Random Effects				
σ^2	0.86			
$\tau_{00\ topic}$	0.10			
ICC	0.11			
N_{topic}	200			
Observations	62926			
Marginal R^2 / Conditional R^2	0.006 / 0.111			

	Originality (novelty)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	-0.04 (-0.08, -0.00)	0.02	-2.04	**0.042**
subcorpus [conspiracy]	0.06 (0.04, 0.08)	0.01	5.97	**<0.001**
date	-0.19 (-0.20, -0.18)	0.00	-46.52	**<0.001**
Random Effects				
σ^2	0.88			

$\tau_{00\ topic}$	0.07
ICC	0.08
N_{topic}	200
Observations	62770
Marginal R^2 / Conditional R^2	0.034 / 0.110

	Divergence (semantic distance)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	-0.05 (-0.09, -0.01)	0.02	-2.73	**0.006**
subcorpus [conspiracy]	0.15 (0.13, 0.16)	0.01	17.11	**<0.001**
structural sophistication	0.03 (0.02, 0.03)	0.00	8.56	**<0.001**
Random Effects				
σ^2	0.93			
$\tau_{00\ topic}$	0.06			
ICC	0.06			
N_{topic}	200			
Observations	95092			
Marginal R^2 / Conditional R^2	0.005 / 0.069			

	Divergence (metaphoricity)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	0.00 (-0.03, 0.03)	0.02	0.07	0.946
subcorpus [conspiracy]	0.04 (0.03, 0.06)	0.01	5.16	**<0.001**
structural sophistication	-0.01 (-0.01, -0.00)	0.00	-2.58	**0.010**
Random Effects				
σ^2	0.96			
$\tau_{00\ topic}$	0.04			
ICC	0.04			
N_{topic}	200			
Observations	95036			
Marginal R^2 / Conditional R^2	0.000 / 0.040			

	Sophistication (lexical)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	-0.01 (-0.07, 0.05)	0.03	-0.40	0.685
subcorpus [conspiracy]	0.10 (0.08, 0.12)	0.01	12.72	**<0.001**
Random Effects				
σ^2	0.78			
$\tau_{00\ topic}$	0.18			
ICC	0.19			
N_{topic}	200			
Observations	95098			
Marginal R^2 / Conditional R^2	0.002 / 0.191			

	Sophistication (structural)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	-0.01 (-0.04, 0.02)	0.01	-0.54	0.591
subcorpus [conspiracy]	0.04 (0.02, 0.06)	0.01	4.45	**<0.001**
Random Effects				
σ^2	0.96			
$\tau_{00\ topic}$	0.04			
ICC	0.04			
N_{topic}	200			
Observations	95105			
Marginal R^2 / Conditional R^2	0.000 / 0.038			

	Appropriateness			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	-0.00 (-0.06, 0.05)	0.03	-0.19	0.852
subcorpus [conspiracy]	-0.11 (-0.13, -0.10)	0.01	-15.49	**<0.001**

word count	0.49 (0.48, 0.49)	0.00	179.93	**<0.001**
Random Effects				
σ^2	0.64			
$\tau_{00\ topic}$	0.13			
ICC	0.17			
N $_{topic}$	200			
Observations	95105			
Marginal R^2 / Conditional R^2	0.233 / 0.361			

3 References

Acar, S., & Runco, M. A. (2015). Thinking in multiple directions: Hyperspace categories in divergent thinking. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(1), 41–53. https://doi.org/10.1037/a0038501

Arzouan, Y., Goldstein, A., & Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, *1160*, 69–81. https://doi.org/10.1016/j.brainres.2007.05.034

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Benoit, K., & Matsuo, A. (2020). *spacyr: Wrapper to the "spaCy" "NLP" Library*. https://CRAN.R-project.org/package=spacyr

Benoit, K., Muhr, D., & Watanabe, K. (2021). *stopwords: Multilingual Stopword Lists*. https://CRAN.R-project.org/package=stopwords

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, *3*(30), 774. https://doi.org/10.21105/joss.00774

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3*(null), 993–1022.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Carmody, S. (2022). *ngramr: Retrieve and Plot Google n-Gram Data*. https://CRAN.R-project.org/package=ngramr

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593. https://doi.org/10.1016/j.cognition.2007.03.004

de Marneffe, M.-C., Dozat, T., Silveira, N., Haverinen, K., Ginter, F., Nivre, J., & Manning, C. D. (2014). Universal Stanford dependencies: A cross-linguistic typology. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 4585–4592. http://www.lrec-conf.org/proceedings/lrec2014/pdf/1062_Paper.pdf

Fradis, A., Mihailescu, L., & Jipescu, I. (1992). The distribution of major grammatical classes in the vocabulary of Romanian aphasic patients. *Aphasiology*, *6*(5), 477–489. https://doi.org/10.1080/02687039208249485

Gold, R., Faust, M., & Goldstein, A. (2010). Semantic integration during metaphor comprehension in Asperger syndrome. *Brain and Language*, *113*(3), 124–134. https://doi.org/10.1016/j.bandl.2010.03.002

Hasebe, Y. (2006). Method of using Wikipedia as Japanese corpus. *Doshisha Studies in Language and Culture*, *9*(2), 373–403. https://doi.org/info:doi/10.14988/pa.2017.0000011063

Kim, M., Crossley, S. A., & Kyle, K. (2018). Lexical Sophistication as a Multidimensional Phenomenon: Relations to Second Language Lexical Proficiency, Development, and Writing Quality. *The Modern Language Journal*, *102*(1), 120–141. https://doi.org/10.1111/modl.12447

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Kyle, K., & Crossley, S. A. (2015). Automatically Assessing Lexical Sophistication: Indices, Tools, Findings, and Application. *TESOL Quarterly*, *49*(4), 757–786. https://doi.org/10.1002/tesq.194

Lakoff, G., & Johnson, M. (2008). *Metaphors We Live By*. University of Chicago Press.

Lüdecke, D. (2018). *Esc: Effect Size Computation For Meta Analysis*. Zenodo. https://doi.org/10.5281/ZENODO.1249218

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, *331*(6014), 176–182. https://doi.org/10.1126/science.1199644

Montani, I., Honnibal, M., Honnibal, M., Van Landeghem, S., Boyd, A., Peters, H., McCann, P. O., Samsonov, M., Geovedi, J., O'Regan, J., Orosz, G., Altinok, D., Kristiansen, S. L., Roman, Explosion Bot, Fiedler, L., Howard, G., Wannaphong Phatthiyaphaibun, Tamura, Y., … GregDubbin. (2021). *explosion/spaCy: V3.2.1: doc_cleaner component, new Matcher attributes, bug fixes and more* (v3.2.1). Zenodo. https://doi.org/10.5281/ZENODO.5764736

Olson, J. A., Nahas, J., Chmoulevitch, D., Cropper, S. J., & Webb, M. E. (2021). Naming unrelated words predicts creativity. *Proceedings of the National Academy of Sciences*, *118*(25). https://doi.org/10.1073/pnas.2022340118

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in*

Natural Language Processing (EMNLP), 1532–1543. https://doi.org/10.3115/v1/D14-1162

Pereira, F., Gershman, S., Ritter, S., & Botvinick, M. (2016). A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data. *Cognitive Neuropsychology*, *33*(3–4), 175–190. https://doi.org/10.1080/02643294.2016.1176907

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36.

Selivanov, D., Bickel, M., & Wang, Q. (2022). *text2vec: Modern Text Mining Framework for R*. https://CRAN.R-project.org/package=text2vec

Viechtbauer, W. (2010). Conducting Meta-Analyses in *R* with the **metafor** Package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Watanabe, K., & Xuan-Hieu, P. (2022). *seededlda: Seeded-LDA for Topic Modeling*. https://CRAN.R-project.org/package=seededlda

5.4 Study 4

Selective exposure drives traffic to conspiracy websites

Alessandro Miani,[1] Théo Schöpfer,[1] Thomas Hills,[2,3] and Adrian Bangerter[1]

[1] Institute of Work and Organizational Psychology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland.

[2] Department of Psychology, University of Warwick, University Road, Coventry CV47AL, UK

[3] The Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, UK

Abstract

Conspiracy theories overlap with extreme ideologies having a detrimental societal impact. To counteract their spread, it is important to understand the routes through which these narratives spread. Cognitive biases and online social bubbles have been previously associated with the spread of conspiracy theories online and extreme partisanship. However, the magnitude of their respective contribution remains unsettled. We address this gap by devising a field study in which we analyze and compare online browsing behavior of an estimated 8M users visiting 2,766 websites classified on ideological types and strength. We find that users' selective exposure is the main drive for accessing ideological and low-quality websites and increases as function of ideological strength. This behavior is more evident in conspiracy, followed by conservative, and reduced in liberal~~,~~ websites. Conspiracy websites differ from political websites in one important aspect: as conspiratorial ideology increases, conspiracy websites show less contribution from social traffic, confirming the primacy of selective exposure. These results replicate on a sample of Italian websites delivering misinformation. We also find that, in absolute terms, online news traffic concentrates on center-liberal bias and that extreme ideological websites are on the fringes. Our results help understand similarities and differences among extreme ideologies and help develop tailored intervention aimed at fighting their spread.

1 Introduction

Conspiracy theories (CTs) are fueled by hate and anger (1–3) and propose violence as a means to re-establish societal order (4–8). As such, belief in CTs overlaps with extreme ideologies (9–11) and are often used by extremist groups for (online) propaganda and lure recruits (12, 13). Cases of lone-wolf culminating in murdering and criminal incidents perpetuated by hand of individuals who believe in CTs have been recorded in the last years (12–16). The spread of CTs is even more societally impactful when large portions of the population endorse health-related CTs (17) that causes the loss of human lives and public funds (18). The societal significance of CTs warrants increasing efforts to understand the pathways through which people access CTs online.

People might reach CTs inadvertently (19), via a link posted on a social network or via a Google search about a particular issue (e.g., COVID-19). Once a conspiratorial worldview has started to build up, it can grow, accelerating in a non-linear fashion due to recursive dynamics (19). CTs are persuasive because they comprise an interconnected network of ideas (20) that give an impression of an overall truth (21) and, because they are freed from veridicality, they can exploit cognitive biases for appealing information (22–24). It follows that exposure to CTs reinforces conspiracy belief which in turn motivates people to search for more conspiracies and join conspiracy groups online (25). Prolonged exposition to CTs, instead of satisfying epistemic and existential needs (26), increases frustration (27) and the likelihood to search for more CTs.

Some scholars propose that the spread of CTs and misinformation is facilitated by the structure and algorithmic ranking (28) of online social networks (e.g., Facebook, YouTube) that promote the formation of echo chambers (29) and foster radicalization towards extreme ideologies (30–32). Following the Capitol Hill storm on January 6th, 2021, the social media accounts of the ex-president Donald J. Trump were suspended and Twitter took down more than 70,000 accounts spreading misinformation (33). Yet, misinformation continued to spread (34) and some scholars suggest that echo chambers do not represent an urgent concern because information diet, at least in the US, is mostly centrist and diverse (35).

Moving from the study of social media structure, scholars have also suggested that cognitive biases, such as *confirmation bias*, might be responsible for shaping individuals' news environment hence providing affordances for the spread of misinformation and CTs (23).

Confirmation bias is a defensive behavior enacted to preserve an ideology that is crucial for making sense of the world (36) and for reducing the unpleasant state of conflict between old and new information (37, 38). Two primary mechanisms drive confirmation bias: selective exposure to congenial information and avoidance of challenging information (37, 39, 40). Confirmation bias is ideologically transversal but is accentuated in people highly committed to their beliefs (41, 42) and with a rigid thinking style (37). In fact, confirmation bias increases as a function of political (43, 44) and conspiratorial (45, 46, 32) ideological strength and interacts with social media environment (47, 48). One explorative study (3) assessed the impact of both social media and cognitive biases on online traffic towards CTs and politically biased websites, showing that both factors drive people to search for information in highly ideological websites.

In short, studies converge showing that both social media and cognitive biases affect information foraging in people committed to strong ideologies such as CTs and political extremes. However, due to heterogeneity of methods and a lack of comprehensive and comparative data, the magnitude of their respective contribution cannot be conclusively drawn. Here, we devise a field study to assess and compare the magnitude of the impact of social media and cognitive biases in information foraging. We assess their role in relative and absolute terms. This is crucial for advancing research on fighting misinformation by tailoring specific intervention, ultimately improving the online informational ecosystem.

2 METHOD

We used a comprehensive sample of websites (N = 2,766) for an estimation of 8-million internet users [see details in supplemental materials (SM) section 1.2.3]. Relying on a methodology previously developed (3), for each website, we extracted measures of incoming traffic divided in percentages of traffic modalities. These modalities are (1) direct, (2) search, and (3) social traffic, which reflect cognitive biases such as (1) selective exposure and (2) openness to challenging information, and (3) social media environment. Direct traffic, i.e., typing the URL of a website in the browser or by recalling it from saved bookmarks, is associated with selective exposure to congenial information (49, 50). In our dataset, echoing other works on selective exposure (41, 43), direct traffic positively correlates with time spent browsing and pages visited (see Table 1 in SM). Search traffic, i.e., accessing a website from a search engine such as Google, is associated with cross-cut exposure hence openness to challenging information (49, 50). Social traffic occurs when a website is accessed via a link posted on social media and therefore reflects social media environment. These measures are expressed as percentage of traffic that reaches a website in a month and is stable over time (see details in SM 1.2.1). We additionally extracted measures of total visits so to derive a more precise estimation of traffic modality in absolute terms.

