



The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms

Heather Wolters, Kasey Stricklin, Neil Carey, and Megan K. McBride

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

Abstract

Disinformation is a growing national security concern with a primary impact occurring in the mind. This study explores the underlying psychological principles facilitating the absorption and spread of disinformation and outlines options for countering disinformation grounded in cognitive and social psychological literature. The report serves as an introduction to the topic for policy- and defense-decision makers with no prior background in psychology or disinformation. It first examines what disinformation is and why it is important before turning to the four key psychological principles relevant to this topic. A key finding is that the principles themselves are neutral and normal cognitive processes that malign actors can exploit to achieve their own objectives.

This document contains the best opinion of CNA at the time of issue.

It does not necessarily represent the opinion of the sponsor.

Distribution

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

9/27/2021

This work was performed under Federal Government Contract No. N00014-16-D-5003.

Cover image credit: "Internet effects on brain, conceptual illustration," Science Photo Library via Britannica ImageQuest, accessed Sept. 21, 2021, https://quest.eb.com/search/132_3125015/1/132_3125015/cite.

Approved by:

September 2021



Jonathan Schroden, Research Program Director
Countering Threats and Challenges Program
Strategy, Policy, Plans, and Programs Division

Request additional copies of this document through inquiries@cna.org.

| REPORT DOCUMENTATION PAGE | | | <i>Form Approved</i> <i>OMB No. 0704-0188</i> | | |
|---|-------------------------|-------------------------------|---|-------------------------------------|--|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS. | | | | | |
| 1. REPORT DATE (DD-MM-YYYY) September 2021 | | 2. REPORT TYPE 1Rev | | 3. DATES COVERED (From - To) | |
| 4. TITLE AND SUBTITLE (U) The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms | | | 5a. CONTRACT NUMBER N00014-16-D-5003 | | |
| | | | 5b. GRANT NUMBER | | |
| | | | 5c. PROGRAM ELEMENT NUMBER 0605154N | | |
| 6. AUTHOR(S) Heather Wolters, Kasey Stricklin, Neil Carter, Megan McBride | | | 5d. PROJECT NUMBER R0148 | | |
| | | | 5e. TASK NUMBER E236.00 | | |
| | | | 5f. WORK UNIT NUMBER | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Center for Naval Analyses 3003 Washington Blvd Arlington, VA 22201 | | | 8. PERFORMING ORGANIZATION REPORT NUMBER DRM-2021-U-029337-1Rev | | |
| 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of the Chief of Naval Operations (N803) 2000 Navy Department Washington, D.C. 20350 | | | 10. SPONSOR/MONITOR'S ACRONYM(S) | | |
| | | | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) | | |
| 12. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited. | | | | | |
| 13. SUPPLEMENTARY NOTES | | | | | |
| 14. ABSTRACT Disinformation is a growing national security concern with a primary impact occurring in the mind. This study explores the underlying psychological principles facilitating the absorption and spread of disinformation and outlines options for countering disinformation grounded in cognitive and social psychological literature. The report serves as an introduction to the topic for policy- and defense-decision makers with no prior background in psychology or disinformation. It first examines what disinformation is and why it is important before turning to the four key psychological principles relevant to this topic. A key finding is that the principles themselves are neutral and normal cognitive processes that malign actors can exploit to achieve their own objectives. | | | | | |
| 15. SUBJECT TERMS disinformation, psychology, psychological principles, policy, defense-decision making, psychology of disinformation | | | | | |
| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON Knowledge Center/Tanya McCants |
| a. REPORT U | b. ABSTRACT U | c. THIS PAGE U | | | |

Executive Summary

The rise of disinformation

Disinformation, or the intentional creation and spread of false information, is a growing national security concern. While the use of information operations is not a new phenomenon—various actors have used them throughout history for a range of objectives—the connectivity that characterizes the world today allows both information and disinformation to spread faster and with a much greater reach. It is not surprising that a rapidly increasing number of adversaries and domestic US actors are coming to understand the utility of the information space for achieving their objectives and are seeking to weaponize it. The use of disinformation has led directly to real-world events and violence, can have a demonstrable impact on a recipient’s behavior, and can lead its promulgators to achieve some goals simply through its existence, regardless of its believability. Because disinformation’s primary impact occurs in the mind, technological, political, or military solutions alone cannot sufficiently mitigate the threat.

Psychological principles associated with disinformation spread

This report is the result of an extensive literature review across multiple domains, including disinformation, psychology, military science, foreign affairs, economics, computer science, and marketing. Through the literature review, we identified four key psychological principles related to the absorption and spread of disinformation: initial information processing; cognitive dissonance; the influence of groups, beliefs, and novelty; and the role of emotions and arousal. This report describes each psychological principle, explains how the principle contributes to the absorption and spread of disinformation, and details ways to mitigate the effect of the principle on the spread of disinformation.

A critical takeaway from the identification of these principles is that they are not unique to absorbing and spreading disinformation. These same principles are key to absorbing and spreading true information as well. Thus, at an individual level, it appears that disinformation is absorbed and spread through normal, routine, and adaptive mechanisms, which malign actors can exploit and manipulate for their own objectives. The four psychological principles we identified are:

- **Initial information processing:** Our mental “processing capacity” is limited; we simply cannot deeply attend to all new information we encounter. Our brains take mental shortcuts to incorporate new information, and those shortcuts can open us up to mistakes. To the extent that we do not process information as deeply as we should, disinformation can be construed as true information.
- **Cognitive dissonance:** Cognitive dissonance describes the discomfort we feel when we are confronted with two competing ideas. We are motivated to reduce the dissonance by changing one attitude, removing (ignoring) the contradictory information, discounting the importance of contradictory information, or increasing the importance of compatible information. If disinformation supports our initial beliefs or creates less dissonance than true information, we are more likely to believe the disinformation.
- **Influence of group membership, beliefs, and novelty (the GBN model):** Not all information is equally valuable to individuals. Our group memberships, our beliefs, and the uniqueness of the information influence whether we absorb and share disinformation. We are more likely to share information with people we consider members of our group, when we believe the information is true, and when it is novel or urgent. If disinformation is coming from a group member with whom we identify, is consistent with our beliefs, or is new information for us, we are more likely to share it.
- **Role of emotion and arousal in our sharing of disinformation:** Just as not all information is equally valuable, not all information affects us the same way. Research demonstrates that we pay more attention to information that makes us feel positively or that arouses us to act. That means we are more likely to share information if we feel awe, amusement, or anxiety than if we feel sadness or contentment. Given that disinformation is, by definition, created by someone, it is more likely to be absorbed and shared if it is constructed to be emotional and arousing.

Countering the absorption and spread of disinformation

The research team identified several techniques in the literature that could be useful for countering disinformation absorption and spread. Two techniques to help counter disinformation were associated with more than one psychological principle: (1) preventive inoculation (i.e., warning people about the effects of disinformation and how to spot it) and (2) encouraging deeper, analytic thinking. The literature also details techniques that could combat the effects of a single psychological principle. For example, to counter the effects associated with groups, beliefs, and novelty, researchers recommend creating unique content that encourages individuals to identify with a broader “group” and increases their access to opposing information. In addition, researchers recommend that disinformation containment

policies emphasize behavioral interventions aimed at countering the psychological principles activated by disinformation, rather than solely focusing on stopping the malicious use of bots, algorithms, and technologies. Before disinformation is shared, it is absorbed by an individual. Thus, interventions that disrupt how an individual absorbs disinformation should interrupt the chain between seeing disinformation and sharing it.

This page intentionally left blank.

Contents

| | |
|--|-----------|
| Introduction..... | 1 |
| Report elements..... | 5 |
| Report organization | 6 |
| Approach and sources..... | 6 |
| What Is Disinformation? | 8 |
| Types of disinformation | 11 |
| Why Disinformation Matters..... | 15 |
| Adversary use of disinformation | 15 |
| Domestic use of disinformation | 18 |
| Is disinformation effective?..... | 19 |
| Summary | 22 |
| Psychological Principles That Can Facilitate Disinformation Adoption and Spread | 24 |
| Initial information processing (dual process theory)..... | 25 |
| Implications of initial information processing for disinformation | 28 |
| What can we do about it? | 30 |
| Cognitive dissonance theory..... | 34 |
| Implications of cognitive dissonance for disinformation | 38 |
| What can we do about it? | 39 |
| The Group, Belief, Novelty (GBN) model | 42 |
| Example of the GBN model using authoritarian attitudes..... | 44 |
| Implications of the GBN Model for disinformation | 45 |
| What can we do about it? | 46 |
| Emotions and arousal..... | 48 |
| Implications of emotionally arousing content for disinformation..... | 50 |
| What can we do about it? | 51 |
| Summary | 52 |
| Conclusion | 53 |
| Glossary | 54 |
| Figures | 56 |
| Tables | 57 |
| References..... | 58 |

This page intentionally left blank.

Introduction

On the morning of September 11, 2014—the anniversary of 9/11—the news in St. Mary Parish, Louisiana, was alarming. At around 8:30 A.M., the director of the regional Office of Homeland Security and Emergency Preparedness received a call from a citizen concerned about a text message suggesting that there had been a chemical spill: “Toxic fume hazard warning in this area until 1:30 PM. Take Shelter. Check Local Media and columbiachemical.com.”

However, as reporters and analysts documented months later, the text message did not occur in isolation. Hundreds of Twitter accounts, many of which appeared to belong to concerned citizens in the area, used the hashtag #ColumbianChemicals to report on the unfolding crisis by asking questions, sharing pictures, and posting videos (Figure 1).

Figure 1. Twitter post showing the Columbian Chemicals plant explosion



Source: John Borthwick, “Media hacking,” Render, Mar. 7, 2015, <https://render.betaworks.com/media-hacking-3b1e350d619c>.

At the same time, reporters and politicians as far away as New York City began to hear about the emergency. One user tweeted at New Orleans *Times-Picayune* reporter Heather Nolan:

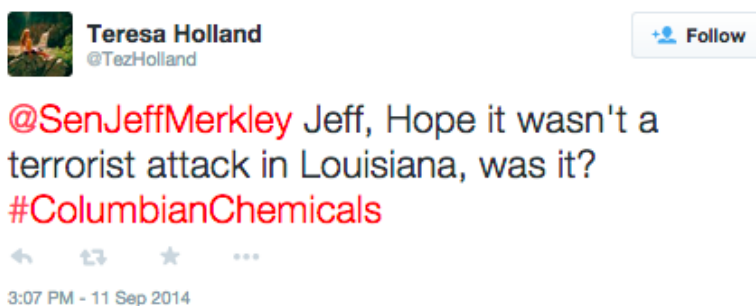
@EricTraPPP: Heather, I'm sure that the explosion at the #ColumbianChemicals is really dangerous. Louisiana is really screwed now.

Another tweeted at political consultant Karl Rove:

@zpokodon9: Karl, Is this really ISIS who is responsible for #ColumbianChemicals? Tell @Obama that we should bomb Iraq!

Yet another user tweeted at Oregon senator Jeff Merkley (Figure 2):

Figure 2. Twitter post asking Oregon senator Jeff Merkley about the Columbian Chemicals plant explosion



Source: Matt Kodama, "#ColumbianChemicals Hoax: Trolling the Gulf Coast for Deceptive Patterns," June 12, 2015, <https://www.recordedfuture.com/columbianchemicals-hoax-analysis/>.

Twitter was not the only platform involved. On YouTube, a video in which "a man showed his TV screen, tuned to an Arabic news channel, on which masked ISIS fighters delivered a speech next to looping footage of an explosion" suggested that ISIS [Islamic State of Iraq and Syria] was responsible for what was now being called an attack.¹ Another video, titled "Panic due to the explosion on a Columbian Chemicals facility in Louisiana," showed an ambulance on the highway. Yet another, titled "Flash from an explosion on a Columbian Chemicals facility in Louisiana," allegedly showed footage of the actual explosion.² A Wikipedia page, citing the video, became active, and a public Facebook page (called "Louisiana News") posted an article about the crisis and an alleged statement from ISIS in which the group claimed

¹ Adrian Chen, "The Agency," *New York Times Magazine*, June 2, 2015, accessed Feb. 2, 2021, <https://www.nytimes.com/2015/06/07/magazine/the-agency.html>.

² "Columbian Chemicals Plant Explosion Hoax," Know Your Meme, 2014, accessed Feb. 4, 2021, <https://knowyourmeme.com/memes/columbian-chemicals-plant-explosion-hoax>.

responsibility for the attack. Finally, both the New Orleans *Times-Picayune* and CNN appeared to pick up the story as tweets of coverage on their websites began to circulate.

Figure 3. Twitter account posting a fake CNN homepage showing the story of the Columbian Chemicals plant explosion



Source: John Borthwick, "Media hacking," Render, Mar. 7, 2015, <https://render.betaworks.com/media-hacking-3b1e350d619c>.

All of this, however, was news to the people working at Columbian, *as there had not actually been any sort of leak or explosion*. The entire event had been fabricated, perpetuated by a

skillfully coordinated social media infrastructure that included a botnet spreading content and “fully functional clones of the websites of Louisiana TV stations and newspapers.”³

This particular hoax had a relatively small reach since none of the tweets went viral, the mainstream media did not report on the false story at the time of the ongoing operation, and its breakthrough at the national level came months later when *New York Times* reporter Adrian Chen linked the effort to Russia’s now infamous Internet Research Agency (IRA). Yet, despite the hoax’s failure, it is an almost perfect example of a disinformation campaign.

As the Columbian Chemicals example demonstrates, disinformation (i.e., false information created with the intent to deceive) has the potential to be concerning when it comes to national security. As adversaries and domestic actors alike increasingly use disinformation to achieve a variety of objectives, the study of how disinformation works and how it can be effectively countered has become a national security priority. Most of the study on this topic has thus far focused on disinformation tactics and tools. This report is different because it looks beyond these topics to examine how disinformation affects the recipient and the role psychology plays in one’s acceptance of disinformation, since disinformation is a method of exercising psychological influence.

The primary impact of disinformation takes place in the mind. While the psychological principles behind the absorption of disinformation are normal principles invoked when individuals take in any type of information, malign actors can exploit these principles to enhance the likelihood of disinformation’s assimilation and spread. Thus, understanding how to counter disinformation’s impact requires a firm understanding of disinformation’s effect on the mind.

To date, very little work on the subject of the psychology of disinformation has been aimed at policy-makers and defense decision-makers. This report seeks to fill that gap by detailing the psychological principles that allow disinformation to flourish in a way that is easy to understand for those with no psychology background. It can also serve as a primer for those new to the general topic of disinformation because it details the definition of disinformation and why it poses a national security threat.

³ Chen, “The Agency.” A botnet is a network of bots working in coordination. For additional information on the role of bots and botnets in propagating disinformation, please see Megan McBride, Zack Gold, and Kasey Stricklin, *Social Media Bots: Implications for Special Operations Forces*, CNA, DRM-2020-U-028199-Final, Sept. 2020, https://www.cna.org/CNA_files/PDF/DRM-2020-U-028199-Final.pdf.

Report elements

To make this complex topic accessible and useful for practical application, there are several important elements of this study. First, for the purposes of this paper, *we make no distinction between disinformation and misinformation*. While the two concepts are distinct, with the intent of the information's creator (i.e., malicious versus benign) as the key difference, this paper instead focuses on the information itself and the effect on the person viewing or hearing it. In most cases, the recipient doesn't know who created the information (e.g., the meme, the video, or the email), so the creator's intent is irrelevant to the psychological impact of these types of disinformation on the audience. We will discuss and include examples of both disinformation and misinformation, categorizing them both as disinformation for simplicity (despite their definitional differences, which we recognize).

Second, the term *disinformation* has come to mean many different things to different people, with a number of distinct, though related, concepts at various times classified as disinformation. In this report, *we focus on the elements of disinformation included in a recent State Department-funded report*, and we make the conscious decision to leave out some of the other concepts occasionally included in this category. First, we exclude such techniques as the creation of fake accounts (bots, trolls, etc.) and fake communities. For our purposes, we consider these tools and actions to be part of "influence operations," a category that includes disinformation but is not exclusively composed of the disinformation itself. Second, while disinformation is a type of influence operation, not all influence operations are disinformation, so we also exclude influence operations that do not include disinformation. In other words, this study centers on the psychology of disinformation, and not on the broader psychology of influence operations. By scoping our report in this way, we are able to speak directly to a pressing national security risk in a clear, concise, and focused manner.

In addition, we are approaching this topic from a *nontechnical perspective*, with the intent of gearing explanations toward those with no prior knowledge of this subject. Therefore, we do not include the universe of potentially relevant psychological principles because that could get overwhelming and unwieldy. The psychological principles discussed in this report emerge from the literature as highly relevant to absorbing and spreading disinformation. Additional information on these principles can be gleaned from the cited material. Moreover, the tone of this study provides clear explanations and examples to make this report useful for its intended audience—policy- and decision-makers.

Report organization

In addition to this introduction, this report has three sections all of which address important aspects of our study questions. We first describe what disinformation is within the scope of this report and hone in on the definition we then use for the remainder of the report. While there are many different and conflicting definitions of “disinformation,” our first section outlines how we are defining the topic for purposes of this report.

Our second section addresses the issue of why the topic of disinformation, and therefore this paper, is important in the first place. While much has been written on disinformation, in most cases this particular question is sidestepped and the importance of disinformation is simply stipulated. To effectively counter this threat, it is critical to understand why it matters. Consequently, we begin with an exploration of why disinformation is important from a national security perspective before turning to the tricky question of whether disinformation is actually effective at achieving its objectives. We conclude with why it is important to study the psychology of disinformation as a part of this discussion.

With a firm grounding in what disinformation is and why it matters, we then turn to the psychological principles relevant to disinformation. While many principles are tangentially related to this topic, we chose to focus on the four mechanisms critical to the absorption and spread of disinformation: initial information processing; cognitive dissonance theory; the group, belief, novelty (GBN) model; and the effect of emotions and arousal. For each principle, we explain the concept in plain language (with examples) before detailing its implications for disinformation and what can be done to counter its effects.

This report ends with a brief conclusion that highlights the importance of continuing this area of study. The conclusion summarizes how this study has advanced the conversation on countering the spread of disinformation and highlights additional questions to address.

