Chapter 15
Combatting Conspiratorial Thinking with Controlled Argumentation Dialogue Environments

Lindsay Fields and John Licato

Abstract The COVID-19 pandemic has been associated with an explosion in misinformation, leading to increased interest in methods to combat the failures in critical thinking which make such misinformation so powerful. In combatting misinformation, simply throwing uncontrolled argumentation at the problem is often counterproductive, partially because the means by which people evaluate arguments are highly subject to cognitive biases. Such biases which promote jumping to unwarranted conclusions have been shown to correlate with conspiratorial belief. We consider the use of Controlled Argumentation Dialogue Environments (CADEs) as a means to mitigate cognitive biases which contribute to belief in COVID-19 conspiracy theories. We will discuss *Warrant Game* (WG) and *Warrant Game for Analogies* (WG-A), CADEs in which two arguers are presented with a divisive issue and two competing positions on that issue. They then compete by iteratively improving warrants for their arguments and attacking those of their opponents. The warrant, when made explicit, makes it easier to determine key features typically associated with argument strength and may reveal hidden assumptions or fundamental reasoning incompatibilities. By presenting an issue and positions which relate to conspiratorial thinking, CADEs may operate as an educational tool for breaking conspiratorial belief into core values and building cognitive skills.

Keywords Conspiracy theories · Controlled argumentation · COVID-19

15.1 Introduction

The ongoing COVID-19 pandemic has led to a major proliferation of misinformation and disinformation (The Lancet, 2020; World Health Organization et al., 2020), and

L. Fields (✉) · J. Licato
Department of Computer Science and Engineering, Advancing Machine and Human Reasoning (AMHR) Lab, University of South Florida, Tampa, Florida, USA
e-mail: ldfields@usf.edu

J. Licato
e-mail: licato@usf.edu

S. Oswald et al. (eds.), *The Pandemic of Argumentation*, Argumentation Library 43,
https://doi.org/10.1007/978-3-030-91017-4_15

this has led to increased interest in methods to combat the failures in critical thinking which make them so powerful. In combatting misinformation, simply inviting people to engage in unstructured argumentation can be counter-productive, in part because the means by which people evaluate arguments are highly subject to cognitive biases, of which the arguers may be unaware. Some of these biases promote jumping to unwarranted conclusions, as opposed to methodical inference, and have been shown to correlate with conspiratorial belief (Denovan et al., 2020; Pytlik et al., 2020; Swami et al., 2014). In this paper, we consider the use of Controlled Argumentation Dialogue Environments (CADEs) as a means to mitigate those cognitive biases which contribute to belief in COVID-19 conspiracy theories. We will show how CADEs can restrict the structure of allowed arguments, such that they will tend to disallow argument patterns which are hallmarks of conspiratorial thinking. In this paper, we focus on showing how two specific CADEs (WG and WG-A) can perform this function, and leave the task of empirically studying the persuasive powers of such argumentation environments to future work.

In this work, we will be using Keeley's (1999) definition of a *conspiracy theory* as a "proposed explanation of some historical event (or events) in terms of the significant causal agency of a relatively small group of persons—the conspirators—acting in secret" (p. 116). This definition does not preclude the possibility that such a theory may be accurate and, as will be discussed in Sect. 15.6, CADEs tolerate this possibility and may allow arguers to refine and strengthen a valid argument. An *argumentative dialogue* is a dialogue whose intended purpose includes to exchange, evaluate, communicate, or otherwise address at least one argument. Acceptance of an argument may consist of either: (1) accepting that the conclusion is true; or (2) accepting that the conclusion would follow from the premises, but not necessarily that the premises are true. *Controlled Argumentation Dialogue Environments* (CADEs) are frameworks for argumentative dialogues which are highly structured, restrictive of the communications allowed between participants, and may be supervised by either a human or an artificially intelligent moderator.

Argumentation is inextricably entwined with persuasion; perhaps the most common purpose of argumentation is to trigger belief change in the self or others. Unstructured argumentation, however, is often unsatisfactory in promoting belief change amongst arguers, in part due to the detrimental impacts of cognitive biases (Kahneman, 2011; Mercier, 2016; Mercier & Sperber, 2011, 2017; Stanovich & West, 2007). Further, emotionality may exacerbate the expression of cognitive biases in argumentation dialogue environments. On social media platforms, for example, certain types of emotion-provoking content are known to attract higher user engagement, or "clicks," and make content more viral (Berger & Milkman, 2012; Brady et al., 2017; Chen, 2020; Ferrara & Yang, 2015; Ksiazek, 2016). Many existing studies on the relationship between emotionality and environment conclude that limiting external influence on the environment and re-framing emotional topics may effectively address the maladaptive effects of emotionality (Choi et al., 2018; Richards & Gross, 2000). However, it is naïve to assume that any such controlled dialogues can be maintained within a social media platform which prioritizes engagement over discussion quality. Instead, we seek to mitigate conspiratorial belief in susceptible

individuals prior to their entry into emotional environments, thereby enabling them to recognize and lessen any harmful persuasive effects of misinformation.

Previous work suggests that CADEs specifically designed to optimize persuasion and minimize emotionality have promise in mitigating the effects of cognitive bias (Cooper et al., 2020). This may be due to Oswald's (2016) belief that if arguers are influenced by cognitively biased inferences, then traces of those biases should become evident in the argumentative discourse. It is also possible that, due to the inherently emotional nature of conspiracy belief, CADEs may facilitate emotionality enhanced memory retention, which has been previously shown to decrease the rate of memory decay in cognitive skills tasks (Steidl et al., 2011). In either case, there is cause to extrapolate that introducing controlled argumentation techniques to COVID-19 conspiracy forums, which are particularly susceptible to both biased and emotional appeals, could build resistance to misinformation.

Existing argumentation methods have shown *cognitive inoculation* effects, whereby participants identify and build resistance against social media misinformation (Basol et al., 2020; Roozenbeek & van der Linden, 2018, 2019); participants were able to maintain this resistance for up to five weeks following the initial inoculation (Roozenbeek, 2020). Further research has also considered the "prebunking" effects of a priori inoculative intervention against conspiracy theories, in particular (Cook et al., 2017; van der Linden et al., 2017). When individuals were presented with both a scientific consensus and misinformation casting doubt on said consensus, their previously-held beliefs saw no significant change (van der Linden et al., 2017), implying that simply presenting susceptible parties with accurate information may be sufficient to combat conspiracy belief. Further, it was found that false-balance media coverage had the greatest impact on perceived consensus, but prebunking which specifically targeted false neutrality had the greatest influence on neutralizing misinformation (Cook et al., 2017). CADEs have the potential to create such prebunking effects dependent on the strategies employed by each arguer. Additionally, the nature of two-party argumentation and the selection of controversial theories removes the potential for false neutrality in the inoculative intervention.