Our list of websites is divided in four types (see detailed description in SM 1.1): pro-science (legitimate science or evidence-based sources), least-biased (the most credible media sources that might present minimal bias), politically biased (sources that are moderately to strongly biased toward liberal/conservative causes), and conspiracy websites (sources based on unverifiable information related to known conspiracies). These websites are classified on a fine-grained political spectrum and conspiratorial level that represent the ideological strength of a worldview (either conspiratorial, liberal, or conservative). Ideological strength ranges from 0 (no ideology) to 5 (strong ideology) for a total of 13 classes of websites. Pro-science and least-biased websites are assigned with the lowest ideological strength, that is zero. Politically-biased and conspiracy websites range from 1 (low ideology: low conspiracy level and liberal/conservative center bias) to 5 (high ideology: extreme conspiracy and extreme liberal/conservative bias). This measure of ideological strength correlates with quality of sources [an aggregated measure of bias, factuality, credibility, and transparency (51); $r_{2615} = -.84$, $p < .001$; note that pro-science and least-biased websites do not differ in quality, $t_{174.498} = .455$, $p = .650$]. Website types, class, and main variables summary statistics are shown in Table 1.

Table 1

Descriptives of groups

Website type	Class	Strength	Quality	N	Total monthly visits	Traffic direct	Traffic search	Traffic social	example website
Pro-Science	PS00	0	0.87 (0.08)	148	14,553,751 (36,313,705) [50,000-235,220,000]	28.14 (12.45) [8.08-66.01]	59 (18.13) [15.68-90.39]	6.62 (7.36) [0.4-44.28]	*healthline.com*
(Left) Biased Liberal	L05	5	0.45 (0.14)	96	27,53,745 (8,663,799) [479-54,610,000]	43.65 (27.15) [2.13-88.37]	18.4 (19.18) [0.34-75.46]	32.14 (31.79) [1.81-92.97]	*dailykos.com*
	L03	3	0.6 (0.09)	264	6,302,221 (16,657,271) [1,620-139,510,000]	34.5 (18.46) [7.55-87.55]	40.21 (21.78) [2.53-87.66]	17.44 (15.31) [1.26-82.28]	*huffpost.com*
	L01	1	0.79 (0.10)	565	72,496,884 (1,132,912,785) [22,800-2.434e+10]	34.41 (15.51) [5.96-89.36]	47.62 (18.02) [1.13-92.34]	10.97 (9.3) [0.08-86.71]	*bbc.com*
Least Biased	LB00	0	0.88 (0.06)	477	20,300,326 (282,420,656) [11,780-5.66e+09]	35.88 (15.5) [5.85-91.85]	47.48 (17.32) [1.93-88.38]	10.53 (9.56) [0.34-79.58]	*wikipedia.org*
(Right) Biased Conservative	R01	1	0.74 (0.15)	296	10,463,412 (26,365,281) [50,000-194,910,000]	38.13 (16.57) [6.68-91.23]	43.77 (17.91) [2.48-88.51]	10.31 (8.21) [0.18-50.57]	*tribunnews.com*

	R03	3	0.51 (0.14)	290	11,721,936 (44,025,419) [45,260-363,280,000]	43.89 (21.95) [7.49-99.69]	27.04 (22.2) [0-80.91]	15.57 (15.29) [0-75.78]	*foxnews.com*
	R05	5	0.25 (0.09)	335	1,831,205 (6,804,957) [25,840-70,150,000]	50.44 (26.38) [1.81-96.26]	12.94 (15.21) [0-77.94]	22.86 (25.03) [0-96.55]	*breitbart.com*
Conspiracy	CT01	1	0.27 (0.11)	38	2,332,578 (3,759,837) [56,360-15,590,000]	26.5 (16.63) [6.03-71.8]	47.86 (28.21) [0.38-90.45]	21.24 (28.94) [0.55-85.83]	*brightside.me*
	CT02	2	0.21 (0.04)	34	799,078 (1,283,574) [50,000-5,640,000]	43.25 (18.93) [15.47-78.26]	30.75 (22.63) [0.67-65.85]	17.88 (23.37) [2.57-78.04]	*mercola.com*
	CT03	3	0.2 (0.06)	52	806,588 (1,503,291) [50,000-7,050,000]	44.57 (19.81) [18.29-79.03]	31.21 (22.72) [2.42-79.48]	16.8 (18.38) [0.4-70.49]	*collective-evolution.com*
	CT04	4	0.19 (0.08)	70	1,757,607 (6,387,145) [50,000-39,510,000]	60.99 (19.21) [14.12-89.96]	17.18 (15.64) [2.08-75.98]	15.28 (15.32) [1.57-64.51]	*gaia.com*
	CT05	5	0.16 (0.10)	101	773,043 (1,727,372) [50,000-10,550,000]	60.21 (18.92) [15.29-90.64]	15.29 (10.74) [1.04-50.6]	14.38 (15.57) [0.87-72.3]	*infowars.com*

Note. Values are expressed as Mean (SD) and [range]. Quality: website quality obtained by (51); N: number of websites in our dataset.

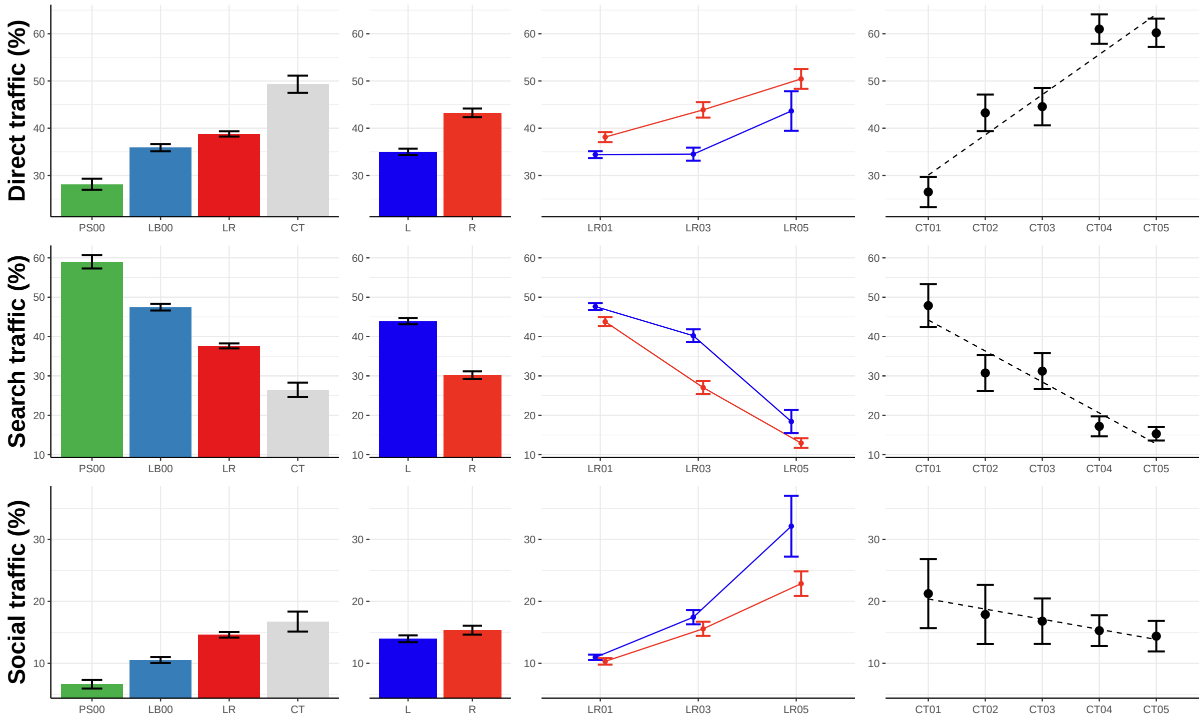
3 RESULTS

We start testing whether traffic modalities (i.e., direct, search, and social) is different between website categories (i.e., from pro-science to conspiracy). We then test whether traffic modalities relate to ideological strength and to liberal and conservative political ideology. The strength of the relationship between ideological strength and traffic modality is further compared between ideological types (i.e., conspiracy, liberal, and conservative) so to comparatively assess the magnitude of this relationship. We then test whether traffic towards websites on different ideological types is mostly driven by individual users or by social groups. We replicate these analyses on an Italian set of websites delivering misinformation, and finally estimate the absolute rate of traffic towards each website class. All models were fitted by adding as covariate the website popularity (an unpopular website, with low total visits, could be accessed directly because it does not appear within the first pages of a search engine) and websites' percentage of uninterested visits (bounce rate). All p-values are < .001 unless otherwise specified.

We started by testing differences of traffic modality by website types (see model fits and post-hoc comparison in SM 2.1). Direct traffic was higher in conspiracy compared to other website types ($F_{5,1926} = 170.696$, $R^2 = .305$), followed by politically biased, least-biased, and pro-science websites (all differences at ps < 0.01). Search traffic showed the reversed pattern ($F_{5,1926} = 79.217$, $R^2 = .168$, all ps < 0.001). Social traffic, mirroring direct traffic patterns, was higher in conspiracy compared to other website types ($F_{5,1925} = 44.865$, $R^2 = .102$, all ps < 0.046 except between pro-science and least-biased websites). As for ideological strength (see SM 2.2), both direct ($\beta = .143$, $F_{3,1928} = 322.992$, $R^2 = .333$) and social ($\beta = .186$, $F_{3,1927} =$

125.547, $R^2 = .162$) traffic were positively related to ideological strength, while search traffic was negatively related ($\beta = -.301$, $F_{3,1928} = 307.069$, $R^2 = .322$) to ideological strength. Compared to liberal websites, conservative websites showed more direct ($\beta = .323$), less search ($\beta = -.550$), and more social ($\beta = .123$, $p = .033$) traffic. See Figure 1.

Figure 1



We then assessed whether the relationship between traffic modality and ideological strength was different across ideological type (i.e., conspiracy, liberal, and conservative) by testing the interaction between ideological strength and ideological type (see SM 2.3). Direct traffic was positively related to ideological strength ($\beta = .549$, $F_{7,1409} = 118.838$, $R^2 = .368$). Compared to conspiracy, liberal websites showed marginally less direct traffic ($\beta = -.183$, $p = 0.053$) but conservative websites did not differ from conspiracy websites ($\beta = .039$, $p = .682$). A slope analysis was performed so to assess the degree of the relationship between direct traffic and ideological strength by ideological type. The conspiracy ($\beta = .303$) and the conservative ($\beta = .157$) slopes were positive while the slope for liberal websites was not significantly different from zero ($\beta = .035$, $p =.192$). All slopes were different from each other (at ps < .0135). Search traffic was also related to ideological strength ($\beta = -.551$, $F_{7,1409} = 102.625$, $R^2 = .334$). Compared to conspiracy, left politically-biased websites showed higher social traffic ($\beta = .292$, $p = 0.002$) but right-wing websites did not differ ($\beta = .008$, $p = .932$). All slopes were negative (βs < -.247) and not different from each other (except between liberal and conservative websites, β = .103, p = .007). As for social traffic, the model ($F_{7,1409}$ = 46.948, R^2 = .185) did not yield a general effect of social traffic on ideological strength (β = -.041, p = .681). Compared to conspiracy, left and right websites showed less social traffic (left: β = -.241, p = 0.03; right: β = -.315, p = 0.005). The conspiracy slope was not different from zero (β = -.003, p = .681), but liberal and conservative websites showed instead a positive relationship (βs > .209). The three slopes were all different from each other (at ps < .026).