Approach and sources

This report relies primarily on a thorough review of the (1) disinformation and (2) psychology literature to pull out the information relevant to the national security space. For the first two sections on the what and why of disinformation, our primary sources are reports from think tanks and academic institutions as well as news reports, with CNA’s previous studies on memes and bots serving as a foundation.⁴ We also included some foreign language sources in the

⁴ Vera Zakem, Megan McBride, and Kate Hammerberg, *Exploring the Utility of Memes for U.S. Government Influence Campaigns*, CNA, DRM-2018-U-017433-Final, 2018, https://www.cna.org/cna_files/pdf/DRM-2018-U-017433-Final.pdf; McBride, Gold, and Stricklin, *Social Media Bots: Implications for Special Operations Forces*.

discussion of adversary use of disinformation to better understand the place disinformation has come to hold in adversary thinking, doctrine, and strategy.

For the third section, we conducted an extensive literature review of psychology journals. Where relevant, we also included literature from other disciplines (e.g., economics, computer science, and marketing). Finally, some of the psychological principles related to modern, technology-enhanced disinformation have roots in foundational experimental psychology. Where applicable, we discuss the root theory first and then apply it to disinformation.

What Is Disinformation?

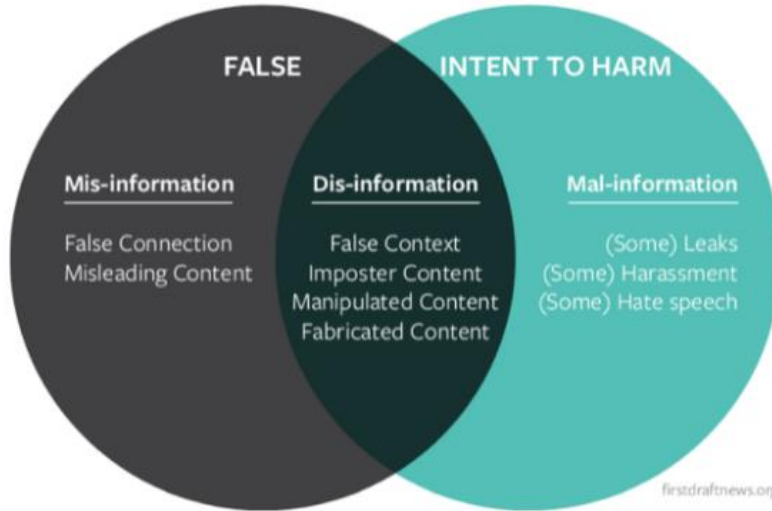
In the messy mainstream and social media ecosystem that characterizes the contemporary world, it is nearly impossible to untangle concerns about disinformation, misinformation, influence operations, adversary interference, propaganda, and online manipulation. Particularly problematic in the context of this report is the distinction between disinformation and misinformation. While these terms are often used interchangeably, they are distinct concepts easily differentiated by attention to the *intent of the content's creator* (Figure 4).

- *Disinformation* is information that is known to be false, and that is spread with the explicit goal of deceiving. In this case, the creator of the information (i.e., the tweet, the post, the email) *intends to deceive*.
- *Misinformation* is information that is false, but that is spread without a desire to deceive. In this case, the creator of the information (i.e., the tweet, the post, the email) has *no intention to deceive*.

It may also be helpful to add a third term to the mix:

- *Mal-information* is information that is known to be true, and that is spread intentionally, with the explicit goal of harming. One example might be sharing private information with the goal of harming the reputation of a politician. Mal-information is slightly different than disinformation and misinformation in that it is not necessarily false (it might contain false elements, but it might also be entirely true). We include it here, though, in an effort to provide a comprehensive oversight of the information ecosystem.

Figure 4. Taxonomy of disinformation, misinformation, and mal-information



Source: Temir Asanov, “Fake News in Modern News Media: Disinformation, Misinformation and Malinformation,” Medium.com, Mar. 17, 2019, <https://medium.com/@tasanoff/fake-news-in-modern-news-media-disinformation-misinformation-and-malinformation-e4fdfa2ab571>.

While this relatively simple taxonomy may seem to solve the problem, the reality is that the correct application of these labels (particularly those of disinformation and misinformation) is contingent on the tricky task of identifying the intent attached to the creation of the content.

To illustrate this, let’s assume, for example, that Alice, Bernard, Claire, and David are active Twitter users.

One day, Alice wakes up planning to go to the beach with friends, but her mother insists that she wear sunscreen. Alice doesn’t want to wear sunscreen because her friends don’t wear sunscreen, so she creates a fake graphic detailing the negative consequences of wearing sunscreen and posts it to Twitter with the hashtag #sciencefacts. She then shows it to her mother in an attempt to get out of wearing sunscreen. In this case, the stakes are relatively low and modest, but we could accurately label this content as disinformation because Alice’s post meets the criterion of that definition: *intentionally shared false information designed to deceive*.

A few days later, Bernard is on the subway and the people sitting across from him are having a heated debate about sunscreen (perhaps because they saw Alice’s post). Bernard finds the argument against wearing it compelling. When he gets home, he creates a graphic to share with his friends. Because Bernard’s graphic is not designed to deceive, it would be accurately

categorized as misinformation: *information that is false but that is spread without the intent to deceive*.

These examples are relatively straightforward because in both cases we know the intention of the individual who created the content. Classifying content becomes far more complicated, though, when information is retweeted, reposted, and forwarded.

For example, imagine that a few weeks later Claire is online and she finds a graphic detailing the negative consequences of wearing sunscreen. Concerned about the data in the graphic, she shares it via social media. While in some ways Claire does the same thing as Alice and Bernard (sharing a graphic via social media), we need to know the intention of the graphic's *creator* in order to accurately categorize it. If Claire shared Bernard's graphic, then her earnest post would be misinformation: *information that is false but that is spread without a desire to deceive*. If, however, Claire had accidentally stumbled across Alice's graphic, then her post would meet the definition of disinformation, and Claire would simply be an unwitting accomplice in Alice's disinformation campaign. The difference, again, is in the intention of the content's original creator.

Because this report focuses on the *psychology of how disinformation works* (and not on, for example, adversary intentions) the distinction between disinformation and misinformation is not critical. It's true that the *perceived* or *assumed* intent of the content's originator is important; we read material generated by comedians (who we assume intend to amuse us) differently from how we read material generated by news organizations (who we assume intend to educate us). These perceived and assumed intentions, though, are facilitated by a set of cognitive shortcuts that we use to negotiate the world. We see the *New York Times* or *Wall Street Journal* banners and assume that we know the content creator's intent. Sometimes disinformation takes advantage of these heuristics—as the Columbian Chemical hoax creators did by creating a fake screenshot of the CNN website. This is possible because most of the time all we have are our perceptions and assumptions; most people don't know the origins of the material they view and share online, or the true intention of the person who created that content. As a result, we receive disinformation and misinformation the same way.

To hammer this home, let's take David's perspective. David is an everyday social media user scrolling through Twitter. Assuming he doesn't know anything about Alice, Bernard, or Claire, the graphics that he sees in their Twitter feeds are exactly the same. The disinformation/misinformation distinction just isn't relevant for David (i.e., for the end user or consumer), and, psychologically, disinformation and misinformation exploit the same principles. Note that, in many cases, these same principles are also exploited by those attempting to share accurate information. News agencies, as one example, will also rely on these principles to ensure that you visit their sites, view their channels, and click their links. The principles themselves are neutral and normal. This report, therefore, will not distinguish

between disinformation and misinformation in discussing the relevant psychological principles; moreover, in an effort to avoid unnecessarily complicating things, we will use the term *disinformation* to refer to both throughout the remainder of this report.

Types of disinformation

While the definitions offered in the section above are accurate and useful, they fall short in capturing the range of material that analysts assess under the rubric of disinformation. A recent State Department-funded report, however, highlights the nuanced and complex nature of this phenomenon:

To understand the disinformation environment, it is useful to dissect the different elements it encompasses. Disinformation can include authentic material used in a deliberately wrong context to make a false connection, such as an authentic picture displayed with a fake caption. It can take the form of fake news sites or ones that are deliberately designed to look like well-known sites. Disinformation can further include outright false information, shared through graphics, images, and videos. It can also take the form of manipulated image and video content, where controversial elements are photoshopped into innocuous contexts to evoke anger or outrage.⁵

In other words, understanding the psychological impact of disinformation isn't merely about understanding why a simple graphic might trick a few people. Disinformation exists in a context—it exploits our cognitive shortcuts and heuristics in order to be effective—that can be fabricated to varying degrees.

To take just the first part of the definition from the State Department report above, disinformation consists of not only patently inaccurate information designed to appear true (such as the Columbian Chemicals hoax), but also accurate information that has been manipulated or taken out of context. As one example, on September 20, 2015, a pro-Russia media outlet falsely claimed that US Ambassador John Tefft had been spotted at an anti-government rally earlier in the day.⁶ In support of its claim, the website quoted Tefft as saying

⁵ Christina Nemr and William Gangware, *Weapons of Mass Distraction*, Park Advisors, 2019, accessed Feb. 2, 2021, <https://www.state.gov/wp-content/uploads/2019/05/Weapons-of-Mass-Distraction-Foreign-State-Sponsored-Disinformation-in-the-Digital-Age.pdf>.

⁶ "US Ambassador to Russia John F. Tefft Is Sent to Opposition Rally in Maryino," Посла США в России Джона Ф. Теффта отправили на митинг оппозиции в Марьино, Ren.TV, Sept. 20, 2015, <http://ren.tv/novosti/2015-09-20/posla-ssha-v-rossii-dzhona-f-teffta-otpravili-na-miting-oppozicii-v-marine>.

that he was attending the rally in order to assess the “caliber” of Russian democracy.⁷ More interesting, the post included a photo of Tefft mid-interview, with the opposition rally visible behind him (Figure 5). The media outlet even tweeted the photo with the following caption: “US Ambassador to Russia John Tefft strolled at an opposition rally in Marino.”⁸

Figure 5. Russian disinformation regarding US Ambassador Tefft



Source: “Посол США в России Джон Ф. Тефт прогулялся на митинге оппозиции в Марьино,” Sept. 20, 2015, <https://twitter.com/rentvchannel/status/645658877426593792>.

⁷ Some of this language remains in the REN TV article. Carl Schreck, “Photoshop Wars: U.S. Ambassador ‘Attends’ Russian Opposition Rally...and the Moon Landing,” Radio Free Europe/Radio Liberty, Sept. 21, 2015, accessed Feb. 3, 2021, <https://www.rferl.org/a/russia-photoshop-us-ambassador-tefft-opposition-rally-ren-tv/27260885.html>.

⁸ “The American Ambassador to Russia John F. Tefft Walked at an Opposition Rally in Maryino,” Посол США в России Джон Ф. Тефт прогулялся на митинге оппозиции в Марьино, Twitter, Sept. 20, 2015, <https://twitter.com/rentvchannel/status/645658877426593792>.

As the US Embassy in Russia promptly pointed out, though, the image was doctored. In fact, they tweeted a response that included the doctored image, the original image, and other doctored images of Ambassador Tefft at the moon landing and a hockey game (Figure 6).

Figure 6. US Embassy response to Russian disinformation regarding US Ambassador Tefft



Source: “The American Ambassador to Russia John F. Tefft Walked at an Opposition Rally in Maryino,” Посол США в России Джон Ф. Тефт прогулялся на митинге оппозиции в Марьино, Twitter, Sept. 20, 2015, <https://twitter.com/rentvchannel/status/645658877426593792>.

Notes: The original image (upper left); the Photoshopped image that REN TV tweeted (upper right); Tefft at the moon landing (lower left); Tefft at a hockey game (lower right). Translation of Twitter post: Ambassador Tefft spent yesterday's weekend at home. But thanks to Photoshop you can be anywhere. #fake #fake.

In this case, the disinformation very clearly consisted of “authentic material used in a deliberately wrong context to make a false connection” and a manipulated image in which an innocuous element (Ambassador Tefft at a news conference) was Photoshopped into a controversial environment (an opposition rally) in order to “evoke anger or outrage.” By

contrast, the Columbian Chemicals disinformation campaign is one that involved “fake news sites or ones that are deliberately designed to look like well-known sites” and “outright false information, shared through graphics, images, and videos.”

That said, the State Department definition—and, by extension, this report, as we have adopted the State Department definition—leaves out a number of techniques that are often included in discussions of disinformation campaigns but that are better classified as influence campaigns. It does not, for example, include instances in which a fake account or persona is used to spread true information, amplify the posts of a real person, galvanize a new community, or infiltrate a closed community to some ulterior end. There is no doubt an element of deception at play in these behaviors, but they don’t meet the definition of disinformation and they rely predominantly on a different set of core psychological mechanisms (some, but not all, of which are explored in this report).

Disinformation poses a critical threat, but there is no standard for how it will look, where it will come from, or what tools might be necessary to identify it. As seen in the State Department definition, it can take many different forms: it can be a skillfully doctored screenshot of the CNN website or a typo-laden meme with no references; it can come from a friend or a media outlet or a stranger; and it might be easily traced via a simple Google query or impossible to track to its source. This project thus takes a different tack, focusing not on how disinformation can be identified but on how it affects—consciously and unconsciously—the recipients that are exposed to it and on how we might mitigate these effects.

Why Disinformation Matters

Since the uncovering of widespread Russian meddling in the 2016 election, the US government, academics, think tanks, and social media companies have poured a wealth of resources into the study of disinformation. Every few months, Congress drags the leaders of the largest tech companies to Capitol Hill to grill them on their plans for countering the threat of disinformation; published reports on the issue now abound, tackling topics including Russian tactics, enabling technologies, and strategies for pushing back. Of interest, however, is that most of these resources seem to take for granted the fact that disinformation matters, rather than spelling out why that is the case. This chapter will endeavor to explain why we should care about disinformation, with a focus on the threat it poses in the national security context.

Adversary use of disinformation

While it is tempting to view adversary use of disinformation as a contemporary problem, the use of such means by adversaries attempting to shift the balance of political or military power is a long-standing issue. As just one example, in the first century CE, Octavian waged a fake news smear campaign—deploying primarily “short, sharp slogans written upon coins in the style of archaic Tweets”—to turn public opinion against Mark Antony and secure the role of first Roman Emperor.⁹ Disinformation, however, has always been contingent on its ability to circulate widely; operating centuries before the printing press, Octavian had to use coins to spread his message. Consequently, it should come as no surprise that adversary use of disinformation has become something of an international crisis in the social media age, when information spreads globally—leaping between continents and across language barriers—in mere seconds. It should also come as no surprise that our adversaries would seek to weaponize this powerful tool.

The objectives of disinformation are situationally specific. In discrete instances, disinformation might be used to obscure the purpose of a military campaign or cast aspersions on an enemy. The type of disinformation that the modern world is struggling with—even when that disinformation is coming from adversaries—is typically far less restrained or limited. The most commonly cited goals for adversary use of disinformation include causing chaos and confusion, sowing discord, distracting from an issue, casting doubt, and making the truth seem

⁹ Izabella Kaminska, “A Lesson in Fake News from the Info-Wars of Ancient Rome,” *Financial Times*, Jan. 17, 2017, accessed Feb. 3, 2021, <https://www.ft.com/content/aaf2bb08-dca2-11e6-86ac-f253db7791c6>.

unknowable.¹⁰ However, while threatening in themselves, these short-term objectives may also be part of a broader long-term strategy for gaining global influence, diminishing US influence, claiming great power status, securing a regime, and more.¹¹ As states have increasingly become aware of the benefits of information for achieving these objectives (including the ease with which they can harness the power of social media, the inexpensive nature of these efforts, and the relatively low risk of engaging in such behavior), disinformation and related types of information influence have come to hold a more prominent place in many adversaries' doctrine, strategy, and thinking.

One of the US's primary adversaries in this space, Russia has come to view the information space as one of the foundational areas in which states compete today.¹² Rather than seeing the use of information (including disinformation) merely as support for traditional military operations during a conflict, Russia views it as integral to and indistinguishable from conventional capabilities, with utility during every phase of conflict and even in peacetime.¹³ In fact, many Russian thinkers have asserted that the boundaries between war and peace are increasingly blurry, with the growing emphasis on nonmilitary means (such as information) asserted as one cause.¹⁴ In 2013, the Russian Chief of the General Staff Valeriy Gerasimov stated that the development of information weapons had the ability to reduce an adversary's combat potential.¹⁵ He further elaborated on this notion in 2019, writing that nonmilitary means are now the primary choice for attaining goals, with military force necessary only when nonmilitary means are unsuccessful.¹⁶ Gerasimov also expressed that, while military means are decisive once armed conflict begins, nonmilitary means make it possible to create the

¹⁰ Dean Jackson, "Issue Brief: How Disinformation Impacts Politics and Publics," National Endowment for Democracy, May 29, 2018, accessed Feb. 2, 2021, <https://www.ned.org/issue-brief-how-disinformation-impacts-politics-and-publics/>.

¹¹ Kasey Stricklin, "Why Does Russia Use Disinformation?," Lawfare, Mar. 29, 2020, accessed Feb. 2, 2021, <https://www.lawfareblog.com/why-does-russia-use-disinformation>.

¹² Daniel Kliman et al., *Dangerous Synergies*, CNAS, 2020, accessed Feb. 1, 2021, <https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/CNAS-Report-Dangerous-Synergies-May-2020-DoS-Proof.pdf?mtime=20200506164642&focal=none>.

¹³ Makhmut A. Gareev, "Anticipate Changes in the Nature of War: Every Era Has Its Own Kind of Military Conflict, and its Own Constraints, and its Own Special Biases," *Voyenno-Promyshlenny Kuryer Online*, June 5, 2013.

¹⁴ *Ibid.*

¹⁵ Valery V. Gerasimov, "The Value of Science is Foresight," *Ценность науки в предвидении*, ВПК, ВПК, Feb. 26, 2013, accessed Feb. 1, 2021, <https://vpk-news.ru/articles/14632>.

