Toward equitable science communication in an algorithmically infused society: Understanding media and message antecedents to knowledge of wicked science issues

By

Shiyu Yang

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Mass Communications)

at the

UNIVERSITY OF WISCONSIN-MADISON

2023

Date of final oral examination: 06/29/2023

The dissertation is approved by the following members of the Final Oral Committee:
Dominique Brossard, Professor and Chair, Life Sciences Communication
Kaiping Chen, Assistant Professor, Life Sciences Communication
Nan Li, Assistant Professor, Life Sciences Communication
James E. Pustejovsky, Associate Professor, Educational Psychology
Dietram A. Scheufele, Professor, Life Sciences Communication

Table of Contents

Disclaimer

This dissertation is supported by the Morgridge Institute for Research and the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these organizations.

Acknowledgments

This dissertation would not have been possible without the help of many people. First, I am deeply grateful for my advisor, Dr. Dominique Brossard. Dominique always provided sound advice and support throughout my years in the Department of Life Sciences Communication (LSC) at the University of Wisconsin-Madison (UW-Madison). Because of Dominique's support, I was able to complete a second M.S. degree in Computer Sciences (CS) while pursuing my Ph.D. in LSC. The CS education provided me with the foundation to conduct the computational social science research involved in this dissertation. From Dominique, I learned so much about what it takes to make a good scholar, teacher, and mentor: attention to both details and big-picture problems, time management skills, interactive teaching strategies, care, responsibility, and confidence in the work one does, just to name a few.

I also wish to thank my committee members, Dr. Dietram Scheufele, Dr. Nan Li, Dr. Kaiping Chen, and Dr. James Pustejovsky. Working with Dietram on different research projects during my time in LSC was an invaluable experience. Dietram always shared incisive and constructive feedback on projects, including this dissertation. I am also very thankful to Nan for her critical feedback on this dissertation and other research projects, her advice on Ph.D. life and the job market, and her encouragement. I learned a great deal from Nan about studying and teaching visual communication of science. My interest in computational social science was inspired by Kaiping when she gave her job talk at UW-Madison almost five years ago. I appreciate her advice on research, teaching, and the job market. I am grateful to James for taking the time to meet with me on multiple occasions and providing help with the analytical methods used in this dissertation.

My gratitude extends to my friends and colleagues in LSC and at UW-Madison who made my time here more joyful (names listed in no particular order): Patrice Kohl, Kate Rose, Emily Howell, Chris Wirz, Richard Heinrich, Nicky Krause, Becca Beets, Mikhaila Calice, Hannah Monroe, Luye Bao, Anqi Shao, Tera Holtz, Don Stanley, Dexter Patterson, Dr. Bret Shaw, Dr. Todd Newman, Dr. Michael Xenos, Dr. Sedona Chinn, Tyler Caraza-Harter, Dr. Justin Hsu, and Dr. Young Wu. I would also like to acknowledge the financial support and the various opportunities for teaching, research, and professional development made possible by LSC and the College of Agricultural and Life Sciences (CALS) at UW-Madison.

Last but not least, I dedicate this dissertation to my family, whose unwavering love and support made everything possible.

Abstract

In an era of increasingly fast-moving and disruptive scientific and technological advances, communicating science to broad segments of the population becomes more important than ever. The persistent challenge to reach diverse population segments, especially those that are traditionally underserved by science outreach efforts, is complicated by the fast-evolving information environment we live in. Many of today's media platforms are driven by artificial intelligence (AI) algorithms, which tailor information to our personal preferences, biases, and contexts. Despite the deepening integration of these algorithmically driven media tools—such as social media platforms and online news aggregators—into our everyday life, it remains unclear how these media tools distribute scientific information in society across diverse population segments. It is also unclear how such media infrastructures might shape important individual and collective outcomes, such as knowledge of emerging science and technologies. Moreover, research and practice should continue to develop communication strategies for addressing information inequity in contemporary media environments.