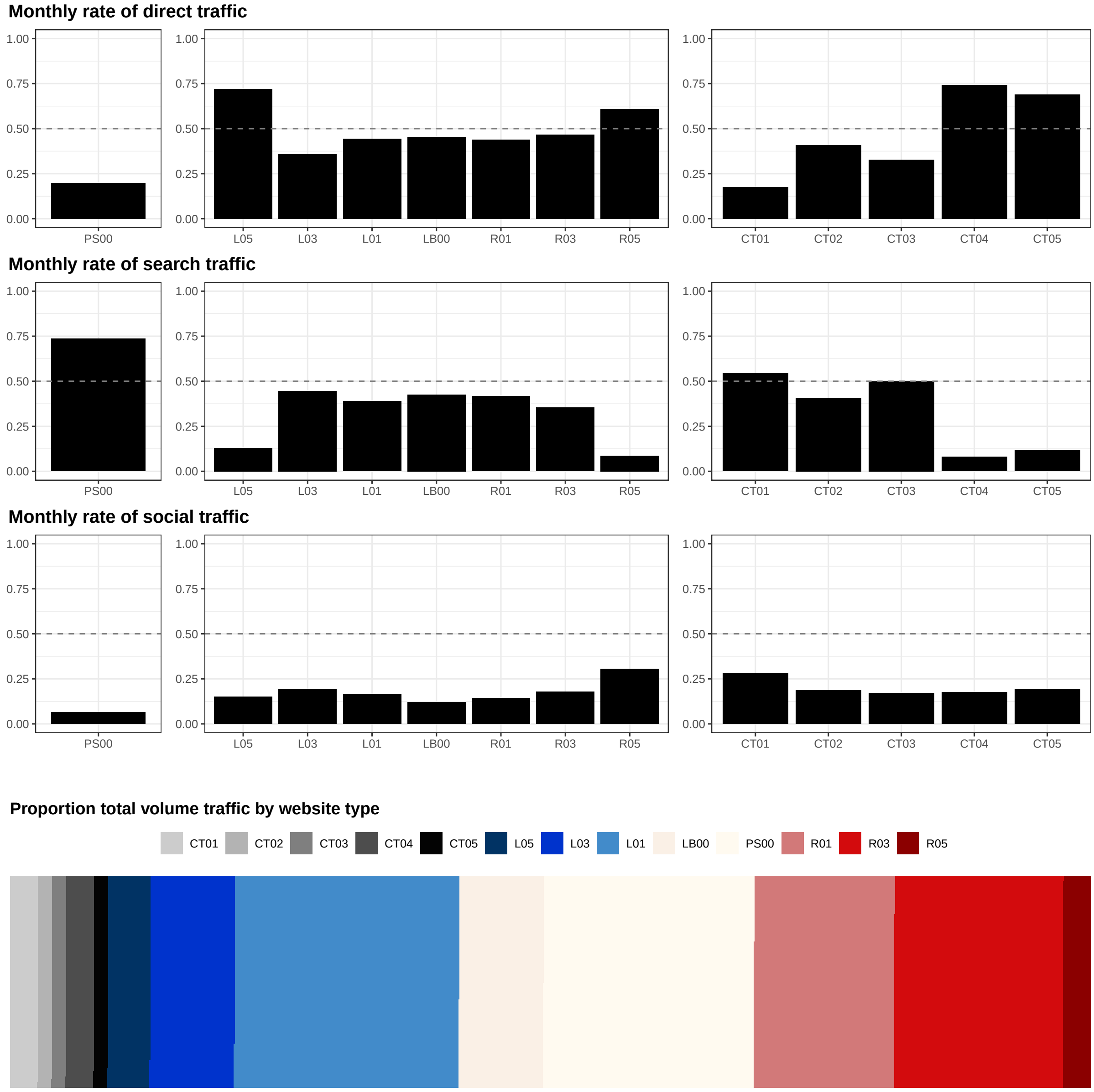
We further tested selective exposure in comparison to social traffic so to assess the impact of self-driven traffic against group-driven traffic. We created a measure, named *individual traffic*, by subtracting social-driven traffic (traffic to websites coming from social media, mail, and referral websites) to individual-driven traffic (direct traffic). In doing so, we obtained a single measure where positive values indicate predominance of individual traffic, whereas negative values indicate predominance of social traffic. We tested whether individual traffic was related to ideological strength and was different between website types. The model ($F_{7,1409}$ = 79.727, R^2 = .28, see SM 2.4) yielded an overall positive effect of individual traffic on ideological strength (β = .305) meaning that as the ideology increases, so does the percentage of individual, over social, traffic increases. Conservative and liberal websites did not differ from conspiracy websites on individual traffic (|βs| < .059, ps > .581). The conspiracy slope was positive (β = .168, p = .0014) and different from liberal (β = -.140) and conservative (β = -.059, p = .014) slopes (at ps < .001). Liberal and conservative slopes were not different to each other (p = .083). See figure 2 in SM 2.4.

We also tested whether access modality, irrespective of ideological type and strength, was related to quality of websites [(51), see SM 2.5]. Direct traffic was negatively related to quality (β = -.242, $F_{3,1855}$ = 306.815, R^2 = .330), while social traffic was positively related to quality of websites (β = .494, $F_{3,1855}$ = 256.825, R^2 = .292). Search traffic showed a positive link (β = -.289, $F_{3,1855}$ = 103.785, R^2 = .142). We replicated these analyses, testing the relationship between traffic modality and website quality (51), with a sample of Italian websites delivering misinformation [(52), see SM 2.6]. Quality of websites was negatively related to direct traffic (β = -.388, p = .003) and positively with search traffic (β = .568). Traffic from social media was not related to quality (β = -.154, p = .300).

Finally, we assessed the total volume visits by website classes and computed the rate of traffic modality from absolute traffic (see SM 2.7). This was done by extracting the absolute number of visits by traffic type from each website percentages. This is useful to assess comparatively the proportion of traffic modalities and to provide an estimate of visits to websites classes (note that prior to this analysis we removed the four topmost visited not-news websites such as Facebook, Wikipedia, Bing, and MSN). The rate of direct, search, and social traffic (including mail and referrals) computed from the total visits volume show patterns similar to those explored so far on averaged percentages. Highly conspiratorial websites and extreme politically-biased websites showed an absolute rate of direct traffic above .50. While direct traffic increases towards high ideological websites, search traffic decreases. In absolute terms, social traffic accounts, on average and for all website classes, for less than 25% of total traffic volume (See first three rows in figure 3). Last row of Figure 3 shows the proportion of traffic generated by each class of website. We estimated total visits by averaging total visits by each website class and then computing the proportion for the whole dataset. We did not simply count visits because different website sample size by class affects total visits volume. The highest proportion of visits was directed towards least-biased and pro-science websites (classes LB00 and PS00, 27.15%) followed by center-left websites (class L01, 21.30%). As a group, conspiracy websites accounted for 8.43% of visits (CT01 = 3.04% CT02 = 1.04% CT03 = 1.05% CT04 = 2.29% CT05 = 1.01%), while extreme liberal and conservative websites (classes L05 and R05) altogether accounted for 5.98% of visits. Visits towards liberal websites were slightly higher (33.11%) than visits towards conservative websites (31.31%).

While here we focused on direct, search, and social traffic, our dataset (freely available at https://osf.io/PUT_FILES_HERE) includes other measures. These measures are related to 1) the partitioning of social media traffic between Facebook, YouTube, Reddit, and Twitter, 2) engagement measures such as number of pages visited and time of each visit, 3) traffic from mail and referral websites, and 4) total visits and bounce rate. While a detailed analysis goes beyond the scope of the current investigation, we nevertheless provide a visual exploration of these measures by website types and quality as well as ideological strength (see SM 1.2 and 2.8) so to stimulate further analyses.

Figure 3 – Absolute estimate of traffic modality rate and



4 Discussion

We analyzed websites' incoming traffic (direct, search, and social) to assess in a comparative and comprehensive way the impact of cognitive biases (selective exposure and openness to challenging information) and social media on information foraging towards ideological (conspiracy and politically biased) websites.

We found that selective exposure, as measured by direct traffic, is the main drive for accessing ideological websites and increases as function of ideological strength. This behavior is particularly marked in conspiracy, followed by conservative, and reduced in liberal websites. Conspiracy and conservative websites' users show similar patterns of information foraging. However, conspiracy websites differ from politically biased (both liberal and conservative) websites in one important aspect: the influence of social group decreases as websites' conspiratorial ideology increases. Selective exposure is negatively associated with website quality. It also is inversely related to openness to challenging information (as measured by search traffic). Overall, these patterns are stable also when absolute traffic is considered: selective exposure is the main drive to extreme ideological conspiratorial and politically extreme websites. Analysis of absolute traffic shows that extreme websites attract only a small proportion of the overall traffic.

Our findings converge with existing survey literature showing that selective exposure relates to ideological strength (41, 42), is amplified in people holding conservative (53) and conspiratorial (45, 46) worldviews (53). This work, nevertheless, extends prior research by comparing the strength of these relationships across different worldviews within a unique dataset.

We compared the magnitude of cognitive biases and social media impact on information foraging finding that social media contribute to a lesser extent than selective exposure towards highly ideological websites. This finding might suggest that individual-level intervention should be prioritized over group-level intervention. In other words, since social traffic towards highly ideological websites contributes to a lesser extent than individual selective exposure, banning social media misinformation accounts [e.g., (33)] would be marginally effective. Differently, individual-level interventions, such as improving digital literacy (54, 55), would be more effective in preventing access towards highly ideological websites.

Mirroring other studies (35, 47), we show that most of the traffic in our dataset is politically centered and slightly slanted on the left, while extreme ideological websites account only for a small percentage. Although seemingly reassuring, some scholars remind that the potential for a polarized future remains (35). We remind, however, that our estimate of total traffic, especially that towards conspiracy websites, might be exaggerated due to sampling bias: compared to the whole internet, we have an exaggerated number of conspiracy websites. Relatedly, our dataset does not include other-than-news type of websites (e.g., entertainment, technology, hobby), whose presence would reduce our estimated shares towards fringes websites. We nevertheless provide averaged monthly traffic, which will be useful for future works. Future studies could estimate the number of conspiracy websites on the internet so to complement our findings.

We derived users' behavior from website traffic measures relying on methods previously developed (3, 47). While we do not know people's ideology and their individual online behavior, compared to survey studies, our method allows to reach a generous sample of users (around 8 million, otherwise impracticable to obtain via surveys) who potentially share a common worldview, shows high ecological validity, and overcomes common problems with study based on self-reported data (56). Future studies could extend our work by implementing web-tracking data with survey data [see e.g., (56)].

While we statistically controlled for important covariates that could affect our results (i.e., website popularity and percentage of uninterested visits), our three main variables of interest are not free from measurement error. First, the presence of bots, aimed at increasing traffic hence revenues from advertising (57), might inflate measures of direct traffic. We tested whether websites differ in advertising revenues and whether revenues are associated with direct traffic. We did not find evidence for such confound (see SM 1.2.2), or at least there were no differences between websites. While we cannot exclude with confidence the presence of bots, we can tentatively claim that such a confound would be evenly spread across website classes and types. Moreover, some conspiracy websites (e.g., Infowars) rely on revenues from their own branded products (58) hence the presence of bots would be even counterproductive. Second, not all search traffic is purely informational (e.g., "covid symptoms"): it is estimated that 10% of all Google queries are navigational [i.e., searching for content within a website, e.g., "covid symptoms wiki" (59)]. Therefore, these queries, practically situated between search and direct traffic, should be considered, theoretically, as selective exposure. Moreover, a comparison of search engines shows that Google reduces exposure to conspiracy theories (60). However, we do not know whether Google affects traffic towards politically biased websites —for which we found an effect comparable in size to that of conspiracy websites. Third, social media traffic is a measure that mixes individual factors such as being part of echo chambers (e.g., Facebook groups of like-minded individuals) and external factors related to the system algorithm such as filter bubbles (e.g., YouTube recommendations). These factors, unfortunately, are difficult to separate in our dataset. We nevertheless created a measure of individual traffic that opposed individual direct traffic to traffic “suggested” by others such as that from social media, mail, and referral websites.

5 References

1. D. B. W. and J. Dawson, From anti-vaxxer moms to militia men: Influence operations, narrative weaponization, and the fracturing of American identity. *Brookings* (2021) (December 8, 2022).

2. A. Fong, J. Roozenbeek, D. Goldwert, S. Rathje, S. van der Linden, The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Process. Intergroup Relat.* **24**, 606–623 (2021).

3. A. Miani, T. Hills, A. Bangerter, LOCO: The 88-million-word language of conspiracy corpus. *Behav. Res. Methods* **54**, 1794–1817 (2021).

4. R. Imhoff, L. Dieterle, P. Lamberty, Resolving the Puzzle of Conspiracy Worldview and Political Activism: Belief in Secret Plots Decreases Normative but Increases Nonnormative Political Engagement. *Soc. Psychol. Personal. Sci.* **12**, 71–79 (2021).

5. D. Jolley, K. M. Douglas, A. C. Leite, T. Schrader, Belief in conspiracy theories and intentions to engage in everyday crime. *Br. J. Soc. Psychol.* **58**, 534–549 (2019).

6. D. Jolley, J. L. Paterson, Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *Br. J. Soc. Psychol.* **59**, 628–640 (2020).

7. A. W. Kruglanski, E. Molinario, M. Ellenberg, G. Di Cicco, Terrorism and conspiracy theories: A view from the 3N model of radicalization. *Curr. Opin. Psychol.* **47**, 101396 (2022).

8. B. Rottweiler, P. Gill, Conspiracy Beliefs and Violent Extremist Intentions: The Contingent Effects of Self-efficacy, Self-control and Law-related Morality. *Terror. Polit. Violence*, 1–20 (2020).

9. R. Imhoff, *et al.*, Conspiracy mentality and political orientation across 26 countries. *Nat. Hum. Behav.* **6**, 392–403 (2022).

10. J.-W. van Prooijen, A. P. M. Krouwel, T. V. Pollet, Political Extremism Predicts Belief in Conspiracy Theories. *Soc. Psychol. Personal. Sci.* **6**, 570–578 (2015).

11. J.-W. van Prooijen, S. M. H. C. Kuijper, A comparison of extreme religious and political ideologies: Similar worldviews but different grievances. *Personal. Individ. Differ.* **159**, 109888 (2020).

12. J. Bartlett, C. Miller, The power of unreason: Conspiracy theories, extremism and counter-terrorism (Demos London, 2010).

13. A. Basit, Conspiracy Theories and Violent Extremism. *Count. Terror. Trends Anal.* **13**, 1–9 (2021).

14. D. Allington, T. Joshi, "What Others Dare Not Say": An Antisemitic Conspiracy Fantasy and Its YouTube Audience. *J. Contemp. Antisemitism* **3**, 35–54 (2020).

15. D. Dave, D. McNichols, J. Sabia, "Political Violence, Risk Aversion, and Non-Localized Disease Spread: Evidence from the U.S. Capitol Riot" (National Bureau of Economic Research, 2021).

16. A. Garry, S. Walther, R. Rukaya, A. Mohammed, QAnon Conspiracy Theory: Examining its Evolution and Mechanisms of Radicalization. *J. Deradicalization*, 152–216 (2021).

17. J. E. Oliver, T. Wood, Medical Conspiracy Theories and Health Behaviors in the United States. *JAMA Intern. Med.* **174**, 817 (2014).