¹⁶ *Ibid.*; Valery V. Gerasimov, "Vectors of the Development of Military Strategy," *Векторы развития военной стратегии*, *Krasnaya Zvezda*, Красная звезда, Mar. 4, 2019, accessed Feb. 1, 2021, <http://redstar.ru/vektory-razvitiya-voennoj-strategii/?attempt=1>.

conditions and influence the operating environment to make the employment of military means more effective.¹⁷ In fact, it is thought that the use of such nonmilitary means as information can even preclude an armed conflict by allowing Moscow to assert influence and shape internal dynamics within an adversary state, creating conditions favorable to Russia.¹⁸

Another US adversary in this space, China has long sought to exercise control over information that circulates regionally, and has recently begun to extend its influence beyond its borders and neighbors.¹⁹ In the last decade, the Chinese Communist Party has sought opportunities to shape the digital information space and, in 2013, Chinese President Xi Jinping stated that the use of innovative techniques to spread narratives positive for China, and promoting the Chinese view globally, was a priority.²⁰ The Chinese government has historically employed these tactics in regions close to its borders, such as Taiwan, to attempt to shape the geopolitical situation in a favorable manner.²¹ These efforts were typically overt propaganda efforts to push Beijing's preferred narrative.²² After the onset of the coronavirus pandemic, China began to take a page out of the Russian playbook, employing more covert disinformation on social media to obscure the virus's roots and shift the blame, typically to the US.²³ While it is unclear whether Beijing will continue to employ these methods after the pandemic has passed, the Chinese government has certainly evolved its thinking on how to use information for influence and the utility of disinformation in reinforcing the government's reputation.

Both Iran's conventional military and the Islamic Revolutionary Guard Corps (a parallel military structure tasked with upholding the ideals of the revolution) underscore the importance of informational control for offensive and defensive purposes alike.²⁴ Iran sees

¹⁷ Gerasimov, "Vectors of the Development of Military Strategy."

¹⁸ Kliman et al., *Dangerous Synergies*.

¹⁹ Ibid., p. 5.

²⁰ Ibid.

²¹ Joshua Kurlantzick, "How China Ramped Up Disinformation Efforts During the Pandemic," Council on Foreign Relations, Sept. 10, 2020, accessed Feb. 2, 2021, <https://www.cfr.org/in-brief/how-china-ramped-disinformation-efforts-during-pandemic>.

²² Sarah Cook, "Welcome to the New Era of Chinese Government Disinformation," *The Diplomat*, May 11, 2020, accessed Feb. 1, 2021, <https://thediplomat.com/2020/05/welcome-to-the-new-era-of-chinese-government-disinformation/>.

²³ "How China Ramped Up Disinformation Efforts During the Pandemic."

²⁴ "IRGC (Islamic Revolutionary Guard Corps)," Counter Extremism Project, accessed Feb. 3, 2021, <https://www.counterextremism.com/threat/irgc-islamic-revolutionary-guard-corps>; Emerson Brooking and Suzanne Kianpour, *Iranian Digital Influence Efforts: Guerrilla Broadcasting for the Twenty-First Century*, Atlantic Council, 2020, accessed Feb. 1, 2021, <https://www.atlanticcouncil.org/wp-content/uploads/2020/02/IRAN-DIGITAL.pdf>.

itself as constantly on the defensive from information efforts emanating from other countries' broadcasters, lobbyists, and so on.²⁵ A recent report on Iran's digital disinformation efforts describes them as "public diplomacy under duress" and notes that pro-Iranian messaging proves most effective when it appears to come from a neutral third party.²⁶ Iran thus endeavors to set up a narrative structure that its adversaries cannot easily meddle with or take down in order to achieve its objectives, which include positioning itself as a leader in the Muslim world and serving as a bastion against perceived US and western regional intervention.²⁷

The pandemic has seen these three adversaries increasingly amplify each other's disinformation, helping China's recent disinformation efforts gain a wider reach than they likely would otherwise. Despite varied objectives, the shared goal of undermining US influence has caused something of a convergence as Russia, China, and Iran work together to diminish the influence of their common adversary. For its part, the US government views disinformation in a patently different way than its adversaries. While many adversaries are increasingly emphasizing the utility of information for achieving a range of objectives, the US is grappling with a range of ethical and legal implications that simply are not issues for its adversaries in this space. As a result, it still does not have a clear strategy for countering information operations, and its efforts in this space tend to be largely reactive.²⁸ In addition, because the US views information influence through a different lens than many of its adversaries, the role of such efforts in adversary strategy and thinking is not always well understood. Therefore, the study of disinformation is important for ensuring that the US does not find itself unwittingly behind or vulnerable in a key area of competition.

Domestic use of disinformation

It would be naïve to write a paper on disinformation—especially one grappling with its real-world impacts—without acknowledging that domestic US actors are active in this space as well. As with adversary use of disinformation, the objectives for domestic actors using disinformation vary with prominent examples including attempts to shift political discourse on a range of controversial issues, such as election integrity, medical advice, and climate change. The objectives may also be motivated by a desire to garner support—measured in the form of votes, subscriptions, or clicks. Again, though, this is hardly a new phenomenon. In 1835,

²⁵ Brooking and Kianpour, *Iranian Digital Influence Efforts: Guerrilla Broadcasting for the Twenty-First Century*.

²⁶ *Ibid.*

²⁷ *Ibid.*

²⁸ Doowan Lee, "The United States Isn't Doomed to Lose the Information Wars," *Foreign Policy*, Oct. 16, 2020, accessed Feb. 3, 2021, <https://foreignpolicy.com/2020/10/16/us-election-interference-disinformation-china-russia-information-warfare/>.

The Sun (a newspaper in New York City) published a series of six articles claiming that life—including unicorns—had been observed on the moon by a one of the world’s foremost astronomers. The hoax wasn’t uncovered for weeks, and the newspaper never printed a retraction, but circulation did allegedly increase. The discovery of the hoax did not put an end to the practice; almost a decade later, *The Sun* again published fake science news. In this case, though, it was famed author Edgar Allen Poe who “sold an ingenious scientific hoax to a newspaper publisher for fifty dollars,” though this time the newspaper apparently issued a retraction.²⁹

In contemporary America, disinformation circulates on a variety of issues ranging from the absurd (e.g., a 2017 story that Hollywood elites were using “the blood of babies to get high”) to the deadly serious (such as that related to the supposed dangers of the COVID-19 vaccine, which could deter large quantities of people from getting the vaccine).³⁰ The reality, though, is that the psychological mechanisms that allow disinformation to spread are the same, despite the motives of the individual behind such content or its potential impacts.

Is disinformation effective?

Given the variety and prevalence of disinformation, the existence of disinformation is not particularly debatable. Rather, the critical question remaining is whether disinformation is actually effective at changing peoples’ minds and helping adversaries achieve their objectives. This effectiveness is incredibly difficult to assess; after all, it is difficult to know whether someone would have taken a certain action or had a certain thought even without the presence of disinformation.³¹

Several recent studies looking into the 2016 presidential election have attempted to answer the effectiveness question, often with disparate results. A 2018 study released by researchers from Ohio State concluded that disinformation likely helped Donald Trump secure victory because around 4 percent of Barack Obama’s 2012 supporters did not vote for Hillary Clinton

²⁹ “Edgar Allan Poe—‘The Balloon Hoax,’” The Edgar Allan Poe Society of Baltimore, Last updated July 15, 2020, <https://www.eapoe.org/works/info/pt049.htm>.

³⁰ David Mikkelson, “Did Keanu Reeves Say Hollywood Elites Use the ‘Blood of Babies’ to Get High?,” Snopes, Nov. 22, 2017, accessed Feb. 1, 2021, <https://www.snopes.com/fact-check/keanu-reeves-blood-drinking/>; Lois Beckett, “Misinformation ‘Superspreaders’: Covid Vaccine Falsehoods Still Thriving on Facebook and Instagram,” *The Guardian*, Jan. 6, 2021, accessed Feb. 3, 2021, <https://www.theguardian.com/world/2021/jan/06/facebook-instagram-urged-fight-deluge-anti-covid-vaccine-falsehoods>.

³¹ “Issue Brief: How Disinformation Impacts Politics and Publics.”

after reading false stories about her.³² By contrast, researchers from the Stanford Institute for Economic Policy Research released findings in 2017 showing that only a small portion of Americans saw the most widely circulated fake stories, suggesting disinformation during the election may not have been that convincing or influential, though the authors did not offer an opinion on whether this disinformation ultimately had an impact on the election outcome.³³ In January 2020, researchers released a study purporting to show that interaction with trolls from the Russian IRA during the 2016 election did not have any impact on Americans' political behaviors or thoughts because the trolls interacted most with individuals who already held strongly partisan attitudes and were unlikely to change their thinking.³⁴ However, the study's small sample size means that more research is needed on this topic for findings to be taken as conclusive. Therefore, none of these studies are decisive for answering the question of disinformation effectiveness. In addition, none of these studies explored the question of whether the disinformation campaigns might have affected voter turnout (i.e., perhaps suppressing Clinton voters and energizing Trump voters).

Most alarmingly, even if consumers of disinformation do not believe the false stories they read or change their attitudes or behaviors as a result, the very existence of disinformation or perception that it is spreading can help adversaries achieve their objectives. The perception that adversaries are excelling in the field of disinformation and may be manipulating events could help promote adversary goals by introducing doubt or anxiety about government institutions, journalistic outlets, and other staples of democracy. Russia, for example, often promulgates a number of narratives on unfolding events, not because it hopes that audiences will believe all of the various, often contradictory, messages, but because doing so makes the truth seem unknowable, and makes citizens doubt the official government versions of events. Russia hopes the long-term effect will lead to democratic systems appearing less appealing, engendering a loss of US influence worldwide, while making political systems similar to the one found in Russia seem stronger and more effective in contrast.³⁵

³² Aaron Blake, "A New Study Suggests Fake News Might Have Won Donald Trump the 2016 Election," *The Washington Post*, April 3, 2018, accessed Feb. 1, 2021, <https://www.washingtonpost.com/news/the-fix/wp/2018/04/03/a-new-study-suggests-fake-news-might-have-won-donald-trump-the-2016-election/>.

³³ Krysten Crawford, "Stanford Study Examines Fake News and the 2016 Presidential Election," *Stanford News*, Jan. 18, 2017, accessed Feb. 2, 2021, <https://news.stanford.edu/2017/01/18/stanford-study-examines-fake-news-2016-presidential-election/>.

³⁴ Christopher Bail et al., "Assessing the Russian Internet Research Agency's Impact on the Political Attitudes and Behaviors of American Twitter Users in Late 2017," *PNAS: Proceedings of the National Academy of Sciences of the United States of America* 117, no. 1 (2020), accessed Feb. 3, 2021, <https://www.pnas.org/content/117/1/243>.

³⁵ "Why Does Russia Use Disinformation?"

While it is unclear if disinformation can actually change an individual's opinion, attitude, or vote on a specific issue, there are many examples where disinformation has affected behaviors. The psychological principle of cognitive dissonance (to be discussed later) posits that one's actions have a powerful effect on one's attitudes since you need to rationalize to yourself why you are doing something. Furthermore, it has been shown to affect the national discourse on a variety of topics and to even sometimes make its way into the mainstream media. This is at least partially because disinformation on social media has proved to spread more widely and more rapidly than accurate information. A large 2018 MIT study of Twitter data found that, according to every normal metric, disinformation wins out over the truth on social media by diffusing "significantly farther, faster, deeper, and more broadly than the truth in all categories of information."³⁶ The authors concluded that this was true because of human nature, rather than the use of bots, because the bots studied in the report amplified just as many true stories as false stories.³⁷

In some cases, the promulgation of disinformation has spurred real-world events, and even violence. Countries such as India and Burma have seen social media used to spread disinformation on religious minorities that led to actual violence.³⁸ An additional salient example was the 2005–2006 rallies in response to the publication of images of Muhammad in a series of European newspapers, which resulted in violence around the world, leading to over 100 deaths and 800 injuries. An often-overlooked component of this dynamic, though, was an informal publication known as the Akkari-Laban dossier. This document, written by Danish-Muslim clerics, was shared widely across the Muslim world. While experts later confirmed that much of its content was accurate, they also noted that it contained inflammatory and false information. Specifically, the dossier implied that one of the images it contained (of a picture of a man dressed as a pig) was meant to depict Mohammad when, in fact, it was an Associated Press photo of a French pig-squealing contest.³⁹ The dossier did contain accurate images of cartoons as well, but the inclusion of this especially inflammatory image makes it an example of disinformation. It is an instance in which "authentic material" (i.e., the Associated Press image) was "used in a deliberately wrong context to make a false connection" (i.e., implying that the image was part of a larger effort to disrespect Islam).⁴⁰ Again, it is impossible to know if this particular image—one among many—was directly responsible for causing the rallies,

³⁶ Soroush Vosoughi, Deb Roy, and Sinan Aral, "The Spread of True and False News Online," *Science* 359, no. 6380 (2018), <https://science.sciencemag.org/content/359/6380/1146>.

³⁷ Ibid.

³⁸ "Issue Brief: How Disinformation Impacts Politics and Publics."

³⁹ Martin Asser, "What the Muhammad Cartoons Portray," BBC, Jan. 2, 2010, http://news.bbc.co.uk/2/hi/middle_east/4693292.stm.

⁴⁰ Nemr and Gangware, *Weapons of Mass Distraction*.

violence, injuries, and death. Its inclusion, though, is widely recognized to have exacerbated an already delicate situation (because Islam identifies the pig as an unclean animal) and thus contributed to significant real-world consequences.

Summary

It is not always immediately clear exactly why disinformation matters for national security or whether the spread of disinformation is even having an effect. While much has been done to map out adversary disinformation tactics and the digital space that enables them, this work often takes for granted the fact that we should be studying disinformation in the first place. There are, however, demonstrable reasons why it is important to pay attention to disinformation. As adversaries and domestic actors alike increasingly turn to information means as a way to achieve a range of objectives, the study of disinformation is also growing more important to ensure that malign actors do not catch the US flat footed. While it is difficult to prove that disinformation is effective in directly changing minds or votes, there is much evidence it can affect behaviors and discourse. In addition, the lack of conclusive study on whether disinformation is effective means that we cannot discount it; just as we do not know exactly how effective disinformation is, we also cannot prove that it is *not* effective. The issue, therefore, requires greater attention and study.

That study should center not just on disinformation tactics but also on how those tactics affect the mind since disinformation's primary impact occurs in the mind. It is by definition a method of exercising psychological influence. Those spreading disinformation hope that a recipient will believe it, thus allowing the promulgators to achieve their objectives. As mentioned, though, disinformation can have an impact even if it is not believable; the mere perception that disinformation exists may be enough to lead to certain effects, such as the erosion of trust in democratic institutions and the media. This perception, of course, is a mental process as well.

The impact of disinformation, however, is not *limited* to the mind. It can have concrete and critical secondhand effects beyond the foregoing examples of real-world violence. For example, the erosion of trust in democratic institutions can affect voting behavior, increased suspicions about medical care can change willingness to receive that care, and distrust of the police can lead to increased resistance to arrest. As a result, it is important to understand not merely the technical issues of how disinformation is transmitted or how it can be prevented—issues that have received considerable attention over the past few years—but also the psychological question of how disinformation works and how it can be countered. Increased familiarity with the psychology of disinformation is important for understanding how the spread of false information aids in the achievement of objectives and, in turn, how to prevent its harmful impacts. Moreover, because countering the effects of disinformation is also a cognitive process, understanding the psychological principles that make disinformation effective can assist in

understanding what principles may counter it. As the use of disinformation and other types of psychological influence become more important to the achievement of malign objectives for both adversaries and domestic actors, understanding how to counter these tactics more efficiently is a national security imperative.

Psychological Principles That Can Facilitate Disinformation Adoption and Spread

This section addresses why people are vulnerable to disinformation by examining innate psychological mechanisms that govern how people process and share information. The primary thesis of this section is that disinformation is processed and shared through *routine and ordinarily adaptive psychological processes*.⁴¹

We begin by describing the psychological principles associated with initial information processing and the subsequent reinforcement of initial judgments through cognitive dissonance. Next, we describe how group membership, beliefs, and novelty affect the way we process and communicate information. Finally, we describe the role of emotion and arousal in people's tendency to share information. In each section, we describe the principles, apply them to disinformation specifically, and describe how normal, routine, and adaptive processes can be used to further spread disinformation. We also describe how these principles can be used to counter the spread of disinformation. Table 1 summarizes these psychological principles, provides a brief explanation, and shows the connection between that concept and disinformation.

⁴¹ Here, the word *adaptive* is used in an evolutionary context. Behavior is adaptive if it helps a person to accomplish a goal.

Table 1. Psychological principles relevant for disinformation adoption and spread

| Principle | Explanation | Application to disinformation |
|---------------------------------------|--|--|
| Initial information processing | We process information as efficiently as possible and that can make us vulnerable to mistakes. | We can accept disinformation as true because we aren't thinking deeply and critically. |
| Cognitive dissonance | When we are confronted with something that goes against our beliefs, we are motivated to resolve the conflict. | We accept disinformation that supports our initial beliefs and try to reject information that disconfirms our initial beliefs. |
| Group, belief, and novelty | We more readily share information with people with whom we identify, when we believe it is true, and when it is novel or urgent. | We accept and share disinformation more readily when it comes from people we know, it appeals to what "our group" believes, and when we think it is new. |
| Emotions and arousal | We pay more attention to information that makes us feel positively or arouses us to act. | We are more likely to share disinformation if it is constructed to elicit high-arousal emotions. |

Source: CNA.

Humans are constantly processing information that comes in through our senses. Encoding, cataloging, deciding on action, and storing all that information could be a full-time job without mechanisms to efficiently process it all. Indeed, we don't have the "computing power" to appropriately process each new piece of information as entirely novel. This section describes multiple psychological theories and their effects. They all come down to the same bottom line: our brains have adaptive mechanisms to triage and organize information as quickly and efficiently as possible.

Initial information processing (dual process theory)

Dual process theory is the foundation for understanding how people process information.⁴² The theory posits that we have two mechanisms to evaluate new information that comes in through our senses. Table 2 describes the key elements of these two processes: automatic thinking and controlled thinking.

⁴² Shelly Chaiken and Yaacov Trope, *Dual-Process Theories in Social Psychology*, (New York: Guilford Press, 1999); Susan T. Fiske and Shelly E. Taylor, *Social Cognition: From Brains to Culture*, 3rd ed. (London: SAGE Publishing, 2016).