We will discuss the recently designed *Warrant Game* (WG) and its successor *Warrant Game for Analogies* (WG-A), CADEs in which two arguers are presented with a divisive issue and two competing positions on that issue. The arguers then compete by iteratively improving warrants for their arguments and attacking those of their opponents. Here the warrant, drawn from Toulmin et al. (1984) influential model of argumentation, is part of an argument which, when made explicit, makes it easier to determine key features typically associated with argument strength and may reveal hidden assumptions or fundamental reasoning incompatibilities. By presenting an issue and positions which relate to conspiratorial thinking (e.g., that COVID-19 was created by the Chinese government as a biological weapon), these CADEs may operate as educational tools for mitigating conspiratorial belief, identifying core values and biases, and building cognitive skills.

15.2 Known COVID-19 Conspiracies

Multiple common conspiracy theories related to COVID-19 have emerged since February 2020 (The Lancet, 2020; Prichard & Christman, 2020; Romer & Jamieson, 2020; Roozenbeek et al., 2020; van der Linden et al., 2020). The nature of these conspiracies has changed over time, as has the scientific consensus regarding the health risks, best practices, and lasting effects of the virus on the public and the world economy. For example, at the beginning of the pandemic, most conspiracies revolved around the existence of the virus and the veracity of expert recommendations (Romer & Jamieson, 2020). However, new developments, such as the introduction and rapid distribution of vaccines, have led to the emergence of entirely new conspiracy theories, which were not present in early 2020 (Brenan, 2021; COVID Collaborative, 2020). We will discuss a few of the more prevalent theories. In many of these theories, the flaw in reasoning seems to be caused by the perceived frequent shifting of scientific consensus (which in turn was a consequence of the novelty of COVID-19). This appears to invite susceptible parties to engage in ad hominem attacks on the source of the consensus and ignore evidence that, on a cursory glance, seems contradictory. We will discuss these, and other, common features of COVID-19 conspiracy theories, as well as how their influence can be reduced by CADEs, in Sect. 15.3.

Almost immediately following the declaration of COVID-19 as a pandemic, conspiracy theories emerged that the virus was bioengineered in China. A survey performed in May 2020 indicated that 23% of Americans believed that the idea that the virus was engineered in a laboratory in Wuhan was "reliable" (Roozenbeek et al., 2020). This is notable, as individuals who are susceptible to believing that the virus was bioengineered are reported to be less likely to comply with social distancing and masking guidelines and less likely to accept a COVID-19 vaccine (Prichard & Christman, 2020; van der Linden et al., 2020). The proliferation of this theory was exacerbated in September 2020 with the release of the controversial "Yan Report," a preprint of a study which claimed that the genome composition of SARS-CoV-2 implied that the virus was man-made. This study has been repeatedly debunked (Koyama et al., 2020; Rassmussen, 2021), however the theory has since been considered by multiple well-known scientists, including Nobel laureate Luc Montagnier (Clavel, 2020) and immunologist Anthony Fauci (Brewster, 2021) leading many conspiratorial thinkers to believe that COVID-19 is a bioweapon and its release is being covered up.

One conspiracy theory that should be noted for its cultural and ethnographic implications is the belief that the COVID-19 vaccines being introduced are used to subversively test on Black Americans. A survey conducted by the National Association for the Advancement of Colored People (NAACP) reported that 71% of Black Americans believed the vaccines were not adequately tested on Black people prior to their launch and 80% had concerns that they would receive less-safe versions of the vaccine (COVID Collaborative, 2020). This theory is the latest in a history of distrust, paranoia, and conspiratorial thinking within the Black community in regard to medical

treatment and, particularly, vaccination. In light of historical events involving the Tuskegee syphilis study (National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, 2020) and the HeLa cell line (Skloot, 2000, 2011), among others, many Black Americans are fearful of participating in front-line medical treatment. The same NAACP survey reported 81% of Black Americans intended to wait to receive the vaccine, as opposed to getting it as soon as it was available, and that 87% believed that earlier versions of the vaccine were less effective than later versions. Distrust in medical authority is further exacerbated by continuing racial disparities in the American healthcare system which directly contribute to higher rates of infant and maternal mortality, complications of chronic conditions, and lower life expectancies in Black Americans compared to their White counterparts (Bajaj & Stanford, 2021). This results in a community-wide conspiratorial environment where only 14% of Black Americans report confidence in the vaccine's safety and only 18% had an intention to be vaccinated in September 2020 (COVID Collaborative, 2020), considerably less than the 58.9% (Romer & Jamieson, 2020) and 65% (Brenan, 2021) of Americans who indicated willingness in March 2020 and December 2020, respectively. As of June 2021, only 9% of individuals receiving at least one dose of the vaccine were Black (Ndugga et al., 2021). This is staggering, especially when considering that Black patients have been reported to account for up to 33% of U.S. COVID-19 cases and make up only 13% of the U.S. population (Dyer, 2020).

Further, many believe that the COVID-19 vaccines, particularly those stemming from research funded by the Bill and Melinda Gates Foundation, are being used as a cover for inserting microchips to monitor global citizens (Evanega et al., 2020; Goodman & Carmichael, 2020). This theory likely stems from a misunderstanding of the Gates Foundation awarding grants to researchers to find better ways to keep track of vaccination information (Goodman & Carmichael, 2020). One such method, proposed by researchers at MIT, was to use a dye containing quantum dots to keep track of those who have already been vaccinated (McHugh et al., 2019). Even though this method never expanded past animal trials, it was cited as proof of an intent to covertly monitor vaccinated individuals.

There are at least two main conspiracy theories connecting 5G technologies with COVID-19 (Meese et al., 2020). One version posits that electromagnetic radiation from 5G lowers the immune system, thereby making the population more susceptible to the virus. This theory shares clear similarities with another conspiratorial claim that sustained exposure to electromagnetic fields causes cancer. However, the more prominent conspiracy theory argues that 5G directly causes COVID-19 via radiation and that the virus initiated in Wuhan, due to the city being the "initial" 5G test site. Although Wuhan was one of the first cities in China to receive 5G (Xinhua, 2019), it has not been confirmed to be the first either in China or globally. It is unclear exactly how these theories began and were so quickly propagated, but they are possibly descendants of previous conspiracies related to electromagnetic radiation and cell phones.