18. N. Nattrass, AIDS and the Scientific Governance of Medicine in Post-Apartheid South Africa. *Afr. Aff.* **107**, 157–176 (2008).

19. R. M. Sutton, K. M. Douglas, Rabbit Hole Syndrome: Inadvertent, accelerating, and entrenched commitment to conspiracy beliefs. *Curr. Opin. Psychol.* **48**, 101462 (2022).

20. A. Miani, T. Hills, A. Bangerter, Interconnectedness and (in)coherence as a signature of conspiracy worldviews. *Sci. Adv.* **8** (2022).

21. G. Pennycook, T. D. Cannon, D. G. Rand, Prior exposure increases perceived accuracy of fake news. *J. Exp. Psychol. Gen.* **147**, 1865–1880 (2018).

22. A. Acerbi, Cognitive attraction and online misinformation. *Palgrave Commun.* **5**, 15 (2019).

23. T. T. Hills, The Dark Side of Information Proliferation. *Perspect. Psychol. Sci.* **14**, 323–330 (2019).

24. J.-W. van Prooijen, J. Ligthart, S. Rosema, Y. Xu, The entertainment value of conspiracy theories. *Br. J. Psychol.* **113**, 25–48 (2022).

25. J. Uscinski, A. M. Enders, C. Klofstad, J. Stoler, Cause and effect: On the antecedents and consequences of conspiracy theory beliefs. *Curr. Opin. Psychol.* **47**, 101364 (2022).

26. K. M. Douglas, R. M. Sutton, A. Cichocka, The Psychology of Conspiracy Theories. *Curr. Dir. Psychol. Sci.* **26**, 538–542 (2017).

27. L. Liekefett, O. Christ, J. C. Becker, Can Conspiracy Beliefs Be Beneficial? Longitudinal Linkages Between Conspiracy Beliefs, Anxiety, Uncertainty Aversion, and Existential Threat. *Pers. Soc. Psychol. Bull.*, 014616722110609 (2021).

28. E. Pariser, The filter bubble: What the Internet is hiding from you (penguin UK, 2011).

29. M. Cinelli, G. D. F. Morales, A. Galeazzi, W. Quattrociocchi, M. Starnini, The echo chamber effect on social media. *Proc. Natl. Acad. Sci.* **118**, e2023301118 (2021).

30. A. Bessi, *et al.*, Science vs Conspiracy: Collective Narratives in the Age of Misinformation. *PLOS ONE* **10**, e0118093 (2015).

31. D. Mocanu, L. Rossi, Q. Zhang, M. Karsai, W. Quattrociocchi, Collective attention in the age of (mis)information. *Comput. Hum. Behav.* (2015) https:/doi.org/10.1016/j.chb.2015.01.024.

32. F. Zollo, *et al.*, Debunking in a world of tribes. *PLoS ONE* (2017) https:/doi.org/10.1371/journal.pone.0181821.

33. J. Tollefson, Tracking QAnon: how Trump turned conspiracy-theory research upside down. *Nature*, d41586-021-00257-y (2021).

34. J. Lasser, *et al.*, Social media sharing of low-quality news sources by political elites. *PNAS Nexus* **1**, pgac186 (2022).

35. A. Guess, B. Nyhan, B. Lyons, J. Reifler, Avoiding the echo chamber about echo chambers. *Kn. Found.* **2**, 1–25 (2018).

36. M. E. Koltko-Rivera, The Psychology of Worldviews. *Rev. Gen. Psychol.* **8**, 3–58 (2004).

37. W. Hart, *et al.*, Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychol. Bull.* **135**, 555–588 (2009).

38. L. Festinger, *A theory of cognitive dissonance* (Stanford University Press, 1957).

39. R. S. Nickerson, Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Rev. Gen. Psychol.* **2**, 175–220 (1998).

40. N. J. Stroud, *Selective Exposure Theories*, K. Kenski, K. H. Jamieson, Eds. (Oxford University Press, 2014).

41. S. Knobloch-Westerwick, J. Meng, Looking the Other Way: Selective Exposure to Attitude-Consistent and Counterattitudinal Political Information. *Commun. Res.* **36**, 426–448 (2009).

42. C. S. Taber, M. Lodge, Motivated Skepticism in the Evaluation of Political Beliefs. *Am. J. Polit. Sci.* **50**, 755–769 (2006).

43. A. Westerwick, B. K. Johnson, S. Knobloch-Westerwick, Confirmation biases in selective exposure to political online information: Source bias vs. content bias. *Commun. Monogr.* **84**, 343–364 (2017).

44. G. D. H. Pearson, S. Knobloch-Westerwick, Is the Confirmation Bias Bubble Larger Online? Pre-Election Confirmation Bias in Selective Exposure to Online versus Print Political Information. *Mass Commun. Soc.* **22**, 466–486 (2019).

45. N. Georgiou, P. Delfabbro, R. Balzan, Conspiracy-Beliefs and Receptivity to Disconfirmatory Information: A Study Using the BADE Task. *SAGE Open* **11**, 215824402110061 (2021).

46. S. A. K. Kuhn, R. Lieb, D. Freeman, C. Andreou, T. Zander-Schellenberg, Coronavirus conspiracy beliefs in the German-speaking general population: endorsement rates and links to reasoning biases and paranoia. *Psychol. Med.*, 1–15 (2021).

47. A. Guess, B. Nyhan, J. Reifler, Exposure to untrustworthy websites in the 2016 US election. *Nat. Hum. Behav.* **4**, 472–480 (2020).

48. A. Guess, J. Nagler, J. Tucker, Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Sci. Adv.* **5** (2019).

49. M. Wojcieszak, E. Menchen-Trevino, J. F. F. Goncalves, B. Weeks, Avenues to News and Diverse News Exposure Online: Comparing Direct Navigation, Social Media, News Aggregators, Search Queries, and Article Hyperlinks. *Int. J. Press.* **27**, 860–886 (2022).

50. A. S. Cardenal, C. Aguilar-Paredes, C. Galais, M. Pérez-Montoro, Digital Technologies and Selective Exposure: How Choice and Filter Bubbles Shape News Media Exposure. *Int. J. Press.* **24**, 465–486 (2019).

51. H. Lin, *et al.*, "High level of agreement across different news domain quality ratings" (PsyArXiv, 2022) https:/doi.org/10.31234/osf.io/qy94s (December 25, 2022).

52. F. Carrella, A. Miani, S. Lewandowsky, IRMA: the 335-million-word Italian coRpus for studying MisinformAtion in *The 17th Conference of the European Chapter of the Association for Computational Linguistics*, (submitted).

53. D. Romer, K. H. Jamieson, Conspiratorial thinking, selective exposure to conservative media, and response to COVID-19 in the US. *Soc. Sci. Med.* **291**, 114480 (2021).

54. A. M. Guess, K. Munger, Digital literacy and online political behavior. *Polit. Sci. Res. Methods*, 1–19 (2022).

55. M. Mosleh, G. Pennycook, A. A. Arechar, D. G. Rand, Cognitive reflection correlates with behavior on Twitter. *Nat. Commun.* **12**, 921 (2021).

56. A. Guess, B. Nyhan, J. Reifler, Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *Eur. Res. Counc.* **9**, 4 (2018).

57. K. Springborn, P. Barford, Impression Fraud in On-line Advertising via ${$Pay-Per-View$}$ Networks in *22nd USENIX Security Symposium (USENIX Security 13)*, (2013), pp. 211–226.

58. H. V. den Bulck, A. Hyzen, Of lizards and ideological entrepreneurs: Alex Jones and Infowars in the relationship between populist nationalism and the post-global media ecology. *Int. Commun. Gaz.* **82**, 42–59 (2020).

59. B. J. Jansen, D. L. Booth, A. Spink, Determining the informational, navigational, and transactional intent of Web queries. *Inf. Process. Manag.* **44**, 1251–1266 (2008).

60. A. Urman, M. Makhortykh, R. Ulloa, J. Kulshrestha, Where the earth is flat and 9/11 is an inside job: A comparative algorithm audit of conspiratorial information in web search results. *Telemat. Inform.* **72**, 101860 (2022).

Selective exposure drives traffic to conspiracy websites

Supplemetal Material

Contents

1 Method

1.1 Material

Our list of websites was gathered (in November 2020) surveying the whole content of mediabiasfactcheck (MBFC, mediabiasfactcheck.com). MBFC contains manual annotations and bias analyses for over 2,000 (as of November 2020) mostly news, websites. According to the method used in MBFC, each website's bias is evaluated on four criteria such as biased wording headlines (e.g., the source uses loaded words to convey emotion to sway the reader), factual sourcing (e.g., the source reports factually and backs up claims with well-sourced evidence), story choices (e.g., the source reports news from both sides), and political affiliation (e.g., the source endorses a particular political ideology). Factual reporting is based on the factual sourcing used for assessing bias. For each website, a minimum of 10 headlines and 5 news stories are assessed by MBFC experts. Low and very low factual reporting sources are those that need to be fact-checked for intentional fake news, conspiracy, and propaganda. Furthermore, although criticized (Albarracin et al. 2021), MBFC has been used as ground-truth by other researchers to study fake news, media bias classifiers, automatic fact-checking systems, and conspiracy websites (Baly et al. 2018; Cardenal et al. 2019; Miani, Hills, and Bangerter 2021; Pennycook and Rand 2019; Risius, Aydinguel, and Haug 2019), and has high level of agreement with other annotation systems (Lin et al. 2022).

In MBFC, websites are classified as either pro-science, least-biased, politically biased, or conspiracy.

- **Pro-science** sources consist of legitimate science or provide evidence-based information through the use of credible scientific sourcing. Legitimate science follows the scientific method, is unbiased and does not use emotional words. These sources also respect the consensus of experts in the given scientific field and strive to publish peer reviewed science. Some sources in this category may have a slight political bias but adhere to scientific principles.
- **Least-biased** sources have minimal bias and use very few loaded words (wording that attempts to influence an audience by using appeal to emotion or stereotypes). Reporting is factual and usually sourced. These are the most credible media sources.
- **Politically biased** sources are moderately to strongly biased toward liberal/conservative causes through story selection and/or political affiliation. They may use strong loaded words, publish misleading reports and omit reporting of information that may damage liberal/conservative causes. Some sources in this category may be untrustworthy.
- **Conspiracy** sources may publish unverifiable information related to known conspiracy theories (e.g., the New World Order, Illuminati, false flags, aliens, antivaccination propaganda, etc.) and unverifiable information that is not always supported by evidence. These sources may be untrustworthy for credible/verifiable information.

Although we do not know who access the websites in our dataset, studies suggest that there is an overlap between political side of users and websites. Tracking users behavior and estimating the website conservative/liberal shares for 459 websites, Gentzkow and Shapiro (2011) show that conservative and liberal online news outlets are accessed by people endorsing conservative and liberal ideologies (Gentzkow and Shapiro 2011) – but note that people on the conservative side are more likely to endorse conspiracy theories (Romer and Jamieson 2021).

1.2 Measures

Similar to other works (Cardenal et al. 2019), we used website incoming traffic as a proxy for reconstructing users behavior. Measures of websites incoming traffic was collected (in November 2020) from Similarweb (SW, https://www.similarweb.com), a web analytic company that extracts traffic measures for about 100 million websites in 190 countries across 10 billion content pages. Measures are expressed as monthly average. Contrary to other tools (e.g., Alexa or Google Analytics), SW offers a wide range of free features, otherwise accessible upon a registration or monthly subscription.

From SW, we collected the following measures.

- **Traffic modality** refers to the modality by which users access websites. Expressed in percentages, it is

partitioned into *direct*, *search*, *social*, *mail*, and *referral* (that sum up to 100%).
  - **Direct traffic** is the percentage of traffic generated by users who access a website by typing the URL into their browser, or recalling it from a saved bookmark or any links from outside the browser (but not mail). Direct traffic is associated with selective exposure (Cardenal et al. 2019).
  - **Search traffic** is the percentage of traffic generated by users who access a website via a search query through a search engine (e.g., Google). Note that compared to other search engines (Yahoo, Bing, DuckGoGo, and Yandex), Google almost entirely reduces exposure to conspiracy theories (Urman et al. 2022). Note also that around 10% of all Google queries are navigational (Jansen, Booth, and Spink 2008), i.e., are queries aimed at finding information within a specific website (e.g., "wiki covid"). Search traffic is associated with openness to cross-cut uncongenial information (Cardenal et al. 2019).
  - **Social traffic** is the percentage of traffic generated by users who access a website by clicking on a post in a social media (e.g., Facebook, Reddit, Twitter, YouTube). This measure allows to assess the impact of social bubbles (Pariser 2011; Sunstein 2018). The percentage of social traffic is further partitioned into traffic coming from specific social media such as *Facebook*, *Reddit*, *Twitter*, and *YouTube*.
  - **Mail traffic** is the percentage of traffic generated by users who access a website by clicking on a URL from an email.
  - **Referral traffic** is the percentage of traffic generated by users who access a website by clicking on a URL from a webpage.
  - In order to test the impact of social group on traffic modality, we derived a measure named **individual traffic**. This measure is computed by taking direct traffic and subtracting to it the sum of social, mail, and referral traffic.
- **Engagement** refers to the users' behavior once reached a website and relates to how well users engage with a website.
  - **Pages per visit** reports the average number of pages visited in each visit.
  - **Visit duration** reports the average time spent (in seconds) on a website for each visit.
  - **Bounce rate** is the website's percentage of traffic generated by users who enter a site, take no action, and leave after visiting only one page. It is thought as a control to partial out the proportion of "unintentional" visits.
- **Total visits** is the number of monthly total visits to a website. In theory, an unpopular website is accessed directly because it does not appear within the first pages of a search engine (i.e., low rank). Rank and Visits correlate highly ($r_{1912} = -.97$).