Table 2. Dual process theory

| Process | Characteristics | How it works |
|----------------------------|---|--|
| Automatic thinking | Fast, effortless, uses few resources | People rely on previous experience with similar information and best guesses (heuristics), easy-to-process information (fluent), and information that is easy to remember to determine how to think about new information and act on it with very little “thought” |
| Controlled thinking | Slow, deliberate, requires many cognitive resources | People rely on analytical and deep thinking involving questioning assumptions to understand and act upon new information |

Source: CNA.

The first process, called *automatic*, requires very little effort and is fast, efficient, and “cognitively cheap.” Essentially, when the information comes in, it is immediately associated with something already known and is dealt with or “filed away” appropriately. For example, if someone were driving the same road to work every day and came upon their exit sign, they would automatically turn on their signal and proceed to the exit ramp. This action requires very little thought. Kahneman argues that attention requires effort and that we conserve mental energy by focusing on only a few things and letting automatic processes handle the rest.⁴³

The second process is called *controlled*. Controlled processes are slower, more cognitively involved, and require deliberate attention. Going back to the driving example, if the exit you were expecting to see was closed, your brain would begin controlled processes to determine alternative routes for getting to work. While the theory posits two processes, researchers acknowledge that there is really a gradient from the most automatic to the most controlled processes.⁴⁴ Fully automatic processes are those that are completely unintentional and happen with little or any cognitive awareness. As processes become more controlled, your brain recognizes novel information and attends more deeply to it. And for processes that are more fully controlled, people are more likely to seek additional information and think about whether the information they have is accurate.

⁴³ Daniel Kahneman, *Thinking, Fast and Slow*, (New York: Farrar, Straus, and Giroux, 2011).

⁴⁴ Fiske and Taylor, *Social Cognition: From Brains to Culture*.

The amount of information that people can respond to at any given time is finite,⁴⁵ so they default to the processing option that requires the lowest level of effort.⁴⁶ This makes our brains highly efficient because we do not waste resources thinking more deeply about things than we need to. There is a downside, however: we can make mistakes. Returning to the exit ramp example, if the exit sign that we see every day is blocked or closed, using the automatic processes to proceed as usual can be dangerous. If we fail to recognize the blocked exit information as important to attend to, we can make mistakes.

Heuristics are one of the primary means used to determine the appropriate level of attention to give information. Heuristics are mental shortcuts we use to make sense of the world.⁴⁷ Heuristics are developed through previous experience and connections we make between similar pieces of information. For example, if you have lived in a given climate for several years, you have a heuristic about what the likely temperature will be in January. Most of the time, you will be close to correct and you can dress for the day without checking the temperature. However, *some* days you will be wrong—it will be abnormally warm or cold—and your choice of clothes will be inappropriate. By definition, heuristics are correct (or close enough) most of the time, but can lead to huge errors at other times.

In addition to heuristics, information fluency affects our likelihood of processing information and the depth at which we process it. Information fluency refers to how easily people process information. Processing is more fluent if the information is easier to understand (this includes font, color contrast, accent, and cadence in addition to actual message content).⁴⁸ In addition, information that is easier to understand is more likely to be incorporated into what someone believes, regardless of whether it is true.⁴⁹ Information that is easy to understand “feels right” and so it seems true.⁵⁰ Once that information is incorporated into our scheme, it becomes

⁴⁵ Researchers demonstrated that individuals can reliably recall five to nine pieces of information reliably. Beyond that, they are inconsistent or unable to accurately recall. (T. L. Saaty and Mijgan Sagir Özdemir, “Why the Magic Number Seven Plus or Minus Two,” *Mathematical and Computer Modelling* 38 (2003), doi: 10.1016/S0895-7177(03)90083-5).

⁴⁶ Fiske and Taylor have coined the term *cognitive miser* to describe the tendency to process information at the lowest acceptable level of effort.

⁴⁷ Daniel Kahneman and Amos Tversky, “On the Psychology of Prediction,” *Psychological Review* 80, no. 4 (1973).

⁴⁸ Christian Unkelbach, “Reversing the Truth Effect: Learning the Interpretation of Processing Fluency in Judgments of Truth,” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33, no. 1 (2007), doi: 10.1037/0278-7393.33.1.219.

⁴⁹ *Ibid.*

⁵⁰ Tommy Shane, “The Psychology of Misinformation: Why We’re So Vulnerable,” *First Draft News*, June 30, 2020, <https://firstdraftnews.org/latest/the-psychology-of-misinformation-why-were-vulnerable/>.

previous knowledge that we can use to build more heuristics and make future judgments about new information.

Another important cognitive psychological principle that influences what we remember is the primacy and recency effect.⁵¹ It is easier for people to remember information that is presented first (primacy) and presented most recently.⁵² Studies demonstrating the primacy and recency effect traditionally focused on which words in a list people could remember. The primary and recency effect was displayed when people were more likely to remember the first few and the last few words presented. Studies have also shown that the primacy effect is particularly important for determining what website links people choose.⁵³ People most often click the links listed first in a search, for example.

Implications of initial information processing for disinformation

The phenomenon of spreading disinformation may appear to be the result of ideology and deep commitment to the material. However, some research suggests that's not the case. Pennycook and Rand demonstrated that individuals who were more prone to think deeply about issues (based on their Cognitive Reasoning Test score) were more able to discern fake from real news and were *more* discerning regarding information that was ideologically consistent than information that was ideologically inconsistent.⁵⁴ The implication of this work is that susceptibility to fake news is driven more by "lazy thinking" than it is by entrenched belief structures. Therefore, if people thought more deeply (i.e., used more controlled and less automatic processes), they would be less likely to believe and share disinformation.

Routine initial information processing mechanisms are adaptive and help people to quickly organize information, make sense of it, and decide on action to take. However, the efficiency in these processes leaves people vulnerable to making mistakes and processing information incorrectly. The default automatic processing, relying on heuristics and information fluency, leaves us with "blind spots" that prevent us from recognizing when we need to slow down to distinguish information from disinformation.⁵⁵ This results in credible information and

⁵¹ Bennet B. Murdock, Jr., "The Serial Position Effect on Free Recall," *Journal of Experimental Psychology* 64, no. 5 (1962), doi: <https://doi.org/10.1037/h0045106>.

⁵² Cong Li, "Primacy or Recency Effect? A Long-Term Memory Test of Super Bowl Commercials," *Journal of Consumer Behavior* 9, no. 1 (2009).

⁵³ Jamie Murphy, Charles Hofacker, and Richard Mizerski, "Primacy and Recency Effects on Clicking Behavior," *Journal of Computer-Mediated Communication* (2020).

⁵⁴ Gordon Pennycook and David G. Rand, "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning than by Motivated Reasoning," *Cognition* 188 (July 2019), doi: <https://doi.org/10.1016/j.cognition.2018.06.011>.

⁵⁵ Kahneman, *Thinking, Fast and Slow*.

disinformation being treated the same and can lead one to miscategorize disinformation as true.

Automatic thinking makes it likely that miscategorized information will be used. Familiarity plays in our initial information processing. Rather than rely on controlled processes to evaluate the veracity of new information, if the new information is similar to something we are familiar with, we might use a heuristic and assimilate the new information just as we did the old information. The Columbian Chemicals example, described in the introduction, demonstrates how the natural tendency to use automatic thinking could have influenced the spread of that fake story. As noted, the story spoofed the CNN website banner and logo. People familiar with the CNN logo and reputation as a news outlet were likely to use those heuristics to determine that *this* CNN story was as credible as others. Therefore, they may have shared it to inform others of the supposed harm from the explosion.

Information that is more familiar to us (repeatedly presented) is more likely to be recalled as true.⁵⁶ This is called the illusory truth effect, when familiarity with something gives the illusion that it is more accurate, and that inadvertently increases someone's belief in that information.⁵⁷ This can be especially problematic when attempting to correct disinformation because the correction often repeats the original claim.⁵⁸ Swire provides a clear and illustrative example of this challenge:

Truthfully stating that playing Mozart to your child will *not* boost its IQ mentions the two concepts of "Mozart" and "increased IQ," thereby making the link between the concepts more familiar even though the statement seeks to dispel the Mozart-IQ myth. This inadvertent increase in familiarity may reduce the effectiveness of the correction and may thus contribute to the continued influence effect of misinformation [emphasis in original].⁵⁹

⁵⁶ Nicholas DiFonzo et al., "Validity Judgments of Rumors Heard Multiple Times: The Shape of the Truth Effect," *Social Influence* 11 (2016), doi: <https://doi.org/10.1080/15534510.2015.1137224>.

⁵⁷ Ian Maynard Begg, Ann Anas, and Suzanne Farinacci, "Dissociation of Processes in Belief: Source Recollection, Statement Familiarity, and the Illusion of Truth," *Journal of Experimental Psychology: General* 121, no. 4 (1992), doi: <https://doi.org/10.1037/0096-3445.121.4.446>.

⁵⁸ Some researchers have argued for a "truth sandwich" to counter this phenomenon. In a truth sandwich, you would say the truth, refute each point in the disinformation, then repeat the truth. This puts the truth first and last, using the cognitive principles of primacy and recency. It also mitigates the tendency for repeating the lie to make it more accepted. However, the usefulness of a truth sandwich in combating disinformation has not yet been empirically investigated to our knowledge.

⁵⁹ Briony Swire, Ulrich K. H. Ecker, and Stephan Lewandowsky, "The Role of Familiarity in Correcting Inaccurate Information," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43, no. 12 (2017), doi: [10.1037/xlm0000422](https://doi.org/10.1037/xlm0000422).

Automatic processing does not provide deep consideration of information (source, context, etc.). Thus, when processing automatically, individuals might be more likely to recall repeated information without the context of *why* it was repeated to them. This causes errors in judgment. When those errors in judgment allow us to encode disinformation as true, it is very challenging to undo.

Once we make initial judgments about information, they become part of our scheme and how we understand the world. If that information is inaccurate (because of disinformation or a mistake in processing or remembering), we can be left with encoded inaccurate information to draw from later. There is a memory effect associated with the likelihood of recalling disinformation as opposed to the idea that the disinformation was debunked. The continued influence effect describes how corrected disinformation can continue to influence memory and reasoning.⁶⁰ Examples are the persistence of people's belief that Michael Jackson invented the moonwalk, Edison the lightbulb, and Guillotin the guillotine.⁶¹ Essentially, this theory posits that disinformation, once acquired, is very difficult to correct (unlearn). Corrections through debunking and fact-checking can "fail" to correct because the disinformation can be recalled later as fact separate from the debunked information. Essentially, both pieces of information are stored in a person's memory: the disinformation and the disinformation plus the correction. A correction fails when the disinformation is recalled without the correction. Swire, Ecker, and Lewandowsky found that effective corrections (those that have a lot of detail and affirm the facts) can "wear off" in a week and the participants will retell the inaccurate information.⁶² Dual process theory and the tendency toward automatic processing may be partially responsible for this phenomenon. If the disinformation was easier to understand and recall, or more familiar than the correction information, then when confronted with new information, we are more likely to rely on the disinformation.

What can we do about it?

Given the difficulty of correcting disinformation once present in a person's memory, it is more fruitful to attempt to *prevent* disinformation from taking root than to *correct* disinformation that has been encoded as truth. Here we focus on two relevant avenues for combating disinformation: the first is bolstering our own internal toolkit, and the second is helping those around us. Table 3 summarizes the prevention mechanisms described in detail in this section.

⁶⁰ Ulrich K. H. Ecker et al., "Correcting False Information in Memory: Manipulating the Strength of Misinformation Encoding and its Retraction," *Psychonomic Bulletin & Review* 18 (2011), doi: 10.3758/s13423-011-0065-1.

⁶¹ Bianca Pellegrino, "People Famous For Doing Things They Didn't Do," ListVerse, Sept. 8, 2019, <https://listverse.com/2019/09/08/top-10-people-famous-for-doing-things-they-didnt-do/>.

⁶² Swire, Ecker, and Lewandowsky, "The Role of Familiarity in Correcting Inaccurate Information."

Table 3. Disinformation prevention mechanisms for initial information processing

| | Concept | Description |
|--------------------------------|-------------------|---|
| Internal (with yourself) | Skepticism | Recognizing that people might have hidden agendas, scrutinize information presented to you |
| | Alertness | Warnings about the effects of disinformation can make you more alert to the possibility |
| | Analytic thinking | Purposefully make deliberate, controlled, thoughtful evaluations of new information |
| External (with others) | Create friction | Asking questions that make others process information more deliberately (e.g., asking, "what makes you say that?") |
| | Inoculation | Provide awareness of possible disinformation so others can recognize it when confronted by it later |
| | Nudges | Provide subtle cues of the behavior you want someone to exhibit; in this case, processing information more deliberately |

Source: CNA.

Bolstering your resistance to disinformation

The first preventive mechanism we describe involves building resistance to disinformation in your own information processing. There are three concepts, all of which can be taught, that guard against the inadvertent adoption of disinformation: skepticism, alertness, and analytic thinking.⁶³ All three of these processes for individually combating disinformation require the person to move from a greater reliance on automatic to a greater reliance on controlled processes.

Skepticism is the awareness of possible hidden agendas and a personal desire to understand the evidence.⁶⁴ It works to fight disinformation because skepticism requires the person to use more cognitive resources to evaluate information.⁶⁵ Research indicates that skepticism can be encouraged and trained through exposure to and discussion of pseudoscience and false advertising campaigns.⁶⁶

⁶³ Tommy Shane, "The Psychology of Misinformation: How to Prevent It," First Draft News, July 27, 2020, <https://firstdraftnews.org/latest/the-psychology-of-misinformation-how-to-prevent-it/>.

⁶⁴ Ruth Mayo, "Cognition Is a Matter of Trust: Distrust Tunes Cognitive Processes," *European Review of Social Psychology* 26 (2015), doi: <https://doi.org/10.1080/10463283.2015.1117249>.

⁶⁵ Briony Swire and Ulrich K. H. Ecker, "Misinformation and Its Correction: Cognitive Mechanisms and Recommendations for Mass Communications," in *Misinformation and Mass Audiences*, ed. Brian Southwell, Emily A. Thorson, and Laura Sheble (Austin: University of Texas Press, 2018).

⁶⁶ Rodney Schmalz and Scott Lilienfeld, "Hauntings, Homeopathy, and the Hopkinsville Goblins: Using Pseudoscience to Teach Scientific Thinking," *Frontiers in Psychology* 5 (2014), doi: <https://doi.org/10.3389/fpsyg.2014.00336>.

Alertness is a heightened awareness to the effects of disinformation.⁶⁷ Warning people about the effects of disinformation—such as the continued influence effect (i.e., the way that once disinformation becomes part of our schema for processing the world, it can be difficult to correct)—can reduce the continued reliance on disinformation, but it will not eliminate it.⁶⁸ Research shows that the *specific* description of the continued influence effect had the greatest impact for countering disinformation. More general warnings about the possibility of publishing unverified information are not as effective as a specific warning.

Analytic thinking involves deliberate, controlled, thoughtful evaluation of information as opposed to quick, automatic judgments. Bago, Rand, and Pennycook performed an experiment in which they asked people to make initial (automatic) judgments under time constraints while doing other tasks.⁶⁹ Then they asked the participants to reconsider the information without time constraints. They found that deliberation (controlled, analytic thinking) was associated with more accurate perceptions than initial judgments (automatic) made under constrained conditions. This finding indicates that, with deliberation, people can more accurately discern the true information from the false, even if they first made incorrect initial judgments.

Bolstering resistance to disinformation in others

The second avenue to combat disinformation believed because of initial processing errors involves attempts to influence other people's encoding of disinformation. Similar to the individual and internal mechanisms described previously, these techniques are based on the premise that deeper thinking can overcome fallacies in initial, automatic processes. Three related techniques can be used to interrupt the automatic information processing of others: creating friction, inoculation, and nudges.⁷⁰

Friction is the opposite of fluency. Whereas fluency consists of information that is easy to process, understand, and incorporate, friction occurs when something is difficult to process. Fazio found that adding friction to processing by asking people to explain why they think a headline is true or false reduced their likelihood of sharing stories with false headlines and had

⁶⁷ "The Psychology of Misinformation: How to Prevent It."

⁶⁸ Ulrich K. H. Ecker, Stephan Lewandowsky, and David T. W. Tang, "Explicit Warnings Reduce But Do Not Eliminate the Continued Influence of Misinformation," *Memory and Cognition* 38 (2010), doi: 10.3758/MC.38.8.1087.

⁶⁹ Bence Bago, David G. Rand, and Gordon Pennycook, "Fake News, Fast and Slow: Deliberation Reduces Belief in False (But Not True) News Headlines," *Journal of Experimental Psychology: General* 149, no. 8 (Aug. 2020), doi: 10.1037/xge0000729.

⁷⁰ "The Psychology of Misinformation: How to Prevent It."

no effect on the likelihood of sharing stories with true headlines.⁷¹ The implication here is that, if someone is asked to justify the veracity of something prior to sharing it, they are *less likely* to spread disinformation. Creating friction increases the depth at which individuals have to process information. Those who engage in more reflective thinking have been shown to be more discerning in their social media use. Mosleh and colleagues demonstrated that those who were higher in reflective thinking were more selective about whom they followed, shared news content from more reliable sources, and tweeted about more substantive topics.⁷² It is relevant to note that Mosleh’s study looked at reflective thinking as a trait, not a state that could be induced through friction. However, the finding is consistent with Fazio’s finding regarding induced deep processing.

Inoculation involves building preemptive resistance to disinformation by exposing people to examples of disinformation or disinformation techniques to help them recognize and reject them in the future.⁷³ This is similar to the concept of inoculation to viruses through the use of vaccines, where individuals are given a small dose, or antibodies for, a disease and that allows a body to prepare an informed defense against it.⁷⁴ Cook, Lewandowsky, and Ecker experimentally tested multiple preemptive interventions designed to reduce the effect of disinformation and found that inoculating messages that explain flawed arguments or highlight real scientific consensus on climate change affected climate attitudes.⁷⁵ The most substantial effect was on attitudes of perceived consensus in the scientific community; however, other attitudes (including increased trust in climate scientists and decreased trust in contrarian scientists) were also observed. Those that were inoculated were more likely to express perceived consensus among scientists, trust in climate scientists, and decreased trust in contrarian scientists than those who were not inoculated. The implication of these findings is that inoculation *prior to* presentation of disinformation can buffer the disinformation from affecting the individual’s perceptions.