One of the only conspiracy theories which can be traced directly back to a government official, the belief that COVID-19 is no worse than a standard flu infection and the health risks are overblown, was presented by former U.S. President Donald Trump

in early 2020 (Beer, 2020; Woodward, 2020). Mr. Trump was later revealed to have stated in an interview with Bob Woodward (2020) that, not only was the virus more deadly than the flu, but that he was also intentionally attempting to downplay its severity. Despite this, the initial statements made by the Trump administration, in conjunction with since-debunked claims made in the conspiracy documentary film *Plandemic*, allowed conspiracies attacking the severity of the pandemic to take hold (Evanega et al., 2020; Prichard & Christman, 2020). As late as November 2020, 38% of Americans reported believing the seriousness of COVID-19 was being exaggerated, and 36% reported believing that it was definitely or probably true that the pandemic was a planned conspiracy (Prichard & Christman, 2020). This had serious implications on compliance with public safety measures. Filtration efficiency of bioaerosols was found to be as much as 94% for single-layer cloth masks and up to 99% for disposable medical-grade masks (Clase et al., 2020). Yet, there is continual proof of a strong negative correlation between conspiratorial belief and engaging in mask-wearing safety protocols (Prichard & Christman, 2020); in July 2020, 21% of Americans reported not wearing masks while in public and only 21.1% reported being "very worried" that they, or a member of their household, would become infected (Romer & Jamieson, 2020).

15.3 Features of Conspiracy Belief

To identify the ways in which CADEs can combat conspiratorial thinking, it is helpful to identify common features of conspiracy theories such as those listed in the previous section. Oswald (2016), in particular, detailed many recurring features of conspiracy belief, which we will attribute back to our previously defined COVID-19 related conspiracies. The first such feature is that, in an attempt to refute an official account of an event in favor of a conspiratorial account, an arguer will attempt to cast doubt on the integrity or competence of the official account's source, thereby engaging in an ad hominem attack. Individual susceptibility to falling back on such attacks may increase with historical framing which is perceived to support distrust in authority. For example, Black Americans may be more susceptible to engaging in ad hominem attacks in an attempt to refute vaccine efficacy if the attack is framed from the historical context of previous government-sanctioned, involuntary medical testing on Black people.

Another major feature impacting conspiratorial thinking is that belief in one conspiracy theory correlates with belief in others through unwarranted leaps in reasoning. Thereby, from an argumentative perspective, we can extrapolate that conspiracy theories are likely to rely on arguments from generalization and analogy. This is often exacerbated by the tendency of susceptible parties to misrepresent insufficient or unrelated evidence for acceptable argument premises (Byford, 2011; Keeley, 1999). This method of asserting an unrelated premise is known as *errant data* (Oswald, 2016), and using it as the basis for accepting a conspiratorial conclusion is employed most often in defending anti-establishment focused conspiracy theories,

such as those surrounding vaccine and mask efficacy. If an arguer can misrepresent prior historical context as relevant to their current belief, then they can reframe their conclusion as having stronger support. WG-A (Licato & Cooper, 2019) can be used to address these kinds of unrelated or insufficient analogical connections, as we will demonstrate in Sect. 15.6.

Many proponents of conspiracy theories have a propensity to infer or accept conclusions based solely on the lack of contradictory evidence. This kind of inference of fact based on the absence of contrary evidence is known as the ad ignorantiam, or appeal from ignorance, fallacy (Walton, 1999). This is a very common fallacy for arguers attempting to dismiss COVID-19 as a typical flu. Such arguers will attribute a decrease in reported influenza cases and deaths to a conspiracy to misreport these cases as COVID-19, as opposed to a mitigation of illness due to social distancing and mask protocols. Other arguers have used the lack of an immediately available vaccine as proof of COVID-19's genetic dissimilarity to previous SARS-CoV strains and, subsequently, as proof that COVID-19 was designed as a bioweapon, not as a naturally mutating virus.

Finally, Oswald (2016) noted that arguers of conspiracy theories are highly likely to follow a set dialectical format. Specifically, conspiratorial thinkers will argue from a "position of refutation and challenge," (p. 8) wherein they are more inclined to attack opposition to their argument than to directly defend their own argument. This places the onus of argumentation on the source of the official account, as opposed to on the arguer. CADEs are uniquely equipped to both work within and combat this feature. Because the structural characteristics of CADEs are to defend one's own position and attack the opponent's position, arguers can lean into the natural propensity to attack, but are also guided to contemplate their own argument's deficiencies. Further, by controlling the allowed moves within the dialogue environment, arguers are forced to only attack an opposing argument on merit, as opposed to from a fallacious position.

Cook et al. (2017) present two possible elements to an effective inoculation technique: (1) an "explicit warning" of an impending threat to information accuracy and (2) a refutation of an anticipated argument which exposes the imminent fallacy (p. 4). CADEs are potentially equipped to address the second element. By creating a dialogue environment where arguers are required to both defend their own position and attack their opponent's, while limiting their ability to devolve into unstructured argumentation (along with the biases and distractions that result), CADEs may motivate arguers to make their own cognitive biases explicit.

15.4 Warrant Game

We will now describe our proposed approach for combatting conspiratorial thinking through controlled argumentation. A warrant, in Toulmin's (1984; 2003) model of argumentation, is a statement connecting the premises and conclusion of an argument, showing how the premises permit the inference of the conclusion. While a premise may be any fact or evidence which an arguer uses to support a conclusion, a warrant

is a broader principle connecting the premise to its resulting conclusion. Arguers may use any heuristic means to obtain a warrant, but it must define some causal link between premise and conclusion. For example, given the premise "Humans are mammals" and the conclusion "Humans don't lay eggs," two possible warrants are W_1: "No mammals lay eggs," and W_2: "Most mammals don't lay eggs." Each warrant creates a causal link between the premise and the conclusion, but the links' levels of support differ, as does the potential methods of challenging the argument: W_1 can be disproven by simply pointing out that a platypus is an egg-laying mammal; whereas W_2 requires the contesting arguer to prove that most mammals do lay eggs. Clearly, there is a much higher onus of proof for an arguer to show that most mammals lay eggs than to assert the existence a single platypus. Therefore, the best strategy for a defending arguer is to ensure the strongest, and most generalizable, possible warrant is used for their arguments, thereby shifting the onus of proof onto the attacking arguer.