1.2.1 Measure stability over time

We tested whether our measures are stable over time by testing whether these measures correlate with those collected for LOCO (Miani, Hills, and Bangerter 2021) in July 2020. The tests yielded high correlations, meaning that our measures are stable through time [Total monthly visits: $r_{83} = .987$; direct traffic: $r_{83} = .973$; search traffic: $r_{83} = .996$; social traffic: $r_{83} = .958$ (all $ps < .001$)]. For the same sample of websites, we re-collected these metrics in November 2022. Again, they show high agreement [Total monthly visits: $r_{84} = .959$; direct traffic: $r_{84} = .851$; search traffic: $r_{84} = .916$; social traffic: $r_{84} = .662$ (all $ps < .001$)].

1.2.2 Bots and revenues

It is possible that bots run on websites to increase traffic hence revenues from advertising (Springborn and Barford 2013). Site owners that monetize through advertising running on their websites, buy traffic (known as invalid traffic) so to inflate impressions hence their revenues by frauding advertising agencies. There are online services that sell traffic for websites and social media. Here, we attempt to assess the impact of direct traffic on revenues so to possibly rule out the confound of invalid direct traffic. From the website Mustats (https://www.mustat.com), we collected (in November 2022) measures of revenues for each website in our dataset (obtaining data for 2,387 websites). Revenues metrics are divided into *cost-per-mille* (CPM), money for every 1,000 impressions of a given advertisement, and *cost-per-click* (CPC), money for every click on a given advertisement, per month. We also derived measures of CPC and CPM normalized by monthly visitors, obtaining a total of four revenue measures.

We tested whether revenues were different by groups and classes of websites [classes: $N = 13$ (CT01-CT05, LB00, PS00, LR01-LR05); groups: $N = 4$ (pro-science, least-biased, biased, and conspiracy), see Table 1 in main text] by running eight ANOVAs. Results ($Fs < 2.01, ps > .110$) suggest that revenues did not change between groups. We then ran another series of eight regression models to test whether revenues were impacted by direct traffic for each group (by adding the interaction of group per direct traffic predicting revenues) adding total visits as covariate. For each model, we extracted the standardized slope of direct traffic on revenue by group. Slope beta coefficients ranged between $\beta = -.011$ and $\beta = .007$, with a mean of $\beta = -.00004$, all non-significant. This suggests that revenues are not associated with direct traffic.

Although we could not assess nor quantify the presence of bots, our tests at least show that there are no differences in revenues by group of websites and that direct traffic does not impact revenues. Therefore, if bots are present, such a confound would be at least evenly spread across groups. Moreover, taking the case study of Alex Jones' Infowars (infowars.com), the website generates revenue through the sales of an *Infowars Life* brand and almost solely runs ads for Infowars Life products (Bulck and Hyzen 2020). While we do not know whether other conspiracy websites rely on the same marketing strategy, this example suggest that the presence of bots, at least in some websites, would even be counterproductive.

1.2.3 Unique visitor estimation

We estimated how many unique visitors are covered in our dataset. Although our dataset is composed of 2,766 websites, we collected data for about 50-billion monthly visits (summing all visits in our dataset). Relying on data provided by Guess et al. (2021), that is 19,105,773 URL visits produced in two months by 1,551 participants, we estimated that our dataset covers, within a month, a total of 8,092,280 internet users who performed a reasonable average of 205 visits a day. Note that we also estimated our internet users sample size based on Goel, Hofman, and Sirer (2021), but found an unrealistic sample of 50M users who produced an average of 33 clicks a day (according to our 50-billion total monthly view). Because of this, we keep the more conservative estimation, based on Guess et al. (2021) data, of 8M internet users.

2 Results and model fits

Table 1: Correlation matrix of main variables

	1.	2.	3.	4.	5.	6.	7.
1. Direct traffic							
2. Search traffic	-0.66***						
3. Social traffic	-0.28***	-0.42***					
4. Bounce rate	-0.51***	0.25***	0.26***				
5. Visits (log)	0.01	0.09***	-0.13***	-0.27***			
6. Visit duration (log)	0.45***	-0.21***	-0.26***	-0.65***	0.39***		
7. Pages per visit (log)	0.37***	-0.12***	-0.27***	-0.75***	0.28***	0.75***	
8. Website quality	-0.29***	0.50***	-0.25***	0.07**	0.22***	0.02	0.04

Table 2: Website quality by ideological strength

	Website quality
ideological strength	$-.839\ (-.860,\ -.818)^{***}$
Observations	2,617
Adjusted R^2	.701
F Statistic	$6{,}137.837^{***}$ (df = 1; 2615)
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

2.1 Traffic modality by website type

Table 3: Traffic modality by website type

	Direct traffic	Search traffic	Social traffic
	(1)	(2)	(3)
type [pro-science]	$-.766\ (-.972,\ -.560)^{***}$	$1.225\ (.999,\ 1.450)^{***}$	$-.761\ (-.995,\ -.527)^{***}$
type [least-biased]	$-.558\ (-.714,\ -.401)^{***}$	$.827\ (.657,\ .998)^{***}$	$-.431\ (-.609,\ -.253)^{***}$
type [biased]	$-.372\ (-.513,\ -.231)^{***}$	$.370\ (.216,\ .525)^{***}$	$-.176\ (-.337,\ -.015)^{*}$
bounce rate	$-.535\ (-.574,\ -.496)^{***}$	$.271\ (.228,\ .313)^{***}$	$.262\ (.218,\ .306)^{***}$
total visits (log)	$-.111\ (-.151,\ -.072)^{***}$	$.134\ (.091,\ .177)^{***}$	$-.045\ (-.089,\ .0001)$
Observations	1,932	1,932	1,931
Adjusted R^2	.305	.168	.102
F Statistic	170.696^{***} (df = 5; 1926)	79.217^{***} (df = 5; 1926)	44.865^{***} (df = 5; 1925)
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$		

Table 4: Tukey multiple comparisons for direct traffic

	diff (%)	lwr	upr	p adj
pro-science-conspiracy	-21.17	-26.32	-16.02	< .001
least-biased-conspiracy	-13.37	-17.31	-9.43	< .001
biased-conspiracy	-10.52	-14.07	-6.97	< .001
least-biased-pro-science	7.80	3.37	12.22	< .001
biased-pro-science	10.65	6.57	14.73	< .001
biased-least-biased	2.85	0.47	5.23	0.011

Table 5: Tukey multiple comparisons for search traffic

	diff (%)	lwr	upr	p adj
pro-science-conspiracy	32.54	26.02	39.06	< .001
least-biased-conspiracy	20.94	15.95	25.93	< .001
biased-conspiracy	11.16	6.67	15.65	< .001
least-biased-pro-science	-11.60	-17.21	-6.00	< .001
biased-pro-science	-21.38	-26.55	-16.21	< .001
biased-least-biased	-9.78	-12.79	-6.76	< .001

Table 6: Tukey multiple comparisons for social traffic

	diff (%)	lwr	upr	p adj
pro-science-conspiracy	-10.12	-14.63	-5.61	< .001
least-biased-conspiracy	-6.19	-9.64	-2.74	< .001
biased-conspiracy	-2.14	-5.25	0.96	0.286
least-biased-pro-science	3.93	0.05	7.81	0.046
biased-pro-science	7.98	4.40	11.55	< .001
biased-least-biased	4.05	1.96	6.14	< .001

2.2 Traffic modality by ideological strength

Table 7: Traffic modality by ideological strength

	Direct traffic	Search traffic	Social traffic
	(1)	(2)	(3)
strength	.143 (.120, .166)***	$-$.301 ($-$.324, $-$.278)***	.186 (.160, .211)***
bounce rate	$-$.522 ($-$.560, $-$.484)***	.232 (.193, .270)***	.286 (.243, .329)***
total visits (log)	$-$.080 ($-$.119, $-$.041)***	.049 (.010, .088)*	.009 ($-$.035, .052)
Observations	1,932	1,932	1,931
Adjusted R^2	.333	.322	.162
F Statistic	322.992*** (df = 3; 1928)	307.069*** (df = 3; 1928)	125.547*** (df = 3; 1927)
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$		

Table 8: Traffic modality by political lean

	Direct traffic	Search traffic	Social traffic
	(1)	(2)	(3)
political lean [conservative]	.323 (.228, .419)***	$-$.550 ($-$.654, $-$.446)***	.123 (.010, .237)*
bounce rate	$-$.555 ($-$.606, $-$.505)***	.256 (.201, .310)***	.279 (.219, .338)***
total visits (log)	$-$.087 ($-$.135, $-$.038)***	.098 (.045, .151)***	$-$.061 ($-$.119, $-$.004)*
Observations	1,263	1,263	1,263
Adjusted R^2	.305	.147	.083
F Statistic (df = 3; 1259)	186.027***	73.439***	39.170***
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$		

2.3 Traffic modality and ideological strength by website type

Table 9: Traffic modality and ideological strength by website type

	Direct traffic	Search traffic	Social traffic
	(1)	(2)	(3)
strength	.549 (.381, .717)***	−.551 (−.716, −.385)***	−.041 (−.239, .156)
political lean [liberal]	−.183 (−.368, .002)	.292 (.110, .475)**	−.241 (−.458, −.023)*
political lean [conservative]	.039 (−.149, .227)	.008 (−.177, .193)	−.315 (−.535, −.094)**
bounce rate	−.537 (−.584, −.491)***	.193 (.148, .239)***	.346 (.292, .401)***
total visits (log)	−.052 (−.100, −.005)*	−.012 (−.059, .035)	.040 (−.016, .096)
strength:liberal	−.485 (−.677, −.292)***	.103 (−.087, .293)	.609 (.383, .835)***
strength:conservative	−.265 (−.449, −.081)**	−.083 (−.264, .099)	.420 (.204, .636)***
Observations	1,417	1,417	1,417
Adjusted R^2	.368	.334	.185
F Statistic (df = 7; 1409)	118.838***	102.625***	46.948***
Note:	*p<0.05; **p<0.01; ***p<0.001		

Figure 1: Traffic modality and ideological strength by website type

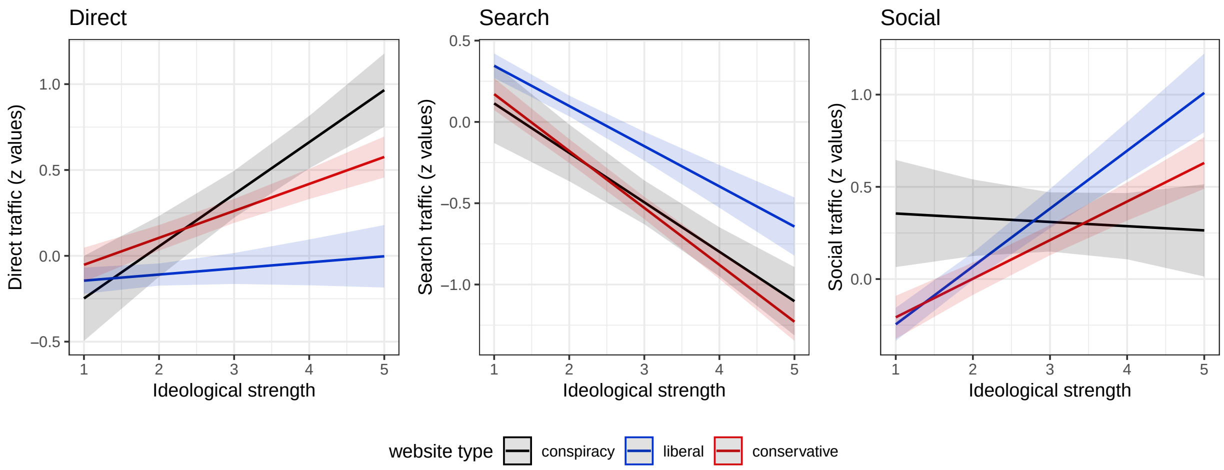


Table 10: Direct traffic and ideological strength by website type (single slopes)

lean2	bias.trend	SE	df	t.ratio	p.value
conspiracy	0.3033	0.0475	1409	6.392	<.0001
liberal	0.0354	0.0272	1409	1.305	0.1921
conservative	0.1570	0.0217	1409	7.241	<.0001

Table 11: Direct traffic and ideological strength by website type (slope comparison)

contrast	estimate	SE	df	t.ratio	p.value
conspiracy - liberal	0.2679	0.0543	1409	4.936	<.0001
conspiracy - conservative	0.1463	0.0519	1409	2.819	0.0135
liberal - conservative	-0.1216	0.0343	1409	-3.548	0.0012

P value adjustment: tukey method for comparing a family of 3 estimates

Table 12: Search traffic and ideological strength by website type (single slopes)

lean2	bias.trend	SE	df	t.ratio	p.value
conspiracy	-0.3042	0.0468	1409	-6.505	<.0001
liberal	-0.2473	0.0268	1409	-9.238	<.0001
conservative	-0.3499	0.0214	1409	-16.376	<.0001