The concept of “nudging,” introduced by Richard Thaler, involves providing small and subtle suggestions to encourage the type of behavior you want someone to perform or the decision

⁷¹ Lisa K. Fazio, “Pausing to Consider Why a Headline Is True or False Can Help Reduce the Sharing of False News,” *Harvard Kennedy School (HKS) Misinformation Review* 1, no. 2 (2020).

⁷² Mohsen Mosleh et al., “Cognitive Reflection Correlates with Behavior on Twitter,” *Nature Communications* 12, no. 921 (2021).

⁷³ “The Psychology of Misinformation: How to Prevent It.”

⁷⁴ Gustav J. Nossal, “Vaccination,” in *Encyclopedia of Life Sciences*, (Hoboken, NJ: Wiley, 1999), www.els.net.

⁷⁵ John Cook, Stephan Lewandowsky, and Ulrich K. H. Ecker, “Neutralizing Misinformation Through Inoculation: Exposing Misleading Argumentation Techniques Reduces Their Influence,” *PLoS ONE* 12, no. 5 (2017), doi: <https://doi.org/10.1371/journal.pone.0175799>.

you want them to make.⁷⁶ Single-serve snack packaging is an example of a nudge. The manufacturer is suggesting the quantity for one sitting, but the person is free to eat more than one single-serve snack in a sitting. In the realm of disinformation, a nudge would subtly prompt more analytical (controlled) thinking.⁷⁷ Pennycook and colleagues measured participants' truth discernment and willingness to share a headline with or without a nudge about accuracy in headlines.⁷⁸ Note that, in this case, they were not provided a lecture on the importance of accuracy or prevalence of inaccuracy on the internet. They were simply asked to make a judgment about the accuracy of a single headline (unrelated to the topic). They were then shown other headlines and asked how likely they were to share them on social media. The researchers found that participants were more likely to share true headlines (relative to false headlines) if they had been asked to rate the accuracy of a single headline.⁷⁹ Asking them to consider accuracy was a subtle nudge before they were asked if they would share a headline. This finding demonstrates that a minor, unobtrusive, and nonthreatening “nudge” can affect the likelihood of sharing false information on social media, even on a topic unrelated to the disinformation that might be shared.

Cognitive dissonance theory

Cognitive dissonance happens when a person is confronted with two competing thoughts. For example, a person might simultaneously think the following: Exercise is good for my body; when I exercise, it hurts. It is uncomfortable to hold two competing ideas/beliefs at one time. Therefore, people are motivated to reduce the conflict or remove the dissonance.

Dissonance theory describes how people are influenced to either accept or reject beliefs, as well as the information/arguments that accompany those beliefs. The theory includes both cognitive and emotional components. It posits that people feel uncomfortable when they have to reconcile conflicting information. Conflicting information is dissonant, whereas nonconflicting information is consonant, consistent, or compatible (we will use these as synonyms here).

⁷⁶ Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth and Happiness*, (New Haven, CT: Yale University Press, 2008).

⁷⁷ “The Psychology of Misinformation: How to Prevent It.”

⁷⁸ Gordon Pennycook et al., “Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy Nudge Intervention,” *Psychological Science* 31, no. 7 (2020), doi: <https://doi.org/10.1177/0956797620939054>.

⁷⁹ Ibid.

When information is incompatible with our beliefs, we react in one of four ways: (1) adding new, consonant cognitions, (2) removing the inconsistent information, (3) reducing the importance of opposing information, or (4) increasing the importance of compatible cognitions.⁸⁰ Festinger's classic example was of smokers encountering information indicating that smoking was bad for their health. In this case, the smoker has four options:

1. Change behavior or adopt new attitude (e.g., stop smoking) (adding new, consonant cognitions).
2. Continue to believe that smoking is not bad for health (remove the incompatible information).
3. Compare risk from smoking to risk from something worse, such as auto accidents (reducing the importance of opposing information).
4. Think about the enjoyment of smoking and its good effects (increase the importance of compatible information and that it might assist with weight control).

If someone saw the Columbian Chemicals story and believed it to be true, cognitive dissonance would be created when the reader was confronted with information that it was untrue. For example, if the person saw the fake CNN story about the explosion first, but then saw a news story showing "live footage" of the plant where all was seemingly normal, cognitive dissonance would occur in the person's mind. When confronted with that contradictory information (i.e., the story initially believed to be turned out to be false), people could (1) change their belief and agree they were misled, (2) refuse to believe the corrected information, (3) question or discredit the source trying to convince them it was a hoax (e.g., believing the correction is a "government cover-up," or (4) remind themselves of the importance of protecting ourselves against terrorist attacks and "err on the side of caution."

Cognitive dissonance affects behavior in several ways. Harmon-Jones (editor of *Cognitive Dissonance: Reexamining a Pivotal Theory in Psychology*) summarizes several paradigms of dissonance theory that have demonstrated aspects of human behavior that can be manipulated.⁸¹ These paradigms are first summarized in Table 4 and then described.

⁸⁰ Leon Festinger, *A Theory of Cognitive Dissonance*, (Evanston, IL: Row, Peterson, 1957).

⁸¹ Eddie Harmon-Jones, ed., *Cognitive Dissonance: Reexamining a Pivotal Theory in Psychology*, (Washington, DC: American Psychological Association, 2019), doi: <https://doi.org/10.1037/0000135-000>.

Table 4. Cognitive dissonance paradigms

| Paradigm | Description |
|-------------------------------|---|
| Norms consistency | People are more likely to believe information that is congruent with their previous norms, beliefs, and actions |
| Induced compliance | People will be more likely to change their beliefs if they act in a way consistent with the new belief for little or no direct benefit to them |
| Belief disconfirmation | People will more likely maintain their belief in disconfirmed information if they believe that a group they identify with still believes it, and they are provided arguments why it is still true |
| Effort justification | People are more likely to maintain their beliefs if they had to expend significant effort to commit to the belief in the first place; effort also includes incurring costs for committing to a belief |
| Free choice | People are more likely to maintain their beliefs if they made a willing choice to believe something in the first place; this is especially true if the choice was hard |

Source: CNA.

The norms-consistency paradigm. People are more likely to accept and defend information that is consistent with their previous norms and actions. For example, people whose actions are consistent with norms are harsher toward those who fail to conform to them. When sixth graders were given an opportunity to cheat, those who cheated revised their attitudes to believe that cheating was not so bad, whereas those who refused to cheat moved their attitudes to be harsher toward those who cheated. Real-world examples would be with how people dress. Those who follow a traditional dress code can harshly judge men who let their hair grow long, or women who dress in revealing outfits. The men and women who violate these norms believe that dress standards are not important. Adults who pay their taxes can be annoyed at those who do not. Those who do not pay taxes feel justified (e.g., “My loopholes are legal,” or “I am genuinely poor”).

The induced-compliance paradigm. When people are paid to say something that they do not believe, their belief will change to reflect what they have outwardly said. This is particularly true when the reward offered is small (say \$2) and less so if the reward is large (say \$20). That is because people can justify saying something against their beliefs for a large amount of money (they can tell themselves: I did not mean it; I did it for the money). However, to justify lying for a small amount of money, they revise their opinion to be more congruent with what they said.

The belief-disconfirmation paradigm. When a group commits to a belief that can be clearly shown to be false, the social circumstances of learning the truth makes a large difference in people’s opinions. In other words, for a group that believes the world will end on a specific date (a prediction that can be clearly and easily proved false when it didn’t happen), it matters how

they learn this information. Festinger provides the example of a group that believed that a flood would engulf the continent but that they (the believing group) had been chosen to be saved. When the anticipated flood did not materialize, the believers who were not with the group at the time changed their beliefs about whether or not the flood would ever occur. However, the believers who were physically with the group—those who were assembled together as the countdown occurred—decided that the flood had simply been temporarily postponed. Thus, the believers who were isolated from the group ultimately rejected the disinformation, and those who were colocated created an alternative form of the original disinformation.

The effort-justification paradigm. When people have had to undergo an unpleasant experience or have exerted a lot of effort to join a group, their commitment to that group will be higher than for groups that do not demand an unpleasant initiation or effort to join. An example is initiation processes at fraternities and sororities. Local chapters with difficult initiations elicit more commitment than do those with easier inductions. This principle is also evident in the observation that converts are often more committed to their religion than natives (i.e., someone who converts to Presbyterianism is likely to be more committed to the faith than someone who was born and raised as a Presbyterian).

The free-choice paradigm. Once someone has made a choice, dissonance is likely to be activated. There will be more dissonance if the choice is difficult, and less dissonance if the choice is easy (e.g., if the superiority of the thing chosen is more obvious). For example, someone choosing between chocolate and strawberry ice cream will experience more dissonance (because the choice is difficult) than someone choosing between chocolate and broccoli ice cream (where the superiority of the chocolate is more obvious). To reduce their experience of dissonance, people will tend to increase their opinion of the alternative they chose, and decrease their opinion of the choice they did not make. In our ice cream example, this means that once I choose chocolate ice cream over strawberry ice cream, my opinion of chocolate will increase and my opinion of strawberry will decrease.

Aronson explains these phenomena in terms of dissonance with a person's self-concept.⁸² This is known as the *self-consistency* explanation, and it posits that these phenomena occur when we need to reduce the dissonance between something we've done or said or thought, and the kind of person that we think we are. For example, our brains will work to reduce our discomfort if our decision to cheat or say bad things about someone is dissonant with our self-concept as a good person.

⁸² Elliot Aronson, "Dissonance Theory: Progress and Problems," in *Theories of Cognitive Consistency: A Sourcebook*, ed. Robert P. Abelson, et al. (Chicago: Rand-McNally, 1968); Elliot Aronson, "The Return of the Repressed: Cognitive Dissonance Theory Makes a Comeback," *Psychological Inquiry* 3, no. 4 (1992), doi: 10.1207/s15327965pli0304_1.

Harmon-Jones puts it this way: ⁸³

Once an individual commits to a given action, any information inconsistent with that commitment is likely to arouse dissonance and prevent the action from occurring. To maintain the commitment in the face of this inconsistent information, the individual selectively enhances the value of the chosen course of action and reduces the value of the unchosen course of action. Doing so makes effective execution of the chosen action more likely. (p. 17)

Implications of cognitive dissonance for disinformation

Dissonance theory has implications concerning whether someone will believe disinformation. First, in the free-choice paradigm, people will emphasize the good aspects of a choice that they freely chose. For example, if Russian disinformation persuades someone to freely join a faked rally (e.g., with paid actors), the person has to justify why they made that choice and went to the rally. People are likely to convince themselves that they attended the rally because it was for a good cause.

This has another implication. If disinformation can limit or shape the reader's response options to a preferred subset of options, then any response will solidify the reader's commitment to that information. Imagine, for example, that someone asks what you might do if you found a wallet on the ground. This could be presented as an open-ended question, but the options could also be shaped and limited by offering a binary choice: take the money for yourself or leave the wallet and walk away. In this case, you aren't presented with an option to return the wallet to the owner, but you are forced to make a choice, thus increasing your commitment to a choice that you may never have chosen on your own.

Disinformation can limit or shape choices by only providing one set of facts, all consistent with the goals of the person pushing the disinformation. For example, disinformation might say that fluoridation causes blindness because (1) my uncle Dave moved to an area with fluoridation and he went blind right after, (2) a survey of people showed that most people believe it does, or (3) the "Honesty in Dentistry Foundation" emphatically states that fluoridation is linked to blindness.⁸⁴

Second, in the belief-disconfirmation paradigm, people will become more resistant to unequivocal disconfirmation if they have others who will confirm the original erroneous belief. Let's say that someone believes that fluoridation causes blindness and sees news of a comprehensive study refuting that belief. They forward the news to a like-minded friend who

⁸³ Harmon-Jones, *Cognitive Dissonance: Reexamining a Pivotal Theory in Psychology*.

⁸⁴ In reality, this was an unscientific survey of the "Honesty in Dentistry Foundation" funded by an anti-fluoridation group.

says, “That flies in the face of everything we know. The study was probably biased for the big toothpaste companies.” That reassurance will be enough to discount the comprehensive study. In another example, if you belong to an online community that believes in UFOs, and you are presented with “proof” that a recent sighting did not occur, you are more likely to reject this proof if your online UFO friends also reject it. This is a reason why the comments sections of articles often contain messages from people who immediately contradict authoritative voices. Having one or more people echoing their disbelief in contrary information is often enough to maintain commitment to a belief that is disconfirmed by facts.

In the effort-justification paradigm, people are more committed to something if they have to work for it. We already discussed that in terms of hazing and initiation *processes* at sororities and fraternities. This also applies to the *content* of information. Providing disinformation that portrays the recipient as hard-working and moral, while others outside the group are lazy and dishonest, finds a willing audience (people like to feel virtuous) and might also make them recall that they had to work hard to be part of the special group of people receiving the disinformation.

In the induced-compliance paradigm, people are more committed to an idea if they perform an act for little or no compensation than for a large compensation. If a person can be nudged to spread disinformation for little or no financial gain, they will be more committed to the ideas they are disseminating than if they are paid significant amounts of money to do so.

In the final paradigm—norms consistency—people who comply with norms are less tolerant and harsher in their judgment of others who break norms. Those who break norms are more tolerant of others who also break them. Those who see and reject disinformation will judge “disinformation spreaders” harshly. Conversely, those who have spread or supported disinformation will reduce any previous commitments they might have made to fight disinformation in general.

What can we do about it?

When countering disinformation absorption through dissonance, preventive inoculation is a promising technique. An offshoot of dissonance theory, developed by McGuire, preventive inoculation applies the analogy of vaccination as a method of preventing people from being persuaded by propaganda.⁸⁵ As explained in the cognition section, inoculation theory states that forewarning people about propaganda will reduce its influence, just as vaccines prevent or lessen the effects of disease. In the case of dissonance, inoculation can remind individuals

⁸⁵ William McGuire, “Resistance to Persuasion Conferred by Active and Passive Prior Refutation of Same and Alternative Counterarguments,” *Journal of Abnormal Psychology* 63, no. 2 (1961).

about aspects of their self-concept and help them rehearse counterarguments to combat disinformation and remove the dissonance in a way that does not introduce disinformation.

However, when attempting to protect against disinformation or refute disinformation, it is important to try not to *create* cognitive dissonance for the receiver. To avoid creating dissonance when using inoculation, messages should be crafted in a way to appeal to previous beliefs:

1. Craft messages that are consistent with the receiver's self-concept.
2. Craft messages that are congruent with previous beliefs.
3. Frame responses in a way that is congruent with your goals and reminds recipients of their group identification.⁸⁶

As a theory of the emotion and cognitive aspects of social influence, dissonance theory posits ways that inoculation can be strengthened. The four key components of using dissonance to strengthen inoculation against disinformation are as follows:

1. *Threat*. By framing the incoming disinformation as a threat, the individual will be emotionally primed to identify the argument as propaganda, fight it, and solidify their previous beliefs. If people are not forewarned and given tools to combat the disinformation, they will be less likely to fight the message or think of counterarguments. Some research suggests that the forewarning of an attack is enough to induce resistance to a message, even without the other key element of refutational preemption (discussed below).⁸⁷ This first element is the motivational/emotional part of the effect.
2. *Refutational preemption*. This part allows the receiver to practice making arguments against the propaganda. By providing the receiver with arguments against the propaganda, the person's initial beliefs become stronger and prime them to think of additional reasons why their initial beliefs are correct. Sometimes people accept disinformation simply because they cannot bring to mind reasons for their prior beliefs.
3. *Involvement*. In order for inoculation to have an impact, it helps that the topic be one that the receiver finds important and salient.⁸⁸ A key part of this is pointing to the

⁸⁶ In the case of disinformation, the group should be defined as an in-group of people who are about to lose something very important. In combating disinformation, this would mean reminding people of the larger groups they are part of, including people with whom they might disagree. We will discuss this more fully in the next section, Group, Belief, Novelty Model (GBN).

⁸⁷ Josh Compton and Bobi Ivanov, "Untangling Threat During Inoculation-Conferred Resistance to Influence," *Communication Reports* 25 (2012), doi: 10.1080/08934215.2012.661018.

⁸⁸ Michael Pfau et al., "Enriching the Inoculation Construct: The Role of Critical Components in the Process of Resistance," *Human Communication Research* 24, no. 2 (1997), doi: 10.1111/j.1468-2958.1997.tb00413.x.

receiver’s interest (for example, “the real reason they are going to say this is to take away your right to unfluoridated waters”). Unless someone has a stake in the subject of the disinformation, he or she is less likely to fight it. As explained earlier, another way to obtain involvement is to get people to act on those initial beliefs.

4. *Delay/dosage.* McGuire’s original theory posited that, just as it takes time for the body to develop defenses against bacteria, it will take time for inoculation arguments to take hold in the receiver. This portion of the theory is controversial since evidence concerning the need for time has been mixed. However, the evidence does point to the fact that successful inoculation can have long-lasting effects. We think that delay might not be as important as repetition (or dosage) of the inoculation. In our view, the issue is not time, but amount of inoculation. Many inoculations require multiple doses. For example, with rabies vaccine, people should get their first dose as soon as possible (if not received early, the disease is almost always fatal), and on days 3, 7, and 14.⁸⁹

Dissonance has pointed to ways that the inoculation effect can be made stronger. One of those ways is to make the propaganda seem like an immediate attack on one’s vested interests. For example, “They will take away your house tomorrow if you do not respond immediately.” Additional ideas are to encourage anger at the attack source. For example, “The message will lie to you and assume you are stupid.” Third, inoculation will be stronger if it encourages the idea that the propaganda/disinformation is a threat to your core beliefs. For example, “If they have their way, they will spit on the ideals of our religion and democracy.” As noted earlier, dissonance theory predicts that inoculation will be stronger if it appeals to our self-concepts as smart, tough, and moral. For example, “Resisting this disinformation means you are clever, stalwart, and principled.”

When preventive inoculation is not possible, “therapeutic inoculation” might be valuable. The analogy here is to the rabies vaccine, which is given to people after they have been bitten by a potentially rabid animal. In the case of the rabies vaccine, a dose is given as close as possible to the attack and additional doses are given in succession. Compton and Pfau report that so far, examples of successful therapeutic inoculation are sparse⁹⁰ and, in our opinion, conflicting. Table 5 shows how these principles can be used for crafting messages.