This is the concept behind the Warrant Game (WG), a CADE in which two arguers compete by iteratively improving warrants and attacking those of their opponents (Licato & Cooper, 2019). By explicitly defining the warrant, arguers are able to build and improve their own argumentative skills, many of which are relevant to the features of conspiratorial thinking, including: (1) determining what methods of attacking an argument are most effective, (2) distinguishing relevant premises from those that are unrelated the argument, and (3) defining whether the conclusion follows from the premises and the strength of the causal link. However, despite its utility in fostering cognitive skills, the warrant is often left implicit in arguments. This omission is often to the detriment of analytical reasoning (Beach et al., 2016; Warren, 2010) and allows conspiratorial thinkers, specifically, to make leaps in reasoning without regard to the bias induced by errant data. Instead, by centering argumentation on the warrant, WG is able to promote reflection on the properties of conspiratorial arguments.

In WG, arguers are presented with a controversial issue and required to produce warrants which either support or refute the conclusion. The arguers then take turns attacking their opponent's warrant on the basis of its connection to the argument. Arguers are allowed to use any of a pre-defined set of attacks (a subset of which are listed in Table 15.1), thereby inducing them to consider the argument's validity while mitigating the impact of bias. A human or artificially intelligent moderator is assigned to determine whether each attack is successful; if an attack succeeds, then the attacking party receives a certain number of points and the defending arguer loses points and is required to improve their warrant based on the nature of the attack. Therefore, WG can provide an overall model for creating and iteratively improving a warrant: (1) create a warrant which causally links the premise and conclusion, (2) determine whether the warrant is susceptible to any of the allowed attacks and revise as necessary to avoid such attacks, and (3) iterate until the warrant is sufficiently strong. Warrant strength is determined by how resistant the warrant is to attacks and an argument's overall strength is given by its strongest warrant.

Adjustments to the standard structure of WG were necessary to appropriately account for the nature of conspiratorial persuasive dialogues. First, a new allowed attack was defined on the basis of *equivocation*, which is the use of ambiguous

Table 15.1 A subset of allowed attacks in WG

Rule name	Points	Description/Tests
Clear if–then structure	+2/−2	The warrant cannot easily be rephrased into an equivalent statement of the form "If X, then Y" without changing its meaning
Premise-antecedent connection	+1/−1	The warrant's antecedent doesn't follow closely from the argument's premises
Consequent-conclusion connection	+1/−1	The argument's conclusion doesn't follow closely from the warrant's consequent
Unnecessary premise	+1/−1	A premise connected to the warrant isn't necessary (as determined by the warrant's antecedent)
Defeating counterexample	+2/−1	There is a counterexample to the warrant: a case where the antecedent is true, but the consequent is false; and this counterexample is significant enough to make the original warrant seem useless as a generalized rule
Warrant generalizability	+1/−1	The warrant is specific to a very limited number of scenarios, rather than being a general rule

language to conceal the truth or to avoid committing to an argument. As such, to make an equivocation attack, an arguer must show that multiple premises or the conclusion use the same term with different meanings. To respond, the defending arguer must either prove that the terms have equivalent meanings or clarify the definition. Second, the initial configuration of WG was based on the starting premise being unambiguous, established fact, which could be accepted by both parties. However, given that many conspiratorial arguments are based on controversial, insufficient, or unrelated premises (see Sect. 15.3), we allowed for arguers to attack their opponent's starting premise at any time. Finally, if an attack on a premise was successful, leading to its alteration, then we allowed the attacking party to challenge the new premise, and any premises resulting therein, even if they had been previously attacked prior to the alteration.

15.5 Warrant Game for Analogies

Warrant Game for Analogies (WG-A) (Cooper et al., 2020; Licato & Cooper, 2019) is a variant of WG used for the evaluation of analogical arguments based on Bartha's (2010) *Articulation Model* (AM). AM is a normative model of analogical argumentation that attempts to explain both what a "good" analogy is, and what kinds of dialogical moves can be considered relevant towards assessing an analogical argument. An analogical argument consists of propositions divided into source and target domains. A pair of analogous propositions is said to be in the *positive analogy* if they

have the same truth value, and in the *negative analogy* if they have opposite truth values. The structure of a WG-A setup is equivalent to two parallel instances of WG: two parallel arguments (the source and target analogies) are forced to share a single warrant, such that the warrant jointly explains the primary inference on both sides. In this way, the warrant takes on the properties of both the *prior association* and the *potential for generalization*, which are the two elements central to a good analogy according to Bartha's AM. Although we will summarize it here, for full details on how WG-A approximates AM, see Licato and Cooper (2019).

WG-A provides a web-based interface wherein two arguers engage in the roles of advocate and critic, and work together to evaluate a given analogical argument. Similar to WG, arguers are provided a pre-determined set of moves which have a high probability of being relevant to the argument. WG-A's central assumption is that, when provided with the source and target domains of an analogical argument, the process of explicitly defining a warrant which connects each domain's facts to its conclusion is roughly equivalent to elaborating a prior association and potential for generalization in AM (Licato & Cooper, 2019). For example, consider the analogy in Fig. 15.1. The argument begins with a set of premises referred to as "facts." We refer to the left box as the *source facts*, and the right box as the *target facts*. The analogy is shown as a pair of conclusions and the overall analogical argument is that if the source facts, target facts, and source conclusion are true, then the target conclusion must follow. The warrant, shown as the "current rule," is defined and iteratively improved by both arguers working together.

A WG-A session proceeds as follows. Two arguers, in the roles of advocate and critic, are presented with a pre-selected set of source facts, target facts, source conclusions and target conclusions. The advocate is tasked with defining the initial warrant such that (1) its antecedent is a generalization of the source and target facts, (2) its consequent is a generalization of the source and target conclusions, and (3) it serves as a causal connection between the source facts and source conclusion, and between the target facts and target conclusion. The critic may then attack the links connecting the warrant to the various facts and conclusions (labeled **L.1–L.5** in Fig. 15.1). For example, if the warrant's antecedent is not a generalization of the

Fig. 15.1 Example analogical argument and warrant in WG-A

source or target facts, then links **L.1** or **L.2** can be attacked, respectively. Furthermore, if a strong, defeating counterexample to the warrant can be found, then **L.3** can be attacked. An attack on a link must be accepted by both parties to be deemed successful (using a resolution process that discourages direct communication and instead requires very structured, template-restricted interactions), at which point the advocate is tasked with improving the warrant. Conversely, the advocate may challenge the attack as invalid and make a case for the warrant's strength. Thus, the final definition of the warrant is subject to multiple constraints, and the parlay of attacks and subsequent edits between arguers serves to iteratively improve the warrant as the game progresses.