Table 13: Search traffic and ideological strength by website type (slope comparison)

contrast	estimate	SE	df	t.ratio	p.value
conspiracy - liberal	-0.0569	0.0535	1409	-1.064	0.5365
conspiracy - conservative	0.0458	0.0512	1409	0.895	0.6437
liberal - conservative	0.1027	0.0338	1409	3.041	0.0068

P value adjustment: tukey method for comparing a family of 3 estimates

Table 14: Social traffic and ideological strength by website type (single slopes)

lean2	bias.trend	SE	df	t.ratio	p.value
conspiracy	-0.0229	0.0557	1409	-0.411	0.6809
liberal	0.3138	0.0319	1409	9.847	<.0001
conservative	0.2093	0.0254	1409	8.227	<.0001

Table 15: Social traffic and ideological strength by website type (slope comparison)

contrast	estimate	SE	df	t.ratio	p.value
conspiracy - liberal	-0.3367	0.0637	1409	-5.287	<.0001
conspiracy - conservative	-0.2322	0.0609	1409	-3.813	0.0004
liberal - conservative	0.1045	0.0402	1409	2.599	0.0256

P value adjustment: tukey method for comparing a family of 3 estimates

2.4 Individual (vs group) traffic

Table 16: Individual traffic by website type

	Individual traffic
ideological strength	.305 (.118, .491)**
political lean [liberal]	−.015 (−.220, .190)
political lean [conservative]	.059 (−.150, .267)
bounce rate	−.565 (−.616, −.513)***
total visits (log)	−.078 (−.131, −.025)**
strength:liberal	−.559 (−.772, −.345)***
strength:conservative	−.412 (−.616, −.208)***
Observations	1,417
Adjusted R^2	.280
F Statistic	79.727*** (df = 7; 1409)
Note:	*p<0.05; **p<0.01; ***p<0.001

Figure 2: Individual traffic by website type

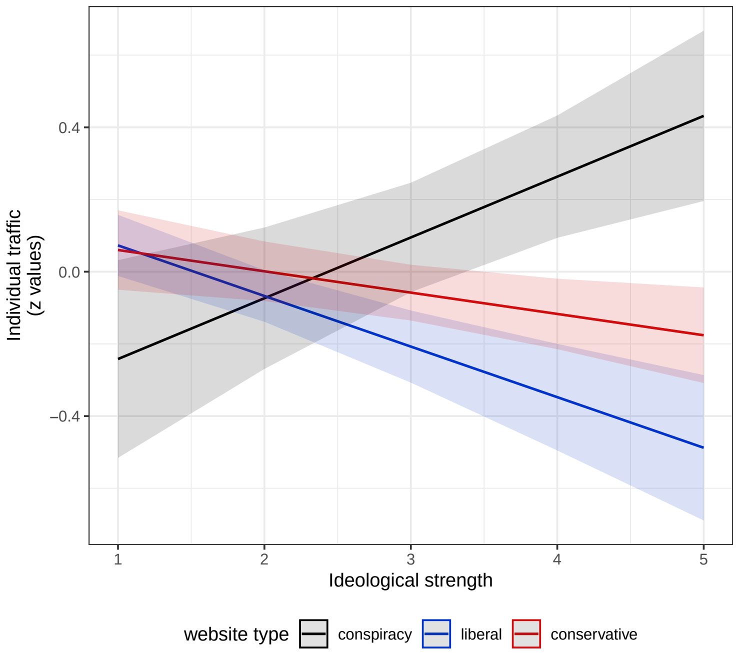


Table 17: Individual traffic by website type (single slopes)

lean2	bias.trend	SE	df	t.ratio	p.value
conspiracy	0.1684	0.0525	1409	3.206	0.0014
liberal	-0.1402	0.0301	1409	-4.661	<.0001
conservative	-0.0590	0.0240	1409	-2.459	0.0140

Table 18: Individual traffic by website type (slope comparison)

contrast	estimate	SE	df	t.ratio	p.value
conspiracy - liberal	0.3086	0.0601	1409	5.135	<.0001
conspiracy - conservative	0.2275	0.0575	1409	3.958	0.0002
liberal - conservative	-0.0811	0.0379	1409	-2.139	0.0826

P value adjustment: tukey method for comparing a family of 3 estimates

2.5 Traffic modality by website quality

Table 19: Traffic modality by website quality

	Direct traffic	Search traffic	Social traffic
	(1)	(2)	(3)
website quality	−.242 (−.283, −.201)***	.494 (.452, .536)***	−.289 (−.335, −.243)***
bounce rate	−.523 (−.562, −.484)***	.231 (.191, .271)***	.288 (.243, .332)***
total visits (log)	−.083 (−.123, −.042)***	.054 (.013, .095)*	.003 (−.042, .049)
Observations	1,860	1,860	1,859
Adjusted R^2	.330	.292	.142
F Statistic	306.815*** (df = 3; 1856)	256.825*** (df = 3; 1856)	103.785*** (df = 3; 1855)
Note:	*p<0.05; **p<0.01; ***p<0.001		

2.6 Traffic modality by website quality (replication with IRMA)

Table 20: Traffic modality by website quality (replication with IRMA)

	Direct traffic	Search traffic	Social traffic
	(1)	(2)	(3)
website quality	−.388 (−.632, −.145)**	.568 (.336, .800)***	−.154 (−.441, .134)
bounce rate	−.400 (−.669, −.130)**	.162 (−.095, .419)	.349 (.031, .666)*
total visits (log)	.227 (−.102, .556)	−.456 (−.769, −.142)**	.271 (−.117, .659)
Observations	49	49	49
Adjusted R^2	.314	.377	.046
F Statistic (df = 3; 45)	8.334***	10.686***	1.773
Note:	*p<0.05; **p<0.01; ***p<0.001		

2.7 Website class proportion

Table 21: Website proportion

Website class	N websites	Sum of total visits	Average visits	Proportion of visits (%)
CT01	38	62,979,600	2,332,578	3.04
CT02	34	19,177,860	799,078	1.04
CT03	52	20,164,700	806,588	1.05
CT04	70	66,789,060	1,757,607	2.29
CT05	101	30,921,710	773,043	1.01
R05	335	291,161,590	1,831,205	2.39
R03	290	2,098,226,550	11,721,936	15.28
R01	296	2,542,609,030	10,463,412	13.64
LB00	476	2,521,031,570	6,271,223	8.18
PS00	148	1,644,573,870	14,553,751	18.97
L01	562	7,515,027,480	16,337,016	21.30
L03	264	1,128,097,550	6,302,221	8.22
L05	96	115,657,299	2,753,745	3.59

2.8 Explorative plots

We provide descriptive figures for all our variables.

- **Plot A):** differences between website types (x axis) on averaged values for variable of interest (y axis). Bars represent the standard error of the mean). PS00: pro-science; LB00: least-biased; LR: biased (including classes L01, L03, L05, R01, R03, R05); CT: conspiracy (including classes CT01, CT02, CT03, CT04, and CT05).
- **Plot B):** scatter plot of website quality (x axis) by variable of interest (y axis). Fitted slope (with 95% CIs shade) and value reported represent Pearson correlation.
- **Plot C):** differences between conspiracy classes (x axis) on averaged values for variable of interest (y axis). Bars represent the standard error of the mean). Fitted slope (with 95% CIs shade) and value reported represent Pearson correlation.
- **Plot D):** differences between liberal (L, including classes L01, L03, L05) and conservative (R, including classes R01, R03, R05) websites (x axis) on averaged values for variable of interest (y axis). Bars represent the standard error of the mean.
- **Plot E):** two-way interaction plot of ideological strength (x axis) and ideological type [red: conservative (including classes R01, R03, R05); blue: liberal (including classes L01, L03, L05)] on averaged values for variable of interest (y axis). Bars represent the standard error of the mean).
- **Plot F):** two-way interaction plot's fitted slopes of ideological strength (x axis) and ideological type [red: conservative (including classes R01, R03, R05); blue: liberal (including classes L01, L03, L05); black: conspiracy (including classes CT01, CT02, CT03, CT04, and CT05)] on averaged values for variable of interest (y axis). Shaded ares represent the 95% CIs.

Figure 3: Explorative plot - Direct traffic

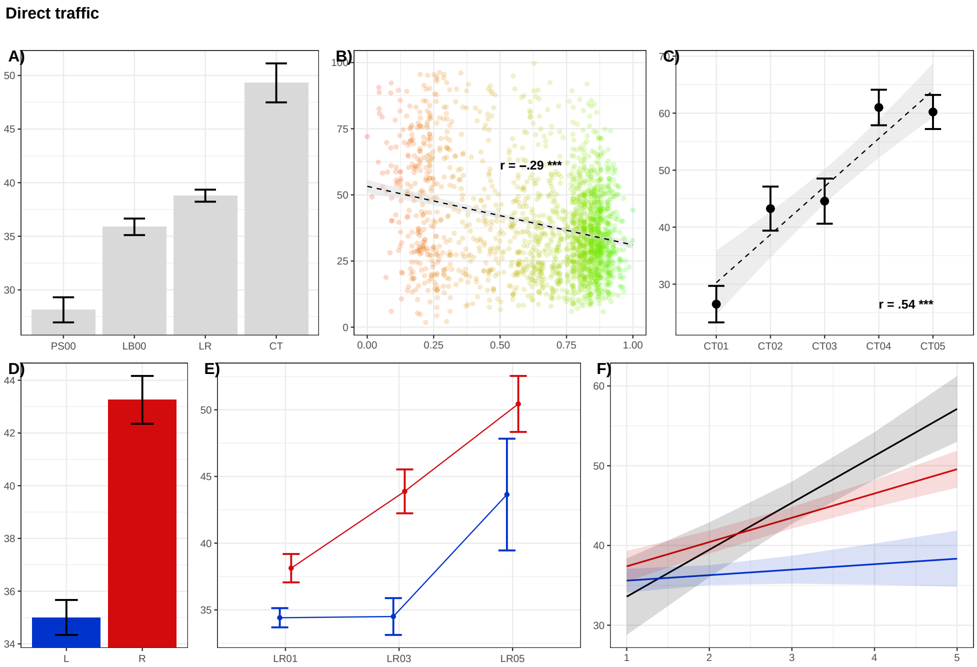


Figure 4: Explorative plot - Search traffic

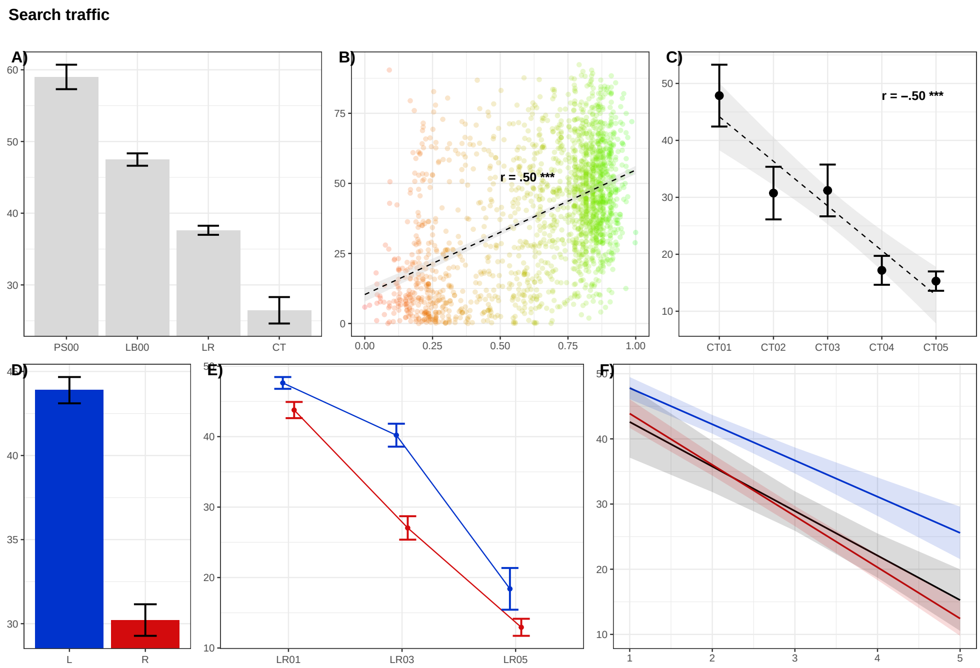


Figure 5: Explorative plot - Social traffic

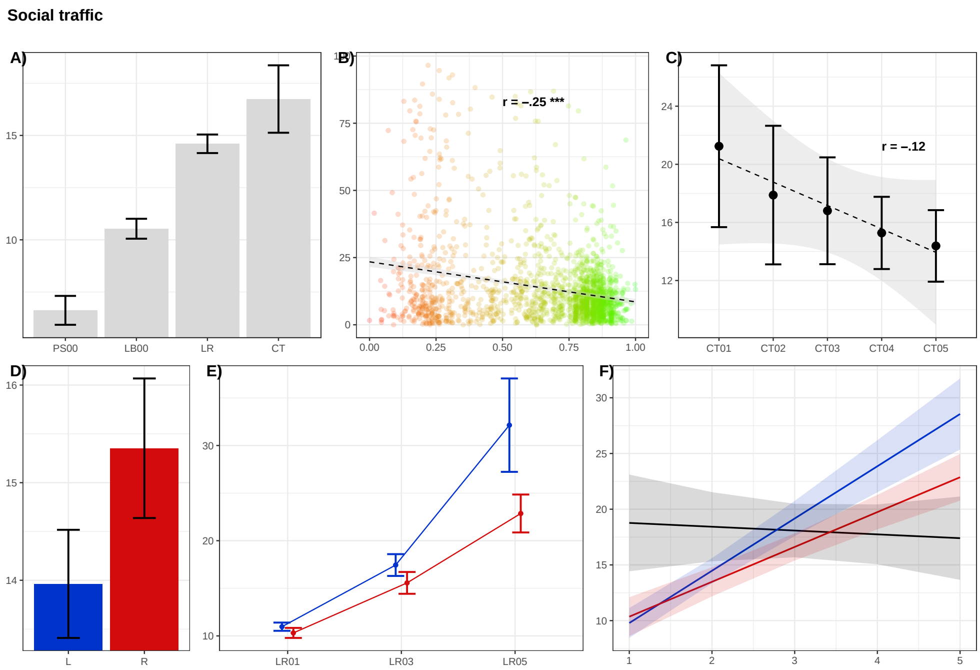


Figure 6: Explorative plot - Visits (log)