Table 5. Key components of using dissonance for inoculation

| | Intervention |
|--------|--|
| Threat | Framing anticipated disinformation as a threat motivates the receiver to gather ways to reject it. |

⁸⁹ Nossal, "Vaccination."

⁹⁰ Josh Compton, “Prophylactic Versus Therapeutic Inoculation Treatments for Resistance to Influence,” *Communication Theory* 30, no. 3 (2020), doi: <https://doi.org/10.1093/ct/qtz004>.

| | Intervention |
|-------------------------|--|
| Refutational preemption | Providing counterarguments, even weak ones, will help the receiver recognize how to react how to reject the disinformation. |
| Involvement | Casting disinformation as an <i>immediate</i> threat activates mechanisms to speed reactions and bypass slower thinking processes, rejecting it more quickly |
| Delay and dosage | Getting the counterargument to the receiver quickly and repeatedly allows time for counterarguments to become familiar and rote before disinformation is believed. |

Source: CNA.

The Group, Belief, Novelty (GBN) model

Not all information is equally valuable to an individual. Some information (or disinformation) resonates with some people more than others. The Group, Belief, Novelty (GBN) model helps explain the likelihood that someone will pass information (rumors specifically) on to others.⁹¹ The theory posits that we accept and share information more readily when it comes from people we know, it appeals to what “our group” believes, and when we think it is new.

Because the theory is built primarily related to rumors (a subset of information), we briefly describe rumors and how they apply to this study on disinformation. Rumors are “unverified and instrumentally relevant information statements in circulation that arise in context of ambiguity, danger, or potential threat and that function to help people make sense of and manage risk.”⁹² While this can include conspiracy theories, they are outside the scope of this paper. For our purposes, all conspiracy theories are rumors but not all rumors are conspiracy theories. In this section, we focus solely on rumors that are not conspiracy theories. Researchers who developed the model cite empirical evidence that rumors are important to study because they alter purchase behaviors,⁹³ “spark” riots in conflict situations,⁹⁴ and

⁹¹ Bernard P. Brooks, Nicholas DiFonzo, and David S. Ross, “The GBN-Dialogue Model of Outgroup-Negative Rumor Transmission: Group Membership, Belief, and Novelty,” *Nonlinear Dynamics, Psychology, and Life Sciences* 17, no. 2 (2013).

⁹² Nicholas DiFonzo and Prashant Bordia, *Rumor Psychology: Social and Organizational Approaches*, (Washington D.C.: American Psychological Association, 2007). doi: <https://doi.org/10.1037/11503-000>.

⁹³ Nicholas DiFonzo and Prashant Bordia, “How Top Professionals Handle Hearsay: Corporate Rumors, Their Effects, and Strategies to Manage Them,” *Public Relations Review* 26 (2000); Alix Freedman, “Rumor Turns Fantasy Into Bad Dream,” *The Wall Street Journal*, May 10, 1991, 1991.

⁹⁴ Donald L. Horowitz, *The Deadly Ethnic Riot*, (Berkeley, CA: Univeristy of California Press, 2001).

influence stock market buying and selling.⁹⁵ In addition, they reference research findings that rumors affect attitudes.⁹⁶

The GBN model is a “two step agent-based mathematical model of negative rumor spread in the context of conflicting groups”⁹⁷ that uses concepts and findings from psychology and sociology research as the basis for its equations. According to the first step, the probability of rumor transmission between two people is based on the following three factors:

1. Group memberships (G) of the receiver and transmitter
2. Strength of their belief (B) in the rumor
3. The perceived novelty (N) of the rumor

The second step models how belief (B) levels and the perceived rumor novelty (N) of participants change over time, using findings from the literature on attitude change.

The first factor (G) encompasses several findings regarding how group membership affects the sharing of rumors. In the case of negative (derogatory) rumors, people share them with their in-group (people with whom they identify with, whether age, race, and gender, occupation, or political leaning). They rarely share rumors with the out-group (people with whom they feel little affiliation). It is not surprising that the rumor target (whether it attacks individuals in the in-group or the out-group) and valence (whether it praises or derogates) affect whether a rumor is shared and spreads.

Rumors that derogate the out-group are called “wedge-driving” rumors because they attempt to drive a wedge between groups. They can also be used for self-enhancing motives (boosting one’s self-esteem and those of the in-group)⁹⁸ or the desire to increase liking between the spreader and the hearer. Conversely, people rarely share derogatory rumors about the out-group with people from that out-group.

The second factor, belief, is the person’s degree of confidence that the information is true. In general, people are much more likely to share information they believe to be true. However, the threshold of degree of belief needed to share information is lower when the stakes are

⁹⁵ Nicholas DiFonzo and Prashant Bordia, “Rumor and Prediction: Making Sense (But Losing Dollars) in the Stock Market,” *Organizational Behavior and Human Decision Processes* 71 (1997).

⁹⁶ DiFonzo and Bordia, *Rumor Psychology: Social and Organizational Approaches*.

⁹⁷ Brooks, DiFonzo, and Ross, “The GBN-Dialogue Model of Outgroup-Negative Rumor Transmission: Group Membership, Belief, and Novelty.”

⁹⁸ Prashant Bordia and Nicholas DiFonzo, “Psychological Motivations in Rumor Spread,” in *Rumor Mills: The Social Impact of Rumor and Legend*, ed. Gary Allen Fine, Veronique Champion-Vincent, and Chip Heath (NY: Aldine Press, 2005).

higher and it involves planning future actions.⁹⁹ For example, college professors involved in tense labor negotiations shared information they believed to be true, but, because of the high stakes involved, they did not require a high degree of confidence in its veracity before sharing it. College students, in a very different context, were reluctant to share information about a recent campus attack that had occurred, for fear of upsetting others with false information.

The third factor, novelty, refers to whether the rumor is novel to the hearer, and whether the hearer thinks it might be new to others. The hearer may spread the rumor as a way to make sense of an uncertain and ambiguous situation. However, once the rumor becomes “old news,” people avoid sharing it because it has become stale. This accounts (in part) for why fact-checking and corrections are less shared than the original rumor—and why therapeutic inoculation is more difficult than is preventive inoculation (which we discussed earlier).

Buchanan explored multiple factors to determine which had the greatest effect on self-reported sharing of disinformation online.¹⁰⁰ He considered attributes of the message, including how authoritative the source was and how much consensus information was available (e.g., amount of “likes” on the platform). He also considered viewer characteristics, such as digital literacy, personality, and demographic variables. Across four studies, people who reported the greatest likelihood of sharing disinformation are those who thought the material was true or those who had preexisting attitudes consistent with the information.

Example of the GBN model using authoritarian attitudes

In this section, we use the construct of authoritarian attitudes to describe how the GBN model can illustrate disinformation sharing. People with authoritarian attitudes have been shown to have three predominant beliefs:¹⁰¹

- Conventional thinking
- Submission to authority
- A belief in aggression toward out-groups

They also tend to be considerably less open to experience and somewhat more conscientious than other people, whereas those on the other end of the scale (“anti-authoritarians”) are more

⁹⁹ Ralph L. Rosnow, James L. Esposito, and Leo Gibney, “Factors Influencing Rumor Spreading: Replication and Extension,” *Language and Communication* 8, no. 29-42 (1988); *ibid.*

¹⁰⁰ Tom Buchanan, “Why Do People Share False Information Online? The Effects of Message and Viewer Characteristics on Self-Reported Likelihood of Sharing Social Media Disinformation,” *PLoS ONE* 15, no. 10 (2020), doi: 10.5255/UKDA-SN-854297.

¹⁰¹ Bob Altmeyer, “The Other Authoritarian Personality,” *Advances in Experimental Social Psychology* 30 (1998), doi: [https://doi.org/10.1016/S0065-2601\(08\)60382-2](https://doi.org/10.1016/S0065-2601(08)60382-2).

open to experience and slightly less conscientious. More recent literature suggests that those with authoritarian attitudes might also be more disagreeable, dogmatic, and lacking in intellectual humility.¹⁰² However, this work is preliminary.

Note that authoritarian beliefs are, to a degree, situation dependent. Research has shown that people could relax these attitudes when they did not feel threatened, or when new experiences led them to question some of their previous beliefs.¹⁰³ Therefore, those with authoritarian attitudes are very likely to share disinformation framed as a threat to conventional beliefs (in the GBN model, it lowers the threshold of belief in the information's veracity before sharing it). Again, using the GBN model, they will share that disinformation with others with whom they identify. Maintaining a sense of being under threat is important for maintaining the cohesion of authoritarian groups.

In the Columbian Chemicals example described earlier, when the photo of the explosion was paired with the idea that Islamist terrorists had caused it, the explosion became threatening to those who are especially threatened by Islamist radicalism and believe that Islamist terrorist attacks in the US are rampant. This made the news appeal to previous beliefs (Beliefs). Sharing the novel photo was not only a civic duty (dissonance theory's explanation) but a way to boost one's self esteem and maintain the group's feelings of threat (Group). The time value of sharing the photo is also important: Sharing the photo immediately (Novelty) gets noticed. Trying to verify the validity of the photo slows you down and makes it less likely that you will be the first with the information.

Implications of the GBN Model for disinformation

The GBN model posits that group identity, degree of belief in the information, and information's novelty are powerful factors in whether information gets shared, with whom we share it, and when we share it. Once disinformation gets shared, it can be amplified and further distorted by those who find the information compelling (as in the old "telephone game" people sometimes play at group gatherings, the original story almost invariably gets distorted in the retelling). These factors coincide with characteristics of the current information environment, where people can stay connected with their smart phones 24/7, allowing people to feel that (1) they are never alone, (2) their voice will always be heard, and (3) they can put their attention

¹⁰² Thomas H. Castello et al., "Clarifying the Structure and Nature of Left-Wing Authoritarianism," *Journal of Personality and Social Psychology* (2021 pre-print), doi: doi 10.31234/osf.io/3nprq.

¹⁰³ Martin Roiser and Carla Willig, "The Strange Death of the Authoritarian Personality: 50 Years of Psychological and Political Debate," *History of the Human Sciences* 15, no. 4 (2002), doi: 10.1177/952695102015004682.

anywhere they want to put it.¹⁰⁴ In a sense, we live in a 24/7 virtual telephone game. Although these characteristics of the information environment can speed up rumor spread, the GBN model provides areas where the telephone game can be interrupted, slowed, or corrected.

What can we do about it?

In this section, we discuss how interventions can address all three aspects of the GBN model. Table 6 summarizes the insights, which are described in the subsequent paragraphs.

The first link in the rumor/disinformation chain is the identification with people who are similar to us, whether in terms of opinions, race, religion, gender, or age. The first implication of GBN is to start with people's narrow group identification.

Table 6. Using components of the GBN model to counter disinformation

| | Intervention |
|---------|--|
| Group | Expand the "group" with which an individual identifies by highlighting aspects of shared destiny and shared values |
| Belief | Increase a person's access to opposing opinions/information |
| Novelty | Create new content to expand the group and open up beliefs |

Source: CNA.

Shared destiny

One of the ways that the effects of disinformation can be reduced is through direct contact with the out-group in a situation with a shared destiny.¹⁰⁵ For example, a set of experiments showed that children in integrated classrooms became more accepting of minorities when they were placed into "jigsaw" groups where each child had to teach the others a piece of information that only they knew. This arrangement made it in the best interest of everyone in the group to listen to one another and root for each other's success with learning their part of the materials. We see this effect in the military, where everyone in the unit depends on particular individuals to do their job well; in fact, they entrust their lives to individuals of different races.

¹⁰⁴ Lauren C. Davis, "The Flight from Conversation," *The Atlantic*, Oct. 7, 2015, <http://www.theatlantic.com>; Sherry Turkle, *Life on the Screen: Identity in the Age of the Internet*, (New York: Touchstone, 1997); Sherry Turkle, *Alone Together: Why We Expect More from Technology and Less from Each Other*, (New York: Basic Books, 2011); Sherry Turkle, *Reclaiming Conversation: The Power of Talk in the Digital Age*, (New York: Penguin Books, 2015); Sherry Turkle, *The Empathy Diaries: A Memoir*, (New York: Penguin Publishing, 2021).

¹⁰⁵ Elliot Aronson and Diane Bridgeman, "Jigsaw Groups and the Desegregated Classroom: In Pursuit of Common Goals," *Personality and Social Psychology Bulletin* 5, no. 4 (1979), doi: <https://doi.org/10.1177/014616727900500405>.

To emphasize shared destiny in the internet age, it would be useful for login screens to feature information about the shared contributions different groups make to a common goal. Military commands depend on each other for resources, information, skillsets, and use of tools. Although login information can become part of the “background noise” quickly, a rotating set of reminders of interdependence and photos of individuals across the globe with whom one works for could reduce the likelihood of forming “gangs” of like-minded individuals unnecessarily sharing information only with an in-group. Knowing that information will be seen by others can have a moderating effect on what one says and what one shares.

It is important to frame arguments to be sure that people feel that they are not isolated in their original beliefs.¹⁰⁶ That is why advertisements often stress that many other people like their product (e.g., “99% of customers who tried our product liked it”). In the case of disinformation, “Everybody knows that this information is true.”

Shared values

Disinformation can be combated by pointing to areas where the out-group and in-group share common values.¹⁰⁷ For example, common values could include admiration of past leaders, respect for the rule of law, fairness, and justice.¹⁰⁸

Easier access to opposing opinions

The second area of GBN is the degree of belief. The GBN model posits that, when information is shared with another individual, that person’s feedback will modify one’s belief in the information. A good option to combat the tendency to share information only with like-minded individuals is to provide easier access to opposing opinions.

Two earlier methods of making it easier to access opposing opinions appeared soon after World War II. The fairness doctrine required holders of broadcast licenses to (1) present discussion of controversial issues of public importance and (2) do so in a way that is honest, equitable, and balanced. In the digital world, this might require access to opposing opinions and information on controversial issues in discussion format or question-and-answer format.

A second method was the equal-time rule. This required candidates of both parties to have access to a certain amount of broadcast time at the same rate as the opposing candidate. In this

¹⁰⁶ Wei-Kuo Lin and Michael Pfau, “Can Inoculation Work Against the Spiral of Silence? A Study of Public Opinion on the Future of Taiwan,” *International Journal of Public Opinion Research* 19, no. 2 (2007), doi: 10.1093/ijpor/edl030.

¹⁰⁷ Kate Woodsome, Danielle Kunitz, and Joy Sharon Yi, “Our Political Divide Is Dangerous. A Neuroscientist and Political Scientist Explain Why,” *Washington Post*, Dec. 24, 2020.

¹⁰⁸ These last suggestions of respect for the rule of law, fairness, and justice might not be accepted by those who are truly authoritarian.

way, broadcasts could not favor a preferred political candidate by denying the opposing candidate a forum to respond. If candidate A bought 30 seconds of TV time for ads, the opposing candidate would buy 30 seconds at the same rate. For the digital world, this might require allowing like-minded social media groups to allow people with different viewpoints some sort of access to the group.

Maintaining novelty of shared destiny and values

The third area of GBN is novelty. One of the common complaints about government information campaigns of the past (e.g., vaccination information, seat belt information) was that it was repetitive and predictable. If disinformation is to be counteracted, true information needs to be presented in different ways (Google updates its screen with a new factoid daily), and preferably by a diverse group of people, with humor and other ways to lighten the message. We see this in advertising campaigns all the time—GEICO constantly updates its ads, and some brands adopt a new slogan on a regular basis. The point is that old information gets more attention when it is repackaged as new.

Emotions and arousal

A person's state of emotion and arousal can also make them more susceptible to disinformation. It is no secret that some information and messages affect people more than others. In fact, marketing campaigns rely on the effect a message has on a person's emotions and subsequent behavior to buy products, donate to a cause, or vote for a candidate. Countless books, courses, and practices have been built on the concept of using messages to influence people's perceptions and behaviors, including the ubiquitous *How to Win Friends and Influence People* by Dale Carnegie and the extensively researched *Influence: Science and Practice* by Robert Cialdini.¹⁰⁹ While there are multiple compelling aspects of a message that can affect the message's ability to persuade, people are more likely to respond to (and share) information that is interesting, elicits positive emotions, or arouses action.¹¹⁰

Novelty, discussed at length in the GBN model section, is a primary determinant of what people consider interesting.¹¹¹ In addition, content can inspire emotional feelings (positive or

¹⁰⁹ Dale Carnegie, *How to Win Friends and Influence People*, (New York: Simon & Schuster, 1936); Robert B. Cialdini, *Influence: Science and Practice*, 4th ed. (Boston: Allyn and Bacon, 2001).

¹¹⁰ Katherine L. Milkman and Jonah Berger, "The Science of Sharing and the Sharing of Science," *Proceedings of the National Academies of Science* 111 no. Supplement 4 (2013), doi: 10.1073/pnas.1317511111; William J. Brady, Ana P. Gantman, and Jay J. Van Bavel, "Attentional Capture Helps Explain Why Moral and Emotional Content Go Viral," *Journal of Experimental Psychology: General* 149, no. 4 (2020).

¹¹¹ Brooks, DiFonzo, and Ross, "The GBN-Dialogue Model of Outgroup-Negative Rumor Transmission: Group Membership, Belief, and Novelty."; *ibid.*

negative) and can motivate or depress a response (high or low arousal). The following table provides examples that are commonly explored in the literature.

Table 7. Examples of positive and negative emotions that differ on level of arousal

| | High arousal | Low arousal |
|------------------|------------------|-------------|
| Positive emotion | Awe or amusement | Contentment |
| Negative emotion | Anxiety or fear | Sadness |

Source: CNA.

Emotion and arousal theories would explain that the sharing of the false Columbian Chemicals story was enhanced because the story elicited highly arousing negative emotions of anxiety and fear. Indeed, the cover graphic explicitly included the word “panic” on the page, further priming the arousing emotion.

Multiple researchers have explored characteristics that make information more likely to be shared. Commonly studied characteristics are message placement (i.e., prominence on a page), length, interest, usefulness of the information, emotions evoked from the story, and arousal evoked from the story. Research demonstrates that positive-emotional content is more viral than negative-emotional content, but there is evidence that high-arousal content (regardless of whether it is positive or negative) is more likely shared than low-arousal content.¹¹² Put simply, the level of arousal the content evokes—or the level of arousal a person feels when he or she sees the content—is even more important than positive or negative content when evaluating the likelihood of sharing information. That means that things that inspire awe, anger, and anxiety are more likely shared than things that inspire sadness. It is important to note that this relationship remains true even when a variety of other factors are held constant, including how interesting the material was, how practical (informative) it was, how prominently it was displayed, how long it was available, how long it was, the influence of the author, and the gender of the author. This research suggests that, when considering how information spreads, focusing on the content of a contagious message is valuable: how does it make people feel? Is that emotion arousing or depressing?