WG-A has been shown to be capable of addressing the problem of ensuring that moves made by players are relevant to the assessment or improvement of the analogical argument under discussion (Licato & Cooper, 2019). Furthermore, when arguers play WG-A, as compared to debating an analogical argument through open-ended text-based chat, there appears to be evidence of marginal improvement in critical thinking skills, though this effect appeared only in a one-week follow-up (Cooper et al., 2020). However, WG-A is restricted to analogical arguments, whereas WG is designed to address all premise-conclusion arguments in general.

15.6 Examples

As an example of the previously-described CADEs' potential to mitigate conspiratorial thinking, we present sample argumentation dialogues for two common COVID-19 conspiracies. We will first use WG to combat the belief that COVID-19 was bioengineered in China. Suppose we provided the players with a typical conspiratorial premise, e.g., "COVID-19 is genetically dissimilar to previous coronavirus strains," and the conclusion "COVID-19 was bioengineered in China." Then it would fall upon the advocate to create a warrant which connects these nodes. The advocate may present the warrant "If a virus is naturally mutated, then it must be genetically similar to its precursors." But although this warrant contains an antecedent and consequent, it is not properly structured: this warrant could easily be challenged by the critic using a "Premise-antecedent connection attack" (Table 15.1), e.g., by saying that genetic similarity is not related to natural mutation. Subsequently, the consequent (the "THEN" part of the warrant) can be said to not connect to the conclusion: genetic similarity is not directly related to bioengineering, this is the relation the arguer is attempting to prove. At this point, the advocate may wish to completely replace the warrant with its contrapositive: "If a virus is not genetically similar to its precursors, then it is not naturally mutated." Note that the premise and warrant antecedent are now parallel. We may now begin the iterative stage of gameplay with the argument presented in Fig. 15.2.

The critic has a few options for attack at this point. Let us assume that they choose to attack **L.3** by stating that a virus not being naturally mutated does not necessarily imply that it is bioengineered. The critic may also provide a counterexample to

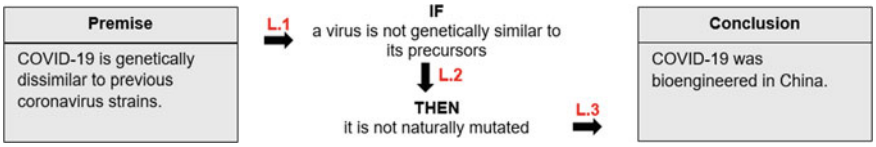

Fig. 15.2 Initial warrant for WG example of "COVID-19 was bioengineered in China"

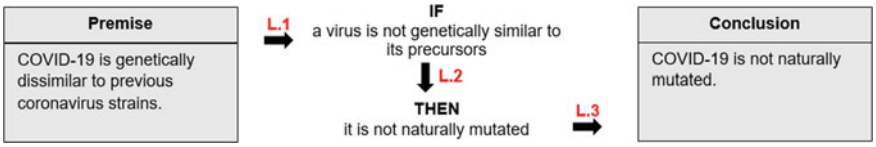

Fig. 15.3 Modified conclusion for WG example of "COVID-19 was bioengineered in China"

strengthen this attack, e.g., an individual with HIV being infected by a new, distinct HIV strain, which then forms a recombinant superinfection strain (Redd et al., 2013). This is a valid attack within the current environment and should be accepted by the advocate. As such, the advocate is now compelled to modify either of **L.3**'s nodes, i.e., the warrant's consequent or the conclusion. Assume they choose to modify the conclusion to now say "COVID-19 is not naturally mutated." The resulting structure, as seen in Fig. 15.3, now leaves **L.3** less vulnerable to attack. However, the critic can now attack the initial premise on the basis of equivocation: because neither player has defined the acceptable bounds of genetic similarity (and therefore, genetic dissimilarity), the initial premise may utilize a different meaning of 'genetic similarity' than that used in the antecedent.

To respond, the advocate must now either show that the premise and warrant are using the same definition of genetic similarity or clarify the definition and adjust the warrant's antecedent accordingly. In either case, this requires them to explicitly define genetic similarity within the context of their argument. Ideally, a rational actor would at this point realize that the premise "COVID-19 is genetically dissimilar to previous coronavirus strains" is impracticably vague at best, and unequivocally false at worst. But even if the advocate decided to continue with the game, the remainder of the argument is subject to attacks, the corrections of which would effectively neuter the argument's conspiratorial component. In order to correct this argument, the advocate would have to concede the game and reevaluate their argument with more defensible components. Note that the given series of attacks allows the critic to deconstruct the argument without having to directly attack the irrelevant aspect of the conclusion, namely that the virus was bioengineered in China, specifically. Multiple attack strategies are possible based on the strength of the warrant, and this is just one example. However, by invalidating the premise, WG forces the arguer to either weaken their conclusion to something less conspiratorial (e.g., "COVID-19 is not naturally mutated") or to improve an indefensible premise ("COVID-19 is genetically dissimilar to previous coronavirus strains").

WG can continue indefinitely, in an iterative loop of attack and response. However, let us now shift gears and provide an analogical example, using WG-A to combat the conspiratorial belief that COVID-19 vaccinations are being used to covertly test on Black Americans. Suppose we provided the players with errant data via the source fact "Scientists claimed that participants would receive medical treatment for syphilis during the Tuskegee study" and presented the target fact "Scientists claim that Americans will receive vaccines to prevent COVID-19" as a relevant relationship. Then suppose we provide the source conclusion "The Tuskegee study was used to covertly test the effects of syphilis on Black men," which is a historical fact, and the target conclusion "The COVID-19 vaccines are being used to covertly test on Black people," which is a conspiratorial belief. It would fall upon the advocate to create a starting warrant which connects these nodes. They may present the conspiracy-reminiscent warrant "If scientists are offering medical treatment to Black people, then the medical treatment is being used to covertly test on Black people." The corresponding WG-A environment for this initial argument is presented in Fig. 15.4.

The critic now has the option to attack the links to the advocate's warrant, update the facts provided, or to add new facts. Let us assume that they choose to attack **L.3** by providing a counterexample where scientists offer medical treatment to Black people without covertly testing on them (e.g., ethical studies to treat sickle cell anaemia). The advocate accepts this attack as valid and is now compelled to modify the warrant's antecedent to say, "If scientists are offering medical treatment to Black people and they are lying about the treatment provided, then the medical treatment is being used to covertly test on Black people." This updated warrant, shown in Fig. 15.5, prevents the critic from launching the same attack, but its antecedent is now no longer a clear generalization of the source and target facts (due to the antecedent now including an intent to deceive), thus opening up the possibility of **L.1** and **L.2** attacks. We assume the critic chooses to attack **L.2** by saying that neither player has established that scientists are lying about the treatment provided with the COVID-19 vaccines.