This dissertation investigates how the media and their message-level factors might affect diverse social segments' knowledge of three wicked science issues, namely human gene editing, artificial intelligence, and COVID-19 and its related vaccines, in today's algorithmically infused information environment. It begins by reviewing the current algorithmically infused information environment, explain disparities in science knowledge and how the new information environment shapes them, and comparing the three wicked science issues that make up the contexts of inquiry. This dissertation then uses a sock puppet algorithm audit design to explore the extent to which algorithmically driven social media platforms such as YouTube recommend science content based on users' racial and socioeconomic status (SES) profiles (Study A). From

there, this dissertation uses three public opinion survey datasets to examine whether use of different social media platforms affects the gaps in factual and perceived knowledge of wicked science issues among Americans with different racial and SES makeup (Study B). Finally, using an experiment, this dissertation examines how message characteristics such as information modality (visual versus text-based) and rhetorical mode (narrative versus logical-scientific) could be leveraged against inequalities in understanding of and engagement with health science information among individuals with varying levels of science literacy (Study C).

Results show that social media algorithms, such as the YouTube algorithm, can indeed expose sociodemographically diverse audiences to different subsets of information even when people are actively searching for the same science issues, although the degree of information tailoring may depend on the specific search topic (e.g., how heavily the issue is discussed by different sources on YouTube). Higher-SES audiences, especially higher-SES White American audiences, are likely to receive a wider range of video and channel recommendations when searching for science issues than lower-SES audiences. In addition, use of different social media platforms is overall associated with wider gaps in factual knowledge of the three science issues and narrower gaps in perceived knowledge of the three wicked science issues among racial minorities than among Whites. Finally, exposure to visual or narrative messages about COVID-19 vaccine safety significantly reduces the gap in factual knowledge of COVID-19 vaccine safety between high and low science literacy groups, while exposure to narrative (but not visual) messages also reduces the gap in message elaboration between the groups. However, combining visuals and narratives does not further enhance message effectiveness. The theoretical, methodological, and practical implications of this dissertation are discussed.

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
AI	Artificial intelligence
AGI	Artificial general intelligence
API	Application programming interface
CCM	Center for Community Media at the City University of New York
CDC	Centers for Disease Control and Prevention
CHTC	Center for High Throughput Computing at the University of Wisconsin-Madison
COVID-19	Coronavirus disease 2019
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPR-Cas9	Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9
DNA	Deoxyribonucleic acid
HGE	Human gene editing
NASEM	National Academies of Sciences, Engineering, and Medicine
NSF	National Science Foundation
PSA	Public service announcement
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SES	Socioeconomic status
STEM	Science, technology, engineering, and mathematics
STM	Structural topic modeling
U.S.	United States
VM	Virtual machine

Chapter 1

Introduction and overview

In an era of increasingly fast-moving and disruptive scientific and technological advances, communicating science to broad segments of the populations becomes more important than ever. On the one hand, a scientifically and technologically skilled population is more likely to achieve economic success and competitiveness and respond to challenges that emerge in society, many of which can be understood or addressed at least in part through science (National Academies of Sciences, Engineering, and Medicine [NASEM], 2016b). On the other hand and more importantly, it is a democratic imperative to actively engage broad publics who may hold different perspectives and values in civic decision-making around issues of social importance, including those related to science and technology (NASEM, 2016b, 2017a). However, science communication and outreach efforts have traditionally failed to reach certain population segments such as racial minorities and those who are already educationally, financially, and informationally disadvantaged (Christopherson et al., 2021). In the United States (U.S.), men and individuals who have college degrees are more likely to actively consume science news (Funk et al., 2017a); museum visitors are primarily White (Gold, 2021) and highly educated (Wilkening Consulting, 2022); and even popular science content aimed for a broad audience rarely reaches beyond groups that are already highly educated and knowledgeable about and receptive to science (Akin et al., 2017). Nonetheless, societies need to meaningfully engage all stakeholder groups to address the challenges of emerging and controversial science and technologies and make meaningful decisions moving forward (for an overview, see Scheufele, 2022). Consequently, in the U.S., both public and private funding agencies are calling for support for advancing equity in science communication research and practice and engaging traditionally underrepresented and underserved communities with the scientific enterprise (see, e.g., (Christopherson et al., 2021; National Science Foundation, 2022).