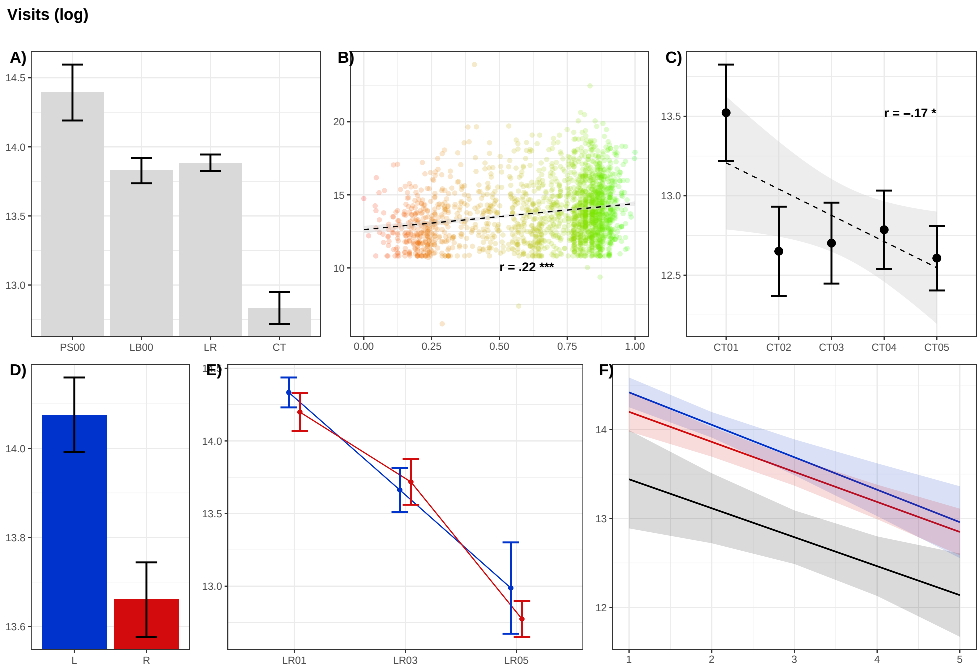


Figure 7: Explorative plot - Bounce rate

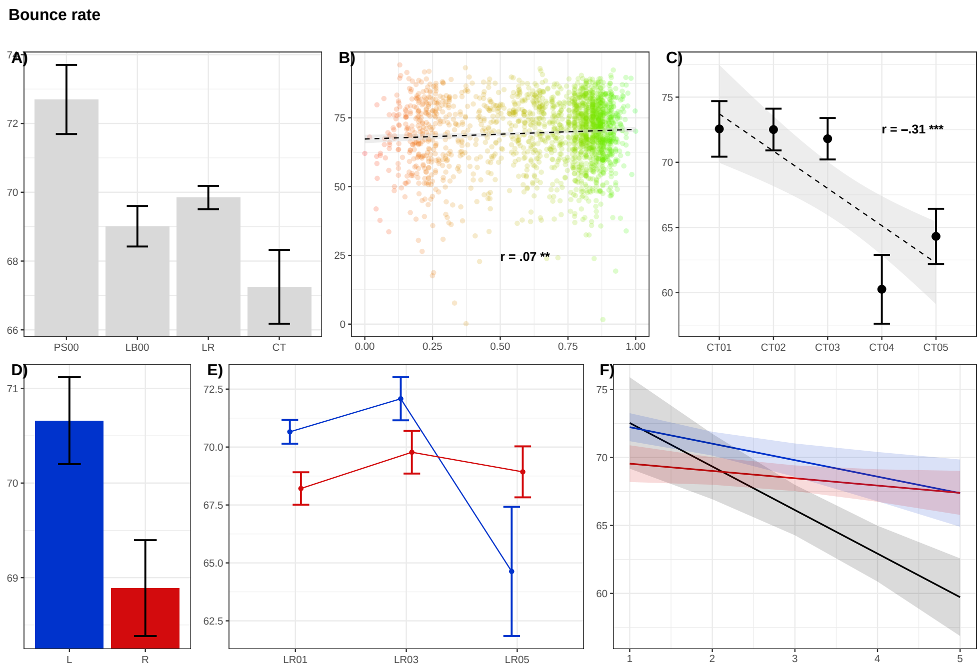


Figure 8: Explorative plot - Traffic from referreals

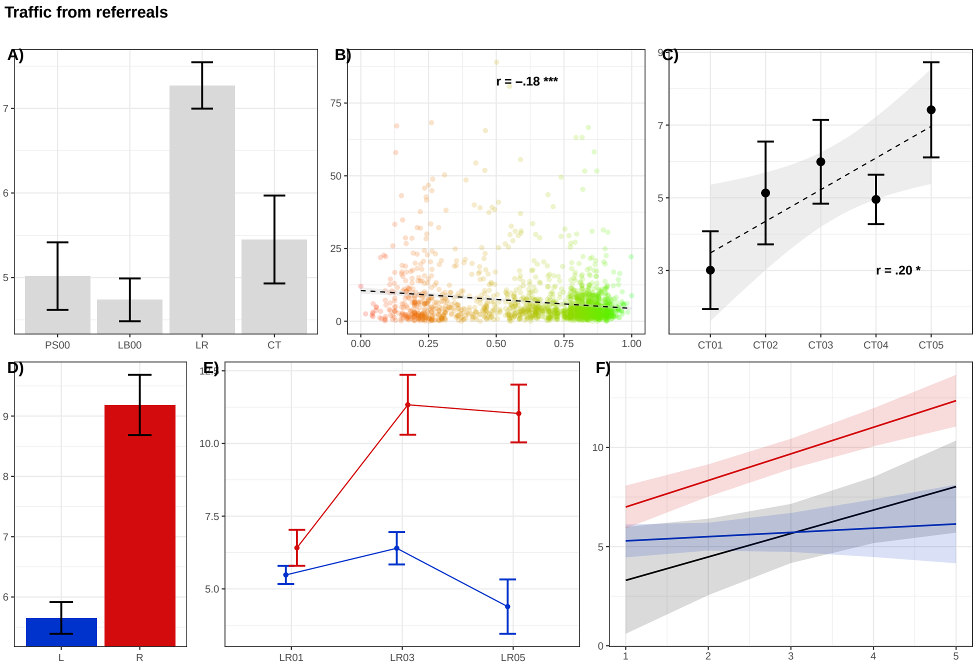


Figure 9: Explorative plot - Traffic from mail

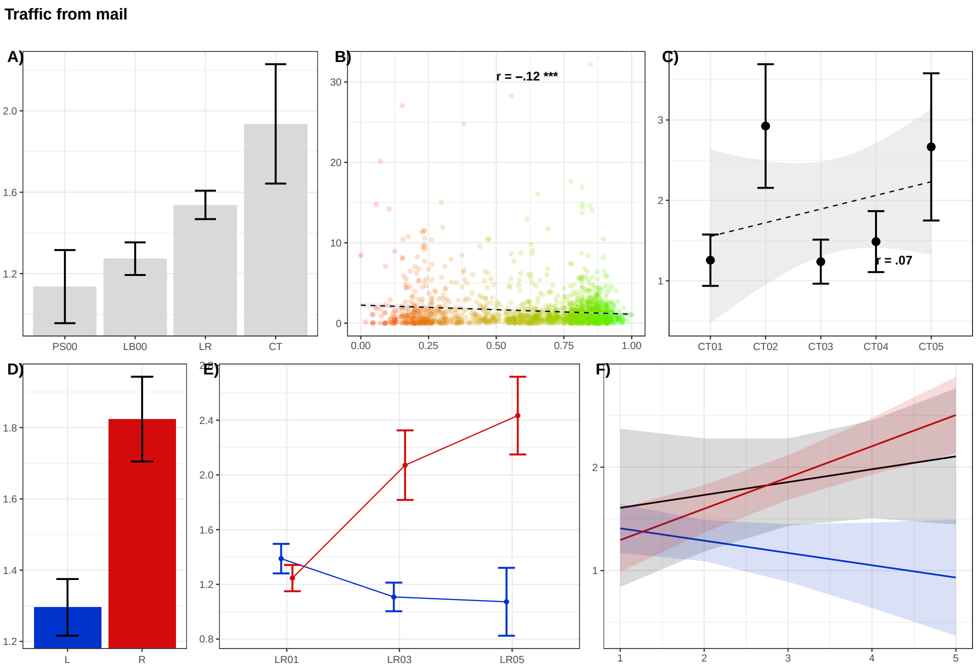


Figure 10: Explorative plot - Visit duration (log)

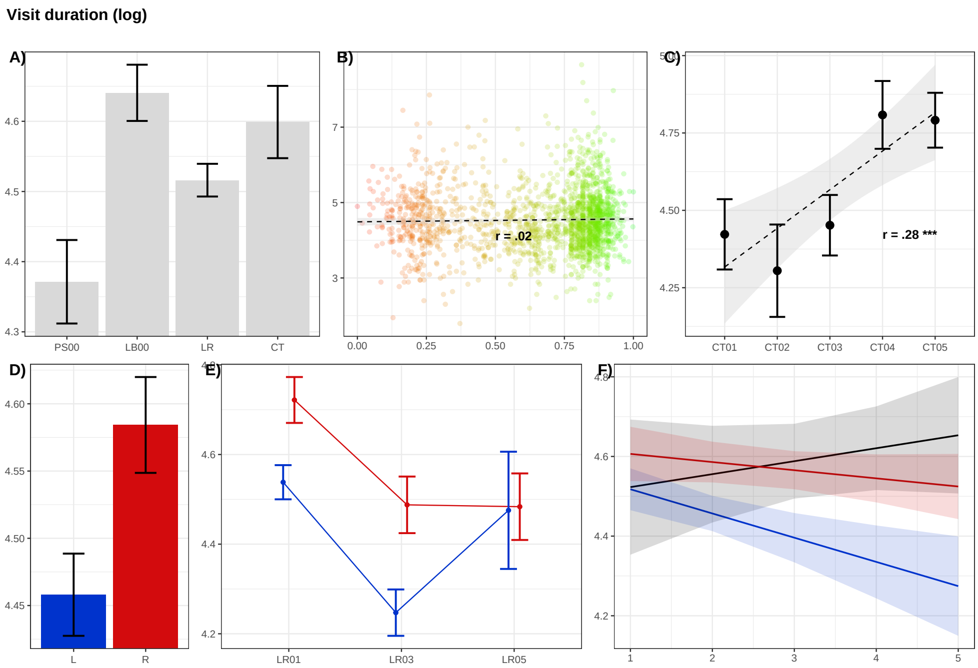


Figure 11: Explorative plot - Pages per visit (log)

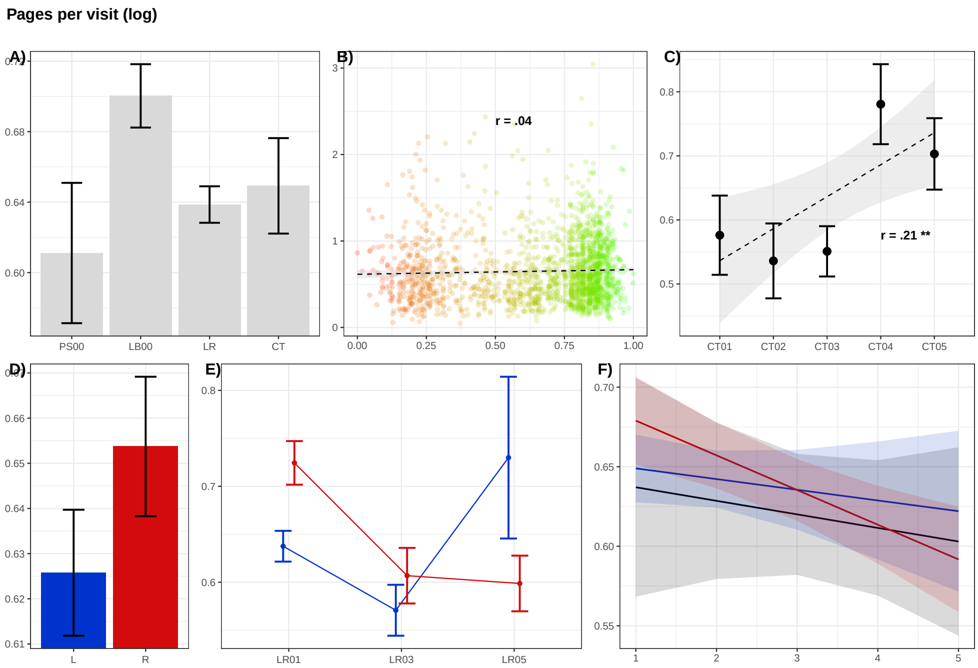


Figure 12: Explorative plot - Traffic from social (Reddit)