The advocate may respond by introducing new facts into the source and target domains: "Scientists did not provide medical treatment for syphilis during the Tuskegee study" and "Scientists are not providing medical treatment to prevent

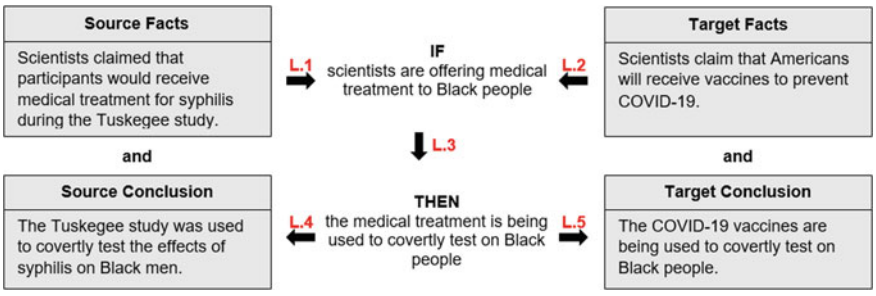

Fig. 15.4 Initial warrant for WG-A example of "COVID-19 vaccines are being used to covertly test on Black people"

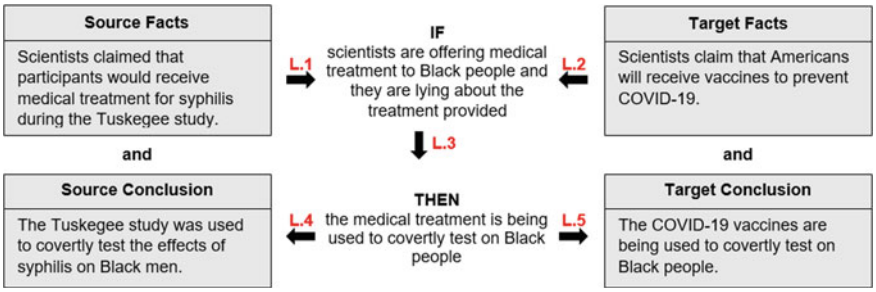

Fig. 15.5 Modified warrant for WG-A example of "COVID-19 vaccines are being used to covertly test on Black people"

COVID-19 with the vaccines." Note that the new target fact exposes the advocate's belief that the vaccines are either ineffective or fake. Both players must accept the phrasing of new facts before they are included, and here we assume the critic refuses to accept this fact addition on the basis that their opponent has not proven the COVID-19 vaccines are ineffective. The advocate may attempt to make the argument that the vaccines are fake, further clarifying their own biases, but this is equally unproven. In the course of an unstructured discussion, this side issue can quickly devolve into a full debate of its own, detracting from the focus on the original argument. By disallowing direct dialogue between participants, WG-A avoids this potential complication. In the present case, if the advocate is not able to quickly and convincingly show that the new facts they propose to introduce are supported, they will not be allowed, and the advocate will need to try a different tactic (ideally, one which draws on claims that can be better supported).

We have thus reached a point where the overall argument is no longer clearly defensible, at least in its conspiratorial form—a similar pattern of attacks will require modifications of the warrant and source/target facts in such a way that will either open the argument up to an increasing number of attacks or weaken the conclusion substantially. As can be seen in this instance, WG-A does not necessarily invalidate the target conclusion (i.e., the conspiratorial belief) entirely, and does not preclude the possibility that the conclusion may be true. This CADE merely confronts the analogy, itself, and allows arguers to (1) identify the basis of their own conspiratorial belief ("COVID-19 vaccines are either ineffective or fake,") and (2) recognize errant data as irrelevant to the argument. Therefore, both arguers have improved argumentative skills to recognize a similarly fallacious argument in the future.

15.7 Conclusion

As discussed previously, decreasing conspiratorial belief has far-reaching implications for minimizing the health risks of COVID-19. There is proof of a strong

negative correlation between conspiratorial belief and engaging in safety protocols (Romer & Jamieson, 2020), despite conspiratorial thinkers' concern for their own health outcomes (Prichard & Christman, 2020). Echo chamber effects exacerbate conspiratorial polarization on social media (Del Vicario et al., 2016; Schmidt et al., 2018) and, as such, any contradictory evidence must be presented in an environment separate from emotional influence.

Warrant-centered reasoning is useful to improving conspiratorial argumentation in ways not limited to the following: By making warrants explicit, counterexamples to them can be found, and in response the warrants can be assessed and improved iteratively. This provides a framework to incorporate counterarguments, and more tightly link premises to their conclusions. This iterative improvement is the driving factor behind WG, described in Sect. 15.4. Further, by allowing arguers a means to iteratively improve arguments to which they are exposed, it is possible for CADEs to both inoculate susceptible individuals against conspiratorial misinformation and allow them to reduce conspiratorial arguments to their relevant facts when they are exposed.

A long-term goal of WG and WG-A is to break down warrant-centered reasoning into steps that are amenable to implementation, and moderation, by artificially intelligent algorithms. In keeping with that goal, if we were provided a dataset of argument premises and conclusions made by reasoners from some particular community, a suitably powerful warrant induction reasoning tool may be able to identify the argumentative norms used by these reasoners, thereby allowing a comparison of such norms across datasets. As such, CADEs have the potential to allow researchers to identify common argumentative features of conspiracy theories and classify their susceptible populations into categories for further intervention.

Finally, the value of a warrant induction tool in personal argumentation is particularly appealing, as it suggests the ability to identify flaws in one's arguments before those arguments are made public. We hope, in the future, to study the possible inoculative effects of CADEs like WG and WG-A. By making conspiratorial thinkers aware of their own reasoning tendencies, and identifying inconsistencies, we may be able to further inoculate these reasoners against their own common cognitive biases and reduce the proliferation of misinformation. As a natural consequence, it is possible that applying warrant-based reasoning to a valid, but flawed, conspiratorial argument may expose the irrelevant aspects and make the overall argument stronger. This method is not intended to completely dissolve all conspiratorial thinking, indiscriminately, but rather to aid conspiratorial thinkers in recognizing their own cognitive biases, clarifying their arguments, and ultimately becoming better reasoners.

It should be noted that the goal of this paper was to show that some of the patterns symptomatic of conspiratorial thinking can be minimized using controlled argumentation dialogue environments, such as WG and WG-A. However, this is but one piece of the puzzle. Further work is required to study whether participants using CADEs are more likely to be convinced of the arguments that arise from them. Although preliminary empirical work with WG-A has suggested its use has some power to influence reasoning patterns (Cooper et al., 2020), conspiratorial thinking may often be accompanied with powerful emotional motivators that may lead to a backlash

effect. The question of how best to combat unfounded conspiratorial thinking thus remains open, but it is our hope that work such as that reported here can serve as a useful tool in that regard.