The persistent challenge to reach diverse population segments, especially those that are traditionally underserved by science outreach efforts, is complicated by the fast-evolving information environment we live in, and research on the societal impacts of new information technologies tends to lag behind the technological advancements themselves. Many of today's media platforms are driven by artificial intelligence (AI) algorithms, which tailor information to our personal preferences, biases, and contexts so that we keep coming back to these platforms (Brossard & Scheufele, 2022). Despite the deepening integration of these algorithmically driven media tools—such as social media platforms and online news aggregators—into our everyday life, it remains unclear how these media tools distribute scientific information in society across diverse population segments. It is also unclear how such media infrastructures might shape important individual and collective outcomes, such as knowledge of emerging science and technologies. Moreover, research and practice should continue to develop communication strategies for addressing information inequity in contemporary media environments.

To that end, this dissertation explores how some media and their message-level factors might affect diverse social segments' knowledge of wicked science issues—issues that have no definitive formulation or solution (Rittel & Webber, 1973)—in today's algorithmically infused information environment. As a starting point and using a sock puppet algorithm audit design, I explore to what extent algorithmically driven social media platforms such as YouTube recommend science content based on users' racial and socioeconomic status (SES) profiles. From there, my dissertation uses three public opinion survey datasets to examine whether use of different social media platforms affects the gaps in knowledge of complex science issues among

Americans with different racial and SES makeup. Finally, using an experiment, my dissertation examines how message characteristics such as information modality and rhetorical mode could be leveraged against inequalities in scientific understanding. Before taking a closer look at the contents of this dissertation, I summarize extant research in order to: (a) give an overview of the current algorithmically infused information environment, (b) identify the gaps in research explaining inequities in science knowledge and how the new information environment shapes them, and (c) review and compare three wicked science issues—human gene editing (HGE), AI, and COVID-19 vaccines, which make the contexts of this study.

The algorithmically infused information environment

The way in which people receive, share, and interpret scientific information has fundamentally changed with advances in our information environments (Brossard & Scheufele, 2022). Online media platforms are quickly outdating traditional legacy media in disseminating scientific information as today people are regularly getting news from online sources, including social media. In 2022, 67% of Americans got their news online (42% from social media) while the share of Americans who got their news from TV and print media was 48% and 15%, respectively (Newman et al., 2022).

Moreover, online media platforms are largely driven by AI algorithms that are beginning to outpower scientists, journalists, and government agencies in gatekeeping scientific information. For instance, AI is being used to augment many journalistic tasks ranging from churning through massive datasets to chase down facts and spot patterns for story ideas, to writing news snippets across varieties of topic domains, to serving up algorithmic newsfeed recommendations, and to moderating the comment sections on news websites (e.g., Broussard et al., 2019; Kobie, 2018; WashPostPR, 2017). AI also largely drives what we see on social media platforms. Designed to outperform human capacities to sift through an overabundance of information and to capitalize on human cognitive and emotional weaknesses (Scheufele et al., 2021), those profit-driven algorithmic tools narrowcast information toward audience members based on their demographics, preferences, and a wealth of user histories and digital trace data, ultimately determining who gets to see what content online (Brossard & Scheufele, 2022).