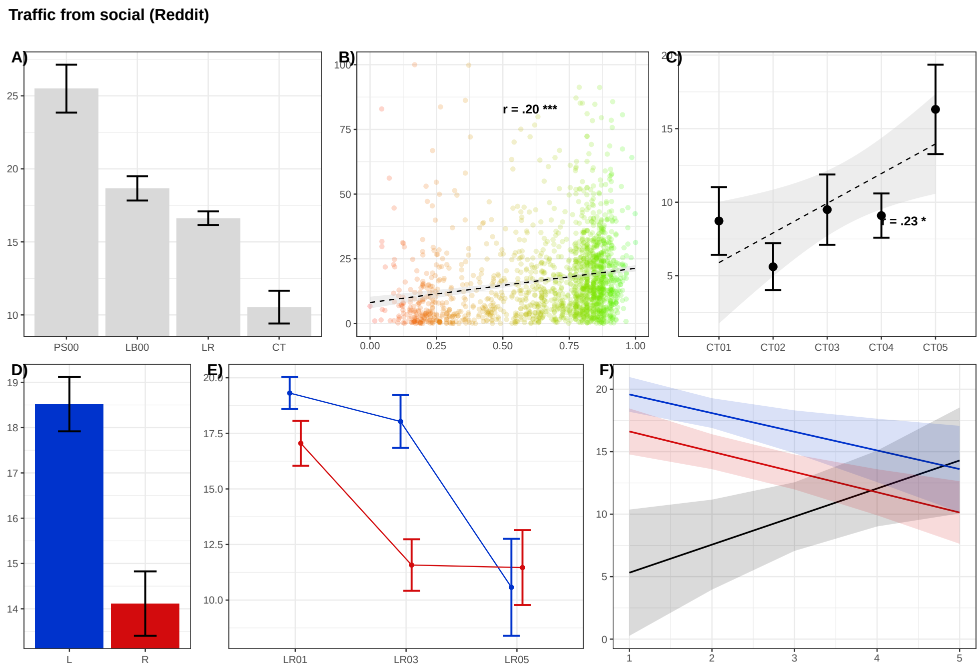


Figure 13: Explorative plot - Traffic from social (YouTube)

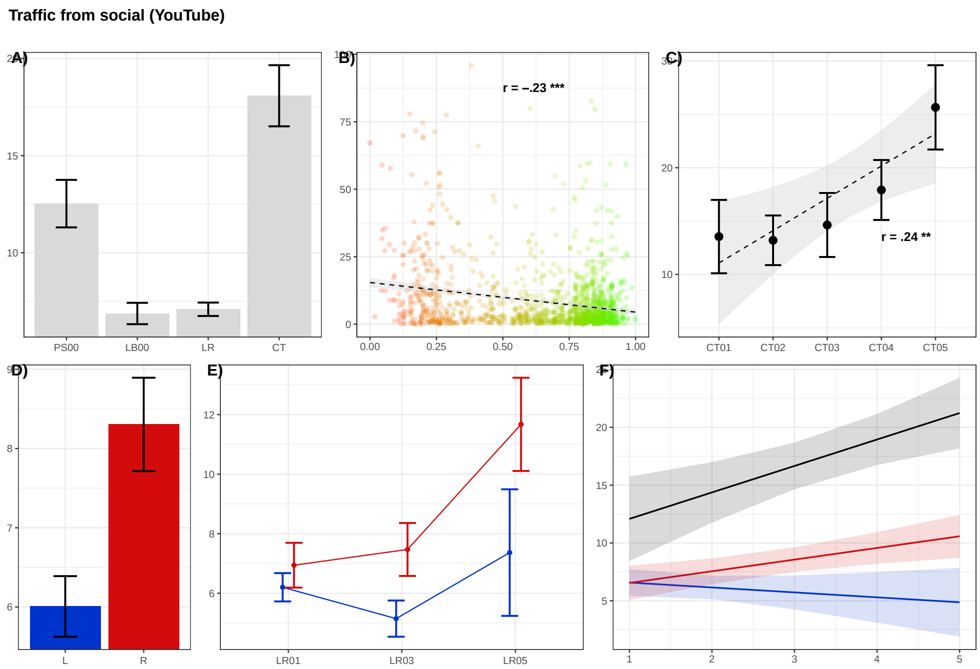


Figure 14: Explorative plot - Traffic from social (Facebook)

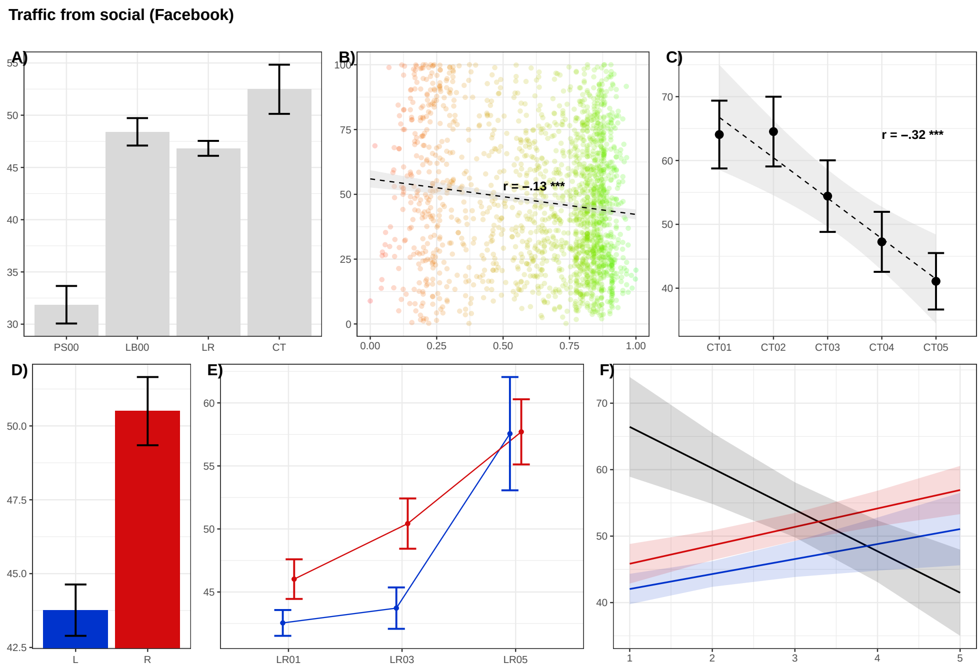
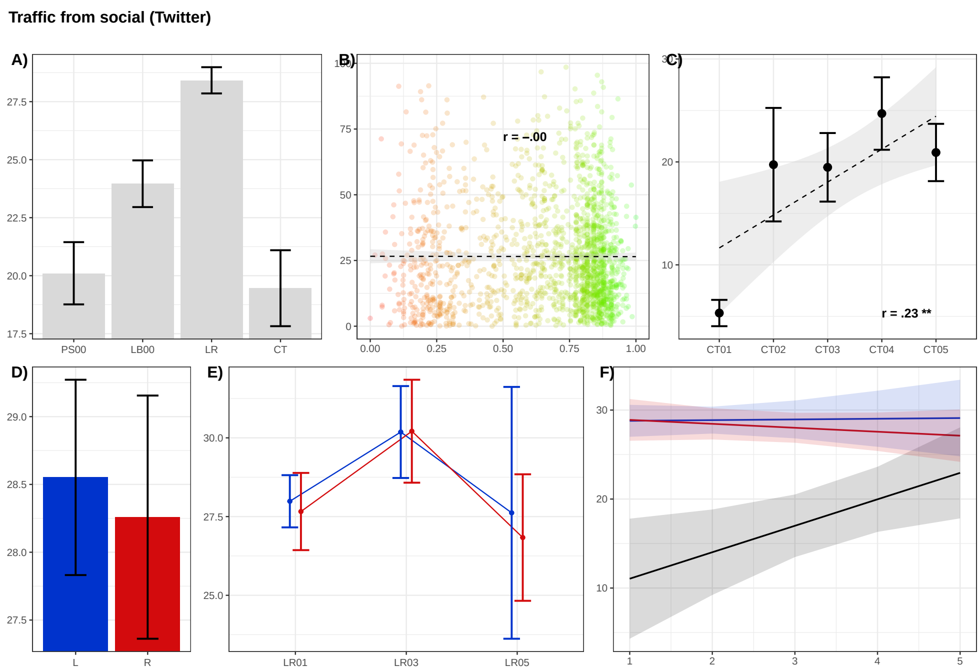


Figure 15: Explorative plot - Traffic from social (Twitter)



References

Albarracin, Dolores, Julia Albarracin, Man-pui Sally Chan, and Kathleen Hall Jamieson. 2021. *Creating Conspiracy Beliefs: How Our Thoughts Are Shaped*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108990936.

Baly, Ramy, Georgi Karadzhov, Dimitar Alexandrov, James Glass, and Preslav Nakov. 2018. "Predicting Factuality of Reporting and Bias of News Media Sources." In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 3528–39. Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-1389.

Bulck, H. Van den, and A. Hyzen. 2020. "Of Lizards and Ideological Entrepreneurs: Alex Jones and Infowars in the Relationship Between Populist Nationalism and the Post-Global Media Ecology." *International Communication Gazette* 82 (1): 42–59. https://doi.org/10.1177/1748048519880726.

Cardenal, Ana S., Carlos Aguilar-Paredes, Carol Galais, and Mario Pérez-Montoro. 2019. "Digital Technologies and Selective Exposure: How Choice and Filter Bubbles Shape News Media Exposure." *The International Journal of Press/Politics* 24 (4): 465–86. https://doi.org/10.1177/1940161219862988.

Gentzkow, Matthew, and Jesse M. Shapiro. 2011. "Ideological Segregation Online and Offline *." *The Quarterly Journal of Economics* 126 (4): 1799–839. https://doi.org/10.1093/qje/qjr044.

Goel, Sharad, Jake Hofman, and M. Sirer. 2021. "Who Does What on the Web: A Large-Scale Study of Browsing Behavior." *Proceedings of the International AAAI Conference on Web and Social Media* 6 (1): 130–37.

Guess, Andrew, Pablo Barberá, Simon Munzert, and JungHwan Yang. 2021. "The Consequences of Online Partisan Media." *Proceedings of the National Academy of Sciences* 118 (14): e2013464118. https://doi.org/10.1073/pnas.2013464118.

Jansen, Bernard J., Danielle L. Booth, and Amanda Spink. 2008. "Determining the Informational, Navigational, and Transactional Intent of Web Queries." *Information Processing & Management* 44 (3): 1251–66. https://doi.org/10.1016/j.ipm.2007.07.015.

Lin, Hause, Jana Lasser, Stephan Lewandowsky, Rocky Cole, Andrew Gully, David Gertler Rand, and Gordon Pennycook. 2022. "High Level of Agreement Across Different News Domain Quality Ratings." Preprint. PsyArXiv. https://doi.org/10.31234/osf.io/qy94s.

Miani, Alessandro, Thomas Hills, and Adrian Bangerter. 2021. "LOCO: The 88-Million-Word Language of Conspiracy Corpus." *Behavior Research Methods* 54 (4): 1794–1817. https://doi.org/10.3758/s13428-021-01698-z.

Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. penguin UK.

Pennycook, Gordon, and David G. Rand. 2019. "Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality." *Proceedings of the National Academy of Sciences* 116 (7): 2521–26. https://doi.org/10.1073/pnas.1806781116.

Risius, Marten, Okan Aydinguel, and Maximilian Haug. 2019. "Towards an Understanding of Conspiracy Echo Chambers on Facebook." In *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Romer, Daniel, and Kathleen Hall Jamieson. 2021. "Conspiratorial Thinking, Selective Exposure to Conservative Media, and Response to COVID-19 in the US." *Social Science & Medicine* 291 (December): 114480. https://doi.org/10.1016/j.socscimed.2021.114480.

Springborn, Kevin, and Paul Barford. 2013. "Impression Fraud in On-line Advertising via ${$Pay-Per-View$}$ Networks." In *22nd USENIX Security Symposium (USENIX Security 13)*, 211–26.

Sunstein, Cass. 2018. *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press. https://doi.org/10.1515/9781400890521.

Urman, Aleksandra, Mykola Makhortykh, Roberto Ulloa, and Juhi Kulshrestha. 2022. "Where the Earth Is Flat and 9/11 Is an Inside Job: A Comparative Algorithm Audit of Conspiratorial Information in Web Search Results." *Telematics and Informatics* 72 (August): 101860. https://doi.org/10.1016/j.tele.2022.101860.

Table of (literal) content

Table 5.1: Top 60 lemmatized features (excluding stop words) used in this thesis and their frequency

feature	frequency	feature	frequency
conspiracy	1624	numb	127
theory	779	information	123
document	545	within	123
topic	483	individual	123
website	447	journal	121
social	350	share	119
belief	292	lexical	118
traffic	267	narrative	117
word	266	content	116
text	261	search	115
use	208	show	113
high	188	different	112
measure	180	seed	112
bias	177	one	111
corpus	166	table	111
psychology	162	network	110
science	158	non	108
compound	158	people	106
loco	156	group	104
value	151	section	104
spread	150	low	104
political	150	model	103
think	149	provide	99
medium	148	difference	98
analysis	141	datum	98
effect	139	online	97
extract	138	douglas	94
language	137	set	93
research	136	source	93
test	133	creativity	93