Acknowledgements This publication is based upon work from COST Action CA-17132, *European Network for Argumentation and Public Policy Analysis* (http://publicpolicyargument.eu), supported by COST (European Cooperation in Science and Technology).

References

Bajaj, S. S., & Stanford, F. C. (2021). Beyond Tuskegee—vaccine distrust and everyday racism. *New England Journal of Medicine, 384*(5), e12. https://doi.org/10.1056/NEJMpv2035827

Bartha, P. F. (2010). *By parallel reasoning: The construction and evaluation of analogical arguments*. Oxford University Press.

Basol, M., Roozenbeek, J., & van der Linden, S. (2020). Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of Cognition, 3*(1), 2. https://doi.org/10.5334/joc.91

Beach, R., Thein, A. H., & Webb, A. (2016). *Teaching to exceed the English language arts common core standards: A critical inquiry approach for 6–12 classrooms* (2nd ed.). Routledge Publishing.

Beer, T. (2020, September 10). All the times Trump compared Covid-19 to the flu, even after he knew Covid-19 was far more deadly. *Forbes*. https://www.forbes.com/sites/tommybeer/2020/09/10/all-the-times-trump-compared-covid-19-to-the-flu-even-after-he-knew-covid-19-was-far-more-deadly/

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

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences, 114*(28), 7317–7318. https://doi.org/10.1073/pnas.1618923114

Brenan, M. (2021, February 10). Two-thirds of Americans not satisfied with vaccine rollouts. *Gallup News*. https://news.gallup.com/poll/329552/two-thirds-americans-not-satisfied-vaccine-rollout.aspx

Brewster J. (2021, June 16). Here's what Dr. Fauci has said about Covid's origins and the lab leak theory. *Forbes*. https://www.forbes.com/sites/jackbrewster/2021/06/16/heres-what-dr-fauci-has-said-about-covids-origins-and-the-lab-leak-theory/

Byford, J. (2011). *Conspiracy theories: A critical introduction*. Palgrave Macmillan.

Chen, V. Y. (2020). Examining news engagement on Facebook: Effects of news content and social networks on news engagement. *Mass Communication and Society, 23*(6), 833–857. https://doi.org/10.1080/15205436.2020.1798462

Choi, E., Johnson, D. A., Moon, K., & Oah, S. (2018). Effects of positive and negative feedback sequence on work performance and emotional responses. *Journal of Organizational Behavior Management, 38*(2–3), 97–115. https://doi.org/10.1080/01608061.2017.1423151

Clase, C. M., Fu, E. L., Joseph, M., Beale, R. C. L., Dolovich, M. B., Jardine, M., Mann, J. F. E., Pecoits-Filho, R., Winkelmayer, W. C., & Carrero, J. J. (2020). Cloth masks may prevent transmission of COVID-19: An evidence-based, risk-based approach. *Annals of Internal Medicine*, M20. https://doi.org/10.7326/M20-2567

Clavel, G. (2020, April 17). Luc Montagnier, prix Nobel controversé, accuse des biologistes d'avoir créé le coronavirus [Luc Montagnier, controversial Nobel laureate, accuses biologists of having created the coronavirus]. *Huffington Post France*. https://www.huffingtonpost.fr/entry/luc-mon

tagnier-le-controverse-prix-nobel-accuse-des-biologistes-davoir-cree-le-coronavirus_fr_5e998515c5b63639081ccebc

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLoS ONE, 12*(5). https://doi.org/10.1371/journal.pone.0175799

Cooper, M., Fields, L., Badilla, M. G., & Licato, J. (2020). WG-A: A framework for exploring analogical generalization and argumentation. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd cognitive science society conference* (pp. 1502–1508). Cognitive Science Society.

COVID Collaborative. (2020). *COVID collaborative survey: Coronavirus vaccination hesitancy in the Black and Latinx communities*. https://www.covidcollaborative.us/content/vaccine-treatments/coronavirus-vaccine-hesitancy-in-black-and-latinx-communities

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*, 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

Dyer, O. (2020). Covid-19: Black people and other minorities are hardest hit in US. *The BMJ, 369*, m1483. https://doi.org/10.1136/bmj.m1483

Evanega, S., Lynas, M., Adams, J., & Smolenyak, K. (2020). *Coronavirus misinformation: Quantifying sources and themes in the COVID-19 'infodemic.'* Cornell Alliance for Science. https://allianceforscience.cornell.edu/wp-content/uploads/2020/10/Evanega-et-al-Coronavirus-misinformation-submitted_07_23_20-1.pdf

Ferrara, E., & Yang, Z. (2015). Quantifying the effect of sentiment on information diffusion in social media. *PeerJ Computer Science, 1*, e26. https://doi.org/10.7717/peerj-cs.26

Goodman, J., & Carmichael, F. (2020, May 30). *Coronavirus: Bill Gates 'microchip' conspiracy theory and other vaccine claims fact-checked*. BBC. https://www.bbc.com/news/52847648

Kahneman, D. (2011). *Thinking, fast and slow.* Farrar, Straus and Girous.

Keeley, B. L. (1999). Of conspiracy theories. *Journal of Philosophy, 96*(3), 109–126. https://doi.org/10.2307/2564659

Koyama, T., Lauring, A., Gallo, R., & Reitz, M. (2020, September 24). *Reviews of "Unusual features of the SARS-CoV-2 genome suggesting sophisticated laboratory modification rather than natural evolution and delineation of its probable synthetic route."* Rapid Reviews: COVID-19. https://rapidreviewscovid19.mitpress.mit.edu/pub/78we86rp/release/2

Ksiazek, T. B. (2016). Commenting on the news: Explaining the degree and quality of user comments on news websites. *Journalism Studies, 19*(5), 650–673. https://doi.org/10.1080/1461670X.2016.1209977

The Lancet. (2020). The COVID-19 infodemic. *The Lancet Infectious Diseases*, *20*(8), 875. https://doi.org/10.1016/S1473-3099(20)30565-X

Licato, J., & Cooper, M. (2019). Evaluating relevance in analogical arguments through warrant-based reasoning. In C. D. Novaes, H. Jansen, J. A. van Laar, & B. Verheij (Eds.), *Reason to dissent: Proceedings of the 3rd European conference on argumentation* (Vol. 1, pp. 375–394). College Publications