The psychological principles illuminated in this section thus far have described what is happening, but they have not offered explanations as to *why* the relationships exist. Our preliminary research has uncovered two possibilities: (a) an illustration of the Yerkes-Dodson

¹¹² Jonah Berger and Katherine L. Milkman, “What Makes Online Content Viral?,” *Journal of Marketing Research* 49, no. 2 (2012), doi: <https://doi.org/10.1509/jmr.10.0353>. Jonah Berger, “Arousal Increases Social Transmission of Information,” *Psychological Science* 22, no. 7 (2011), doi: <https://doi.org/10.1177/0956797611413294>.

Law that arousal is required to motivate behavior and (b) social psychological theories exploring the individual benefits of sharing information with a larger group.

The Yerkes-Dodson Law is the idea that there is a normal distribution (bell-curve) relationship between performance and arousal.¹¹³ At very low levels of arousal, individuals will not engage in tasks. At very high levels of arousal, individuals are so stimulated that they also cannot engage in tasks. There is an optimal level of arousal, near the midpoint, at which an individual will maximally perform. The research described regarding emotional arousal and sharing of information supports portions of the Yerkes-Dodson law.¹¹⁴ Low-emotional arousal is not as associated with information sharing as higher levels are. However, we are unaware of researchers specifically testing for the bell-curve relationship regarding information sharing (i.e., are there levels of emotional arousal where individuals will no longer be inclined to share information?).

There are multiple social benefits served by sharing information among group members. Sharing emotional information with others can increase social bonding or increase the perception of similarity among group members, both of which can strengthen a community.¹¹⁵ However, these social benefits also motivate other behaviors to protect the group and group identity. Brady, Crockett, and Van Bavel note that, when group identities are threatened from individuals inside or outside the group (through challenging information), those who more strongly associate with the group are more likely to express or share emotional information.

Implications of emotionally arousing content for disinformation

The relevance of emotions and arousal to the likelihood of information sharing is an important key for understanding the spread of disinformation. Research suggests that a reader is more likely to share false information if it arouses him or her. A recent study exploring 126,000 tweets found that the false content shared was more novel and inspired fear, disgust, or surprise while the true information shared inspired anticipation, sadness, joy, and trust.¹¹⁶ In this study, false content inspired emotions that were more arousing than emotions inspired by

¹¹³ Robert M. Yerkes and John D. Dodson, "The Relation of Strength of Stimulus to Rapidity of Habit-Formation," *Journal of Comparative Neurology and Psychology* 18 (1908).

¹¹⁴ Bruce E. Kaufman, "Emotional Arousal as a Source of Bounded Rationality," *Journal of Economic Behavior and Organization* 38, no. 135-144 (1999), doi: 10.1016/S0167-2681(99)00002-5.

¹¹⁵ William J. Brady, M. J. Crockett, and Jay J. Van Bavel, "The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online," *Perspectives on Psychological Science* 15, no. 4 (2020), doi: 10.1177/1745691620917336.

¹¹⁶ Vosoughi, Roy, and Aral, "The Spread of True and False News Online."

true content. In addition, they found that bots were as likely to share true information as false information and that those who shared false information were less likely to be influencers (e.g., fewer followers, less active on social media, less time on Twitter, and less likely to be verified) than those who shared true information. Of interest, this article explored the extent to which false and true information spread and found that an individual instance of true content rarely reached more than 1,000 people, whereas the false content routinely reached 1,000 to 100,000 people. The authors conclude that “the greater likelihood of people retweeting falsity more than truth is what drives the spread of false news, despite network and individual factors that favor the truth.”¹¹⁷

It is possible that multiple psychological principles, activated at the same time, can increase the likelihood that disinformation is absorbed or shared. Kaufman builds on the Yerkes-Dodson Law to describe how higher levels of emotional arousal can interfere with cognitive processes (e.g., short-term memory, organizing thoughts, and rational thinking).¹¹⁸ This suggests that if something is highly arousing, it might increase the likelihood that a person would process the information using automatic thinking.

What can we do about it?

Our literature review did not produce many insights into combating the effects of emotion and arousal on disinformation. Indeed, the articles that discuss the topic primarily describe the phenomenon in field and laboratory settings, but do not explore ways to counter the effect. However, the literature strongly indicates that emotional and arousing content are more likely to be shared, and Russia has tended to favor disinformation stories on topics known to be highly emotional and arousing to US citizens, such as race relations.

Given that disinformation is, by definition, constructed, it can be constructed to contain emotional or arousing messages. Several of the ideas to combat the absorption and spread of disinformation have been discussed in the previous sections. Inoculation, discussed to combat multiple psychological principles that can facilitate disinformation, could be applied to combating emotional and arousing messages. Making people aware of the effect of emotional/arousing information on beliefs and behaviors can give them tools to respond more accurately. Increasing controlled processing through friction or nudging could also help combat emotional and arousing disinformation. Providing cues to think more deeply about information on suspect websites might prime individuals to be more skeptical of emotional/arousing information they receive.

¹¹⁷ Vosoughi, Roy, and Aral, “The Spread of True and False News Online.”

¹¹⁸ Kaufman, “Emotional Arousal as a Source of Bounded Rationality.”

Summary

Understanding how false information spreads is the first step toward containing it. We described four major psychological principles that research demonstrates are related to the absorption and spread of disinformation.

The first psychological principle relates to how people **initially process information**. Our “processing capacity” is limited, so we cannot deeply attend to all new information. We take mental shortcuts to incorporating new information, and those shortcuts can open us up to mistakes. To the extent that we do not process information as deeply as we should, disinformation can be construed as true information.

The second psychological principle, called **cognitive dissonance**, describes the discomfort we feel when we are confronted with two competing ideas. We are motivated to reduce the dissonance by changing one attitude, removing (ignoring) the contradictory information, discounting the importance of contradictory information, or increasing the importance of compatible information. If disinformation supports our initial beliefs or creates less dissonance than true information, we are more likely to believe the disinformation.

The third psychological principle describes the role **group membership, beliefs, and novelty** play in absorbing and sharing disinformation. We are more likely to share information with people we identify with (i.e., consider members of our group) when we believe the information is true and when it is novel or urgent.

The fourth psychological principle describes the **role of emotion and arousal** in our sharing of disinformation. Research demonstrates that we pay more attention to information that makes us feel positively or that arouses us to act. That means we are more likely to share information if we feel awe, amusement, or anxiety than if we feel sadness or contentment.

Further we also summarized the research on countering the effect of disinformation through the psychological principles described. Several techniques discussed could be used to counter disinformation absorption that results from more than one psychological principle. Specifically, preventive inoculation and encouraging deeper, analytic thinking could shield a person from the effects of disinformation. Researchers recommend that disinformation containment policies should emphasize behavioral interventions that confront psychological mechanisms to dissuade the spread of disinformation rather than solely focusing only on bots and algorithms.¹¹⁹

¹¹⁹ Vosoughi, Roy, and Aral, “The Spread of True and False News Online.”

Conclusion

Brett Horvath, the president of Guardians.ai, argues that “disinformation is not the weaponization of knowledge”; rather, he says, “it’s the weaponization of cognition.” He further points out that “a coherent strategy...has to be built on principles: What are you defending, and what are you attacking?”¹²⁰ As such, we contend that the threat posed by disinformation cannot be met solely with technological and political solutions; the psychological principles that facilitate its absorption and spread must also be considered.

Human brains allow us to organize information and respond to it efficiently, but cognitive mechanisms are not infallible. As the research we have cited makes clear, disinformation weaponizes normal and necessary cognitive mechanisms by exploiting their vulnerabilities. But despite their susceptibility to disinformation, cognitive mechanisms present a potential path forward. To date, much has been written about how to respond technologically and politically to disinformation, but far less has been written about how we might respond psychologically to such manipulation. Identifying the best paths forward in this space requires first developing an understanding of the psychological mechanisms that make disinformation effective. As noted throughout this report, preventive inoculation and encouraging deeper, analytic thinking are frequently discussed in the literature as ways to protect against disinformation. Additional research is necessary to determine effective ways to integrate these two techniques into individual-level cognition and community-level intervention, especially in online environments. A more targeted approach may be possible if we explore the four psychological mechanisms being exploited to spread disinformation: initial information processing, cognitive dissonance, GBN, and emotional arousal.

Ultimately, solutions that consider technological and political factors alone cannot sufficiently diminish the large-scale absorption and spread of disinformation. Because disinformation primarily affects the mind, a multipronged response that includes evidence-based psychological interventions will be critical to the work of countering this growing threat.

¹²⁰ Robert Ackerman, “Cyber-Driven Disinformation Is Here to Stay,” Signal, Sept. 5, 2019, <https://www.afcea.org/content/cyber-driven-disinformation-here-stay>.

Glossary

| Psychological Terms | Definitions |
|------------------------------------|---|
| Alertness | A heightened awareness to the effects of disinformation that can prevent disinformation from being absorbed. |
| Analytic thinking | Deliberate, controlled, thoughtful evaluation of new information that can prevent disinformation from being absorbed. |
| Belief-disconfirmation paradigm | After a group commits to a belief that can clearly be proven false, the social circumstances of learning the truth have a large effect on group members' acceptance or denial of that truth. |
| Cognitive dissonance | Dissonance theory describes how people are influenced to either accept or reject beliefs, and the arguments/information that accompany those beliefs. It posits that people feel uncomfortable when they have to reconcile conflicting information. |
| Dual process theory | The foundation for understanding the two ways we process information. The first process, called automatic, requires very little effort and is fast, efficient, and "cognitively cheap." The second process, called controlled, is slower, more cognitively involved, and requires deliberate attention. |
| Effort-justification paradigm | When someone has to undergo an unpleasant experience or exert a lot of effort to join a group, their commitment to that group will be higher than for groups that do not demand an unpleasant initiation or increased effort to join. |
| Free-choice paradigm | People will emphasize the good aspects of a choice that they freely chose. Once someone has made a choice, dissonance is likely to be activated. There will be more dissonance if the choice is difficult and less dissonance if the choice is easy. |
| Friction | Friction occurs when something is difficult to process and can be created by asking questions that make others process information more deliberately (e.g., asking, "what makes you say that?"). It is the opposite of fluency and can prevent disinformation from being absorbed. |
| Group, Belief, Novelty (GBN) Model | The theory posits that we accept and share information more readily when it comes from a group of people we know (G), it appeals to what "our group" believes (B), and when we think it is new (N). |

| Psychological Terms | Definitions |
|--------------------------------|---|
| Heuristics | Mental shortcuts we use to make sense of the world. Heuristics are developed through previous experience and connections we make between similar pieces of information. |
| Illusory truth effect | When familiarity with something gives the illusion that it is more accurate, inadvertently increasing someone's belief in that information. |
| Induced-compliance paradigm | When someone is paid to say something that they do not believe, their belief will change to reflect what they have outwardly said. This is particularly true when the reward offered is small, and less so if the reward is large. |
| Initial information processing | Our brains employ mental shortcuts in an effort to process information as efficiently as possible, which can make us vulnerable to mistakes. |
| Inoculation | To provide awareness of possible disinformation and the effects of disinformation so others can recognize it when confronted by it later. This can prevent disinformation from being absorbed. |
| Information fluency | A person's ability to process information. This increases when information is clear and easy to understand. |
| Negative emotion | Emotions that make people feel bad. Examples include anxiety, fear and sadness. |
| Norms-consistency paradigm | People are more likely to accept and defend information that is consistent with their previous norms and actions. |
| Nudge | To provide subtle cues for an encouraged behavior. This can prevent disinformation from being absorbed. |
| Positive emotion | Emotions that make people feel good. Examples include, awe, amusement, and contentment. |
| Primacy and recency effect | The tendency for people to remember information that is presented first (primacy) or presented most recently (recency). |
| Skepticism | The awareness of possible hidden agendas and a personal desire to understand the evidence. This can prevent disinformation from being absorbed. |
| Therapeutic inoculation | Providing awareness of disinformation that a person has already experienced to attempt to remove that disinformation and mitigate its effect on future beliefs and behavior. |
| Yerkes-Dodson Law | The theory that there is a normal distribution (bell-curve) relationship between performance and arousal. At very low levels of arousal, an individual will not engage in tasks. At very high levels of arousal, the individual is so stimulated that they also cannot engage in tasks. |

Figures

| | | |
|-----------|---|----|
| Figure 1. | Twitter post showing the Columbian Chemicals plant explosion | 1 |
| Figure 2. | Twitter post asking Oregon senator Jeff Merkley about the Columbian Chemicals plant explosion | 2 |
| Figure 3. | Twitter account posting a fake CNN homepage showing the story of the Columbian Chemicals plant explosion | 3 |
| Figure 4. | Taxonomy of disinformation, misinformation, and mal-information | 9 |
| Figure 5. | Russian disinformation regarding US Ambassador Tefft..... | 12 |
| Figure 6. | US Embassy response to Russian disinformation regarding US Ambassador Tefft | 13 |

Tables

| | | |
|----------|---|----|
| Table 1. | Psychological principles relevant for disinformation adoption and spread..... | 25 |
| Table 2. | Dual process theory | 26 |
| Table 3. | Disinformation prevention mechanisms for initial information processing | 31 |
| Table 4. | Cognitive dissonance paradigms..... | 36 |
| Table 5. | Key components of using dissonance for inoculation | 41 |
| Table 6. | Using components of the GBN model to counter disinformation..... | 46 |
| Table 7. | Examples of positive and negative emotions that differ on level of arousal..... | 49 |

References

- Altmeyer, Bob. "The Other Authoritarian Personality." *Advances in Experimental Social Psychology* 30 (1998): 47-91. doi: [https://doi.org/10.1016/S0065-2601\(08\)60382-2](https://doi.org/10.1016/S0065-2601(08)60382-2).
- "The American Ambassador to Russia John F. Tefft Walked at an Opposition Rally in Maryino." Посол США в России Джон Ф. Тефт прогулялся на митинге оппозиции в Марьино. Twitter. Sept. 20, 2015. <https://twitter.com/rentvchannel/status/645658877426593792>.
- Aronson, Elliot. "Dissonance Theory: Progress and Problems." In *Theories of Cognitive Consistency: A Sourcebook*, edited by Robert P. Abelson, Elliot Aronson, William J. McGuire, Theodore M. Newcomb, Milton J. Rosenberg, and Percy H. Tannenbaum. Chicago: Rand-McNally, 1968.
- . "The Return of the Repressed: Cognitive Dissonance Theory Makes a Comeback." *Psychological Inquiry* 3, no. 4 (1992): 303-311. doi: 10.1207/s15327965pli0304_1.
- Aronson, Elliot, and Diane Bridgeman. "Jigsaw Groups and the Desegregated Classroom: In Pursuit of Common Goals." *Personality and Social Psychology Bulletin* 5, no. 4 (1979): 438-446. doi: <https://doi.org/10.1177/014616727900500405>.
- Asser, Martin. "What the Muhammad Cartoons Portray." BBC. Jan. 2, 2010. http://news.bbc.co.uk/2/hi/middle_east/4693292.stm.
- Bago, Bence, David G. Rand, and Gordon Pennycook. "Fake News, Fast and Slow: Deliberation Reduces Belief in False (But Not True) News Headlines." *Journal of Experimental Psychology: General* 149, no. 8 (Aug. 2020): 1608-1613. doi: 10.1037/xge0000729.
- Bail, Christopher, Brian Guay, Emily Maloney, Aidan Combs, D. Sunshine Hillygus, Friedolin Merhout, Deen Freelon, and Alexander Volfovsky. "Assessing the Russian Internet Research Agency's Impact on the Political Attitudes and Behaviors of American Twitter Users in Late 2017." *PNAS: Proceedings of the National Academy of Sciences of the United States of America* 117, no. 1 (2020): 243-250. Accessed Feb. 3, 2021. <https://www.pnas.org/content/117/1/243>.
- Beckett, Lois. "Misinformation 'Superspreaders:' Covid Vaccine Falsehoods Still Thriving on Facebook and Instagram." *The Guardian*. Jan. 6, 2021. Accessed Feb. 3, 2021. <https://www.theguardian.com/world/2021/jan/06/facebook-instagram-urged-fight-deluge-anti-covid-vaccine-falsehoods>.
- Begg, Ian Maynard, Ann Anas, and Suzanne Farinacci. "Dissociation of Processes in Belief: Source Recollection, Statement Familiarity, and the Illusion of Truth." *Journal of Experimental Psychology: General* 121, no. 4 (1992): 446-458. doi: <https://doi.org/10.1037/0096-3445.121.4.446>.
- Berger, Jonah. "Arousal Increases Social Transmission of Information." *Psychological Science* 22, no. 7 (2011): 891-893. doi: <https://doi.org/10.1177/0956797611413294>.
- Berger, Jonah, and Katherine L. Milkman. "What Makes Online Content Viral?". *Journal of Marketing Research* 49, no. 2 (2012): 192-205. doi: <https://doi.org/10.1509/jmr.10.0353>.
- Blake, Aaron. "A New Study Suggests Fake News Might Have Won Donald Trump the 2016 Election." *The Washington Post*. Apr. 3, 2018. Accessed Feb. 1, 2021. <https://www.washingtonpost.com/news/the-fix/wp/2018/04/03/a-new-study-suggests-fake-news-might-have-won-donald-trump-the-2016-election/>.
- Bordia, Prashant, and Nicholas DiFonzo. "Psychological Motivations in Rumor Spread." In *Rumor Mills: The Social Impact of Rumor and Legend*, edited by Gary Allen Fine, Veronique Campion-Vincent, and Chip Heath. NY: Aldine Press, 2005.
- Brady, William J., M. J. Crockett, and Jay J. Van Bavel. "The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online." *Perspectives on Psychological Science* 15, no. 4 (2020): 978-1010. doi: 10.1177/1745691620917336.