McHugh, K. J., Jing, L., Severt, S. Y., Cruz, M., Sarmadi, M., Jayawardena, H. S. N., Perkinson, C. F., Larusson, F., Rose, S., Tomasic, S., Graf, T., Tzeng, S. Y., Sugarman, J. L., Vlasic, D., Peters, M., Peterson, N., Wood, L., Tang, W., Yeom, J., … Jaklenec, A. (2019). Biocompatible near-infrared quantum dots delivered to the skin by microneedle patches record vaccination. *Science Translational Medicine*, *11*(523), eaay7162. https://doi.org/10.1126/scitranslmed.aay7162

Meese, J., Frith, J., & Wilken, R. (2020). COVID-19, 5G conspiracies and infrastructural futures. *Media International Australia, 177*(1), 30–46. https://doi.org/10.1177/1329878X20952165

Mercier, H. (2016). The argumentative theory: Predictions and empirical evidence. *Behavioral and Brain Sciences, 20*(9), 689–700. https://doi.org/10.1016/j.tics.2016.07.001

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(2), 57–74. https://doi.org/10.1017/S0140525X10000968

Mercier, H., & Sperber, D. (2017). *The Enigma of Reason*. Harvard University Press.

National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. (2020, April 22). *U.S. public health service syphilis study at Tuskegee*. Centers for Disease Control and Prevention. https://www.cdc.gov/tuskegee/index.html

Ndugga, N., Pham, O., Hill, L., Artiga, S., Alam, R., & Parker, N. (2021, June 23). *Latest data on COVID-19 vaccinations by race/ethnicity*. Kaiser Family Foundation. https://www.kff.org/coronavirus-covid-19/issue-brief/latest-data-on-covid-19-vaccinations-race-ethnicity/

Oswald, S. (2016). Conspiracy and bias: Argumentative features and persuasiveness of conspiracy theories. In L. Benacquista & P. Bondy (Eds.), *Proceedings of the 11th International conference of the Ontario society for the study of argumentation*. Ontario Society for the Study of Argumentation.

Prichard, E. C., & Christman, S. D. (2020). Authoritarianism, conspiracy beliefs, gender, and COVID-19: Links between individual differences and concern about COVID-19, mask wearing behaviors, and the tendency to blame China for the virus. *Frontiers in Psychology, 11*, 597671. https://doi.org/10.3389/fpsyg.2020.597671

Pytlik, N., Soll, D., & Mehl, S. (2020). Thinking preferences and conspiracy belief: Intuitive thinking and the jumping to conclusions-bias as a basis for the belief in conspiracy theories. *Frontiers in Psychiatry, 11*, 568942. https://doi.org/10.3389/fpsyt.2020.568942

Rasmussen, A. L. (2021). On the origins of SARS-CoV-2. *Nature Medicine, 27*(8–9), 9. https://doi.org/10.1038/s41591-020-01205-5

Redd, A. D., Quinn, T. C., & Tobian, A. A. R. (2013). Frequency and implications of HIV superinfection. *The Lancet Infectious Diseases, 13*(7), 622–628. https://doi.org/10.1016/S1473-3099(13)70066-5

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: The cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology, 79*(3), 410–424. https://doi.org/10.1037/0022-3514.79.3.410

Romer, D., & Jamieson, K. H. (2020). Conspiracy theories as barriers to controlling the spread of COVID-19 in the US. *Social Science & Medicine, 263*, 113356. https://doi.org/10.1016/j.socscimed.2020.113356

Roozenbeek, J. (2020, July 9). *How to "inoculate" people against misinformation and online extremism* [Conference session]. Association for Computational Linguistics Third Workshop on Fact Extraction and Verification. https://virtual.acl2020.org/workshop_W5.html

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., & van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society for Open Science, 7*(10), 201199. https://doi.org/10.1098/rsos.201199

Roozenbeek, J., & van der Linden, S. (2018). The fake news game: Actively inoculating against the risk of misinformation. *Journal of Risk Research, 22*(5), 570–580. https://doi.org/10.1080/13669877.2018.1443491

Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications, 5*(1), 65. https://doi.org/10.1057/s41599-019-0279-9

Schmidt, A. L., Zolla, F., Scala, A., Betsch, C., & Quattrociocchi, W. (2018). Polarization of the vaccination debate on Facebook. *Vaccine, 36*(25), 3606–3612. https://doi.org/10.1016/j.vaccine.2018.05.040

Skloot, R. (2000, April). Henrietta's dance. *Johns Hopkins Magazine*. https://pages.jh.edu/jhumag/0400web/01.html

Skloot, R. (2011). *The immortal life of Henrietta Lacks*. Random House, Inc.

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning, 13*(3), 225–247. https://doi.org/10.1080/13546780600780796

Steidl, S., Razik, F., & Anderson, A. K. (2011). Emotion enhanced retention of cognitive skill learning. *Emotion, 11*(1), 12–19. https://doi.org/10.1037/a0020288

Swami, V., Voracek, M., Steiger, S., Tran, U. S., & Furnham, A. (2014). Analytic thinking reduces belief in conspiracy theories. *Cognition, 133*(3), 575–585. https://doi.org/10.1016/j.cognition.2014.08.006

Toulmin, S. (2003). *The uses of argument (Updated Edition).* Cambridge University Press.

Toulmin, S., Rieke, R., & Janik, A. (1984). *An introduction to reasoning* (2nd ed.). Macmillan Publishing Company.

van der Linden, S., Leiserowitz, A., Rosenthal, S., & Maibach, E. (2017). Inoculating the public against misinformation about climate change. *Global Challenges, 1*(2), 1600008. https://doi.org/10.1002/gch2.201600008

van der Linden, S., Roozenbeek, J., & Compton, J. (2020). Inoculating against fake news about COVID-19. *Frontiers in Psychology, 11*, 566790. https://doi.org/10.3389/fpsyg.2020.566790

Walton, D. (1999). The appeal to ignorance, or argumentum as ignorantiam. *Argumentation, 13*, 367–377. https://doi.org/10.1023/A:1007780012323

Warren, J. E. (2010). Taming the warrant in Toulmin's model of argument. *English Journal, 99*(6), 41–46. https://doi.org/10.2307/20787665

Woodward, B. (2020). *Rage.* Simon & Schuster.

World Health Organization, United Nations, International Telecommunication Union, & International Federation of Red Cross and Red Crescent Societies. (2020, September 23). *Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation.* World Health Organization. https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation

Xinhua. (2019, May 27). Hubei sets up over 300 5G base stations. *XinhuaNet News.* http://www.xinhuanet.com/english/2019-05/27/c_138094302.htm