- Brady, William J., Ana P. Gantman, and Jay J. Van Bavel. "Attentional Capture Helps Explain Why Moral and Emotional Content Go Viral." *Journal of Experimental Psychology: General* 149, no. 4 (2020): 746-756.
- Brooking, Emerson, and Suzanne Kianpour. *Iranian Digital Influence Efforts: Guerrilla Broadcasting for the Twenty-First Century*. Atlantic Council. 2020. Accessed Feb. 1, 2021. <https://www.atlanticcouncil.org/wp-content/uploads/2020/02/IRAN-DIGITAL.pdf>.
- Brooks, Bernard P., Nicholas DiFonzo, and David S. Ross. "The GBN-Dialogue Model of Outgroup-Negative Rumor Transmission: Group Membership, Belief, and Novelty." *Nonlinear Dynamics, Psychology, and Life Sciences* 17, no. 2 (2013): 269-293.
- Buchanan, Tom. "Why Do People Share False Information Online? The Effects of Message and Viewer Characteristics on Self-Reported Likelihood of Sharing Social Media Disinformation." *PLoS ONE* 15, no. 10 (2020): 1-33. doi: 10.5255/UKDA-SN-854297.
- Carnegie, Dale. *How to Win Friends and Influence People*. New York: Simon & Schuster, 1936.
- Castello, Thomas H., Shauna M. Bowes, Sean T. Stevens, Irwin D. Waldman, Arber Tasimi, and Scott O. Lilienfeld. "Clarifying the Structure and Nature of Left-Wing Authoritarianism." *Journal of Personality and Social Psychology* (2021 pre-print). doi: doi 10.31234/osf.io/3nprq.
- Chaiken, Shelly, and Yaacov Trope. *Dual-Process Theories in Social Psychology*. New York: Guilford Press, 1999.
- Chen, Adrian. "The Agency." *New York Times Magazine*. June 2, 2015. Accessed Feb. 2, 2021. <https://www.nytimes.com/2015/06/07/magazine/the-agency.html>.
- Cialdini, Robert B. *Influence: Science and Practice*. 4th ed. Boston: Allyn and Bacon, 2001.
- "Columbian Chemicals Plant Explosion Hoax." Know Your Meme. 2014. Accessed Feb. 4, 2021. <https://knowyourmeme.com/memes/columbian-chemicals-plant-explosion-hoax>.
- Compton, Josh. "Prophylactic Versus Therapeutic Inoculation Treatments for Resistance to Influence." *Communication Theory* 30, no. 3 (2020): 330-343. doi: <https://doi.org/10.1093/ct/qtz004>.
- Compton, Josh, and Bobi Ivanov. "Untangling Threat During Inoculation-Conferred Resistance to Influence." *Communication Reports* 25 (2012): 1-13. doi: 10.1080/08934215.2012.661018.
- Cook, John, Stephan Lewandowsky, and Ulrich K. H. Ecker. "Neutralizing Misinformation Through Inoculation: Exposing Misleading Argumentation Techniques Reduces Their Influence." *PLoS ONE* 12, no. 5 (2017). doi: <https://doi.org/10.1371/journal.pone.0175799>.
- Cook, Sarah. "Welcome to the New Era of Chinese Government Disinformation." *The Diplomat*. May 11, 2020. Accessed Feb. 1, 2021. <https://thediplomat.com/2020/05/welcome-to-the-new-era-of-chinese-government-disinformation/>.
- Crawford, Krysten. "Stanford Study Examines Fake News and the 2016 Presidential Election." Stanford News. Jan. 18, 2017. Accessed Feb. 2, 2021. <https://news.stanford.edu/2017/01/18/stanford-study-examines-fake-news-2016-presidential-election/>.
- Davis, Lauren C. "The Flight from Conversation." *The Atlantic*. Oct. 7, 2015. <http://www.theatlantic.com>.
- DiFonzo, Nicholas, Jason W. Beckstead, Noah Stupak, and Kate Walders. "Validity Judgments of Rumors Heard Multiple Times: The Shape of the Truth Effect." *Social Influence* 11 (2016): 22-39. doi: <https://doi.org/10.1080/15534510.2015.1137224>.
- DiFonzo, Nicholas, and Prashant Bordia. "How Top Professionals Handle Hearsay: Corporate Rumors, Their Effects, and Strategies to Manage Them." *Public Relations Review* 26 (2000): 173-190.
- . "Rumor and Prediction: Making Sense (But Losing Dollars) in the Stock Market." *Organizational Behavior and Human Decision Processes* 71 (1997): 329-353.
- . *Rumor Psychology: Social and Organizational Approaches*. Washington D.C.: American Psychological Association, 2007. doi: <https://doi.org/10.1037/11503-000>.
- Ecker, Ulrich K. H., Stephan Lewandowsky, Briony Swire, and Darren Chang. "Correcting False Information in Memory: Manipulating the Strength of Misinformation Encoding and its Retraction." *Psychonomic Bulletin & Review* 18 (2011): 570-578. doi: 10.3758/s13423-011-0065-1.

- Ecker, Ulrich K. H., Stephan Lewandowsky, and David T. W. Tang. "Explicit Warnings Reduce But Do Not Eliminate the Continued Influence of Misinformation." *Memory and Cognition* 38 (2010): 1087-1100. doi: 10.3758/MC.38.8.1087.
- "Edgar Allan Poe—"The Balloon Hoax'." The Edgar Allan Poe Society of Baltimore. Last updated July 15, 2020. <https://www.eapoe.org/works/info/pt049.htm>.
- Fazio, Lisa K. "Pausing to Consider Why a Headline Is True or False Can Help Reduce the Sharing of False News." *Harvard Kennedy School (HKS) Misinformation Review* 1, no. 2 (2020).
- Festinger, Leon. *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson, 1957.
- Fiske, Susan T., and Shelly E. Taylor. *Social Cognition: From Brains to Culture*. 3rd ed. London: SAGE Publishing, 2016.
- Freedman, Alix M. "Rumor Turns Fantasy Into Bad Dream." *The Wall Street Journal*. May 10, 1991, 1991.
- Gareev, Makhmut A. "Anticipate Changes in the Nature of War: Every Era Has Its Own Kind of Military Conflict, and its Own Constraints, and its Own Special Biases." *Voyenno-Promyshlennyy Kuryer Online*. Jun. 5, 2013.
- Gerasimov, Valery V. "The Value of Science is Foresight." *Ценность науки в предвидении*. ВПК. ВПК. Feb. 26, 2013. Accessed Feb. 1, 2021. <https://vpk-news.ru/articles/14632>.
- . "Vectors of the Development of Military Strategy." *Векторы развития военной стратегии*. Krasnaya Zvezda. Красная звезда. Mar. 4, 2019. Accessed Feb. 1, 2021. <http://redstar.ru/vektory-razvitiya-voennoj-strategii/?attempt=1>.
- Harmon-Jones, Eddie, ed. *Cognitive Dissonance: Reexamining a Pivotal Theory in Psychology*. Washington, DC: American Psychological Association, 2019. doi: <https://doi.org/10.1037/0000135-000>.
- Horowitz, Donald L. *The Deadly Ethnic Riot*. Berkeley, CA: University of California Press, 2001.
- "IRGC (Islamic Revolutionary Guard Corps)." Counter Extremism Project. Accessed Feb. 3, 2021. <https://www.counterextremism.com/threat/irgc-islamic-revolutionary-guard-corps>.
- Jackson, Dean. "Issue Brief: How Disinformation Impacts Politics and Publics." National Endowment for Democracy. May 29, 2018. Accessed Feb. 2, 2021. <https://www.ned.org/issue-brief-how-disinformation-impacts-politics-and-publics/>.
- Kahneman, Daniel. *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux, 2011.
- Kahneman, Daniel, and Amos Tversky. "On the Psychology of Prediction." *Psychological Review* 80, no. 4 (1973): 237-251.
- Kaminska, Izabella. "A Lesson in Fake News from the Info-Wars of Ancient Rome." *Financial Times*. Jan. 17, 2017. Accessed Feb. 3, 2021. <https://www.ft.com/content/aaf2bb08-dca2-11e6-86ac-f253db7791c6>.
- Kaufman, Bruce E. "Emotional Arousal as a Source of Bounded Rationality." *Journal of Economic Behavior and Organization* 38, no. 135-144 (1999). doi: 10.1016/S0167-2681(99)00002-5.
- Kliman, Daniel, Andrea Kendall-Taylor, Kristine Lee, Joshua Fitt, and Carisa Nietzsche. *Dangerous Synergies*. CNAS. 2020. Accessed Feb. 1, 2021. <https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/CNAS-Report-Dangerous-Synergies-May-2020-DoS-Proof.pdf?mtime=20200506164642&focal=none>.
- Kurlantzick, Joshua. "How China Ramped Up Disinformation Efforts During the Pandemic." Council on Foreign Relations. Sept. 10, 2020. Accessed Feb. 2, 2021. <https://www.cfr.org/in-brief/how-china-ramped-disinformation-efforts-during-pandemic>.
- Lee, Doowan. "The United States Isn't Doomed to Lose the Information Wars." *Foreign Policy*. Oct. 16, 2020. Accessed Feb. 3, 2021. <https://foreignpolicy.com/2020/10/16/us-election-interference-disinformation-china-russia-information-warfare/>.
- Li, Cong. "Primacy or Recency Effect? A Long-Term Memory Test of Super Bowl Commercials." *Journal of Consumer Behavior* 9, no. 1 (2009): 32-44.

- Lin, Wei-Kuo, and Michael Pfau. "Can Inoculation Work Against the Spiral of Silence? A Study of Public Opinion on the Future of Taiwan." *International Journal of Public Opinion Research* 19, no. 2 (2007): 155-172. doi: 10.1093/ijpor/edl030.
- Mayo, Ruth. "Cognition Is a Matter of Trust: Distrust Tunes Cognitive Processes." *European Review of Social Psychology* 26 (2015): 283-327. doi: <https://doi.org/10.1080/10463283.2015.1117249>.
- McBride, Megan, Zack Gold, and Kasey Stricklin. *Social Media Bots: Implications for Special Operations Forces*. CNA. DRM-2020-U-028199-Final. Sept. 2020. Accessed Feb. 2, 2021. https://www.cna.org/CNA_files/PDF/DRM-2020-U-028199-Final.pdf.
- McGuire, William. "Resistance to Persuasion Conferred by Active and Passive Prior Refutation of Same and Alternative Counterarguments." *Journal of Abnormal Psychology* 63, no. 2 (1961): 326-332.
- Mikkelson, David. "Did Keanu Reeves Say Hollywood Elites Use the 'Blood of Babies' to Get High?". Snopes. Nov. 22, 2017. Accessed Feb. 1, 2021. <https://www.snopes.com/fact-check/keanu-reeves-blood-drinking/>.
- Milkman, Katherine L., and Jonah Berger. "The Science of Sharing and the Sharing of Science." *Proceedings of the National Academies of Science* 111 no. Supplement 4 (2013): 13642-13649. doi: 10.1073/pnas.1317511111.
- Mosleh, Mohsen, Gordon Pennycook, Antonio A. Arechar, and David G. Rand. "Cognitive Reflection Correlates with Behavior on Twitter." *Nature Communications* 12, no. 921 (2021).
- Murdock, Bennet B., Jr. "The Serial Position Effect on Free Recall." *Journal of Experimental Psychology* 64, no. 5 (1962): 482-488. doi: <https://doi.org/10.1037/h0045106>.
- Murphy, Jamie, Charles Hofacker, and Richard Mizerski. "Primacy and Recency Effects on Clicking Behavior." *Journal of Computer-Mediated Communication* (2020): 522-535.
- Nemr, Christina, and William Gangware. *Weapons of Mass Distraction*. Park Advisors. 2019. Accessed Feb. 2, 2021. <https://www.state.gov/wp-content/uploads/2019/05/Weapons-of-Mass-Distraction-Foreign-State-Sponsored-Disinformation-in-the-Digital-Age.pdf>.
- Nossal, Gustav J. "Vaccination." In *Encyclopedia of Life Sciences*. Hoboken, NJ: Wiley, 1999. www.els.net.
- Pellegrino, Bianca. "People Famous For Doing Things They Didn't Do." ListVerse. Sept. 8, 2019. <https://listverse.com/2019/09/08/top-10-people-famous-for-doing-things-they-didnt-do/>.
- Pennycook, Gordon, and David G. Rand. "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning than by Motivated Reasoning." *Cognition* 188 (July 2019): 39-50. doi: <https://doi.org/10.1016/j.cognition.2018.06.011>.
- Pennycook, Gordon, Jonathan McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. "Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy Nudge Intervention." *Psychological Science* 31, no. 7 (2020). doi: <https://doi.org/10.1177/0956797620939054>.
- Pfau, Michael, Kyle J. Tusing, Ascan F. Komer, Waipeng W. Lee, Linda C. Godbold, Linda J. Penaloza, Violet S. Yang, and Hong, Yah-Huei "Enriching the Inoculation Construct: The Role of Critical Components in the Process of Resistance." *Human Communication Research* 24, no. 2 (1997): 187-215. doi: 10.1111/j.1468-2958.1997.tb00413.x.
- Roisner, Martin, and Carla Willig. "The Strange Death of the Authoritarian Personality: 50 Years of Psychological and Political Debate." *History of the Human Sciences* 15, no. 4 (2002): 71-96. doi: 10.1177/952695102015004682.
- Rosnow, Ralph L., James L. Esposito, and Leo Gibney. "Factors Influencing Rumor Spreading: Replication and Extension." *Language and Communication* 8, no. 29-42 (1988).
- Saaty, T. L., and Müjgan Sagir Özdemir. "Why the Magic Number Seven Plus or Minus Two." *Mathematical and Computer Modelling* 38 (2003): 233-244. doi: 10.1016/S0895-7177(03)90083-5.

- Schmaltz, Rodney, and Scott Lilienfeld. "Hauntings, Homeopathy, and the Hopkinsville Goblins: Using Pseudoscience to Teach Scientific Thinking." *Frontiers in Psychology* 5 (2014): 1-5. doi: <https://doi.org/10.3389/fpsyg.2014.00336>.
- Schreck, Carl. "Photoshop Wars: U.S. Ambassador 'Attends' Russian Opposition Rally...and the Moon Landing." Radio Free Europe/Radio Liberty. Sept. 21, 2015. Accessed Feb. 3, 2021. <https://www.rferl.org/a/russia-photoshop-us-ambassador-tefft-opposition-rally-ren-tv/27260885.html>.
- Shane, Tommy. "The Psychology of Misinformation: How to Prevent It." First Draft News. July 27, 2020. <https://firstdraftnews.org/latest/the-psychology-of-misinformation-how-to-prevent-it/>.
- . "The Psychology of Misinformation: Why We're So Vulnerable." First Draft News. June 30, 2020. <https://firstdraftnews.org/latest/the-psychology-of-misinformation-why-were-vulnerable/>.
- Stricklin, Kasey. "Why Does Russia Use Disinformation?". Lawfare. Mar. 29, 2020. Accessed Feb. 2, 2021. <https://www.lawfareblog.com/why-does-russia-use-disinformation>.
- Swire, Briony, and Ulrich K. H. Ecker. "Misinformation and Its Correction: Cognitive Mechanisms and Recommendations for Mass Communications." In *Misinformation and Mass Audiences*, edited by Brian Southwell, Emily A. Thorson, and Laura Sheble. Austin: University of Texas Press, 2018.
- Swire, Briony, Ulrich K. H. Ecker, and Stephan Lewandowsky. "The Role of Familiarity in Correcting Inaccurate Information." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43, no. 12 (2017): 1948-1961. doi: 10.1037/xlm0000422.
- Thaler, Richard H., and Cass R. Sunstein. *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven, CT: Yale University Press, 2008.
- Turkle, Sherry. *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books, 2011.
- . *The Empathy Diaries: A Memoir*. New York: Penguin Publishing, 2021.
- . *Life on the Screen: Identity in the Age of the Internet*. New York: Touchstone, 1997.
- . *Reclaiming Conversation: The Power of Talk in the Digital Age*. New York: Penguin Books, 2015.
- Unkelbach, Christian. "Reversing the Truth Effect: Learning the Interpretation of Processing Fluency in Judgments of Truth." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33, no. 1 (2007): 219-230. doi: 10.1037/0278-7393.33.1.219.
- "US Ambassador to Russia John F. Tefft Is Sent to Opposition Rally in Maryino." Посла США в России Джона Ф. Теффта отправили на митинг оппозиции в Марьино. Ren.TV. Sept. 20, 2015. <http://ren.tv/novosti/2015-09-20/posla-ssha-v-rossii-dzhona-f-teftta-otpravili-na-miting-oppozicii-v-marine>.
- Vosoughi, Soroush, Deb Roy, and Sinan Aral. "The Spread of True and False News Online." *Science* 359, no. 6380 (2018): 1146-1151. <https://science.sciencemag.org/content/359/6380/1146>.
- Woodsome, Kate, Danielle Kunitz, and Joy Sharon Yi. "Our Political Divide Is Dangerous. A Neuroscientist and Political Scientist Explain Why." *Washington Post*. Dec. 24, 2020.
- Yerkes, Robert M., and John D. Dodson. "The Relation of Strength of Stimulus to Rapidity of Habit-Formation." *Journal of Comparative Neurology and Psychology* 18 (1908): 459-482.
- Zakem, Vera, Megan McBride, and Kate Hammerberg. *Exploring the Utility of Memes for U.S. Government Influence Campaigns*. CNA. DRM-2018-U-017433-Final. 2018. https://www.cna.org/cna_files/pdf/DRM-2018-U-017433-Final.pdf.

This report was written by CNA's Strategy, Policy, Plans, and Programs Division (SP3).

SP3 provides strategic and political-military analysis informed by regional expertise to support operational and policy-level decision-makers across the Department of the Navy, the Office of the Secretary of Defense, the unified combatant commands, the intelligence community, and domestic agencies. The division leverages social science research methods, field research, regional expertise, primary language skills, Track 1.5 partnerships, and policy and operational experience to support senior decision-makers.

CNA is a not-for-profit research organization that serves the public interest by providing in-depth analysis and result-oriented solutions to help government leaders choose the best course of action in setting policy and managing operations.



Dedicated to the Safety and Security of the Nation

DRM-2021-U-029337-1Rev-2

3003 Washington Boulevard, Arlington, VA 22201

www.cna.org • 703-824-2000