These changes in our information environments are forcing researchers to rethink traditional models of media effects (Bennett & Iyengar, 2008; Cacciatore et al., 2016). The magic bullet or hypodermic needle models of strong media effects dominated in the 1930s, which assumed, without significant empirical testing, powerful, direct, and uniform persuasive effects of mass media (McQuail, 2005). After World War II, using public opinion data collected from large-scale panel surveys during U.S. elections, researchers identified limited persuasive effects of mass media, specifically due to a two-step information flow whereby information did not directly trigger opinion changes among members of the public but instead was first passed down to them by opinion leaders and due to partisans' tendency to select information that already conformed to their preexisting attitudes rather than information that would sway their attitudes in new directions (Katz & Lazarsfeld, 1955; Lazarsfeld et al., 1948). In the 1970s, the field returned to the concept of strong media effects with television sets becoming a mainstay in homes and other institutions. Two popular media effects theories, the spiral of silence (Noelle-Neumann, 1974) and cultivation (Gerbner & Gross, 1978), posited that ubiquitous and consonant mass-mediated messages had powerful influences on people's perceptions of what others thought (Noelle-Neumann, 1974) and what the world looked like (Gerbner & Gross, 1978). Since the 1970s and particularly in the 1980s and 1990s, research on agenda setting, priming, and framing effects suggested that the persuasive effects of mass-mediated information could be potentially powerful but depended on audience members' individual characteristics such as their cognitive schemas and value predispositions (McQuail, 2005). Finally, moving to an era of social media, Cacciatore et al. (2016) proposed new preference-based media effects models combining elements of both strong (due to tailored persuasion) and weak (due to preference-based reinforcement) media effects. How well such preference-based models hold in an algorithmically infused information environment requires continued empirical tests, and this dissertation represents a step toward answering this question.

The paradigm shift to algorithmically driven information curation also has implications for the civic science society (Christopherson et al., 2018). Members of the public who engage less with credible science content may be deprived of future opportunities to engage with such content due to preference-based reinforcement in the online information environments and therefore gradually become further disconnected from science over time. In the long run, individuals may become entrenched in their own information filter bubbles (Pariser, 2011), which may lead to the siloing of their science-related views and attitudes. Individuals may also become particularly susceptible to online misinformation in algorithmically curated media environments in part because the AI algorithms capitalize on human cognitive and emotional weaknesses (Scheufele et al., 2021). In addition, how other people interact with online content—through clicking, sharing, commenting, and so forth—not only informs algorithmic content curation (e.g., YouTube Creators Academy, 2017; Hoiles et al., 2017), but also provides cues about the relevance and popularity of the content, which ultimately affect how we interpret the mediated content and the science issues being presented (Anderson et al., 2014; Spartz et al., 2017). For instance, if a climate change YouTube video has a lot of views versus only a few views, individuals who watch the video will be more likely to believe that climate change is an important issue to their fellow viewers, and some individuals will also increase their own perception of issue importance (Spartz et al., 2017). Similarly, when exposed to uncivil comments under a blog article about nanotechnology, individuals will be more likely to perceive bias in the science article as well as greater risks in nanotechnology, compared with when individuals read civil comments under the same science article (Anderson et al., 2014).

Given the deepening infusion of algorithms into everyday life and society (Wagner et al., 2021), researchers have begun to examine how online information platforms and their supporting algorithms shape people's information experiences (see, e.g., Bandy & Diakopoulos, 2021; Hannak et al., 2013; Nechushtai et al., 2023). One particular concern related to algorithmic content curation is that the algorithms underpinning many online information platforms may be biased. For instance, researchers are concerned that the algorithms of search engines and social media may feed users of these platforms content that is systematically or progressively biased toward certain political parties or ideologies, potentially radicalizing members of the audience along political lines (Hosseinmardi et al., 2021; Ribeiro et al., 2020; Robertson et al., 2018). In addition, how social media algorithms may prioritize inaccurate or harmful content in their recommendations has also received some scholarly attention (Hussein et al., 2020; Papadamou et al., 2021; Srba et al., 2023).

Nonetheless, results from this body of empirical research are mixed at best, and more studies are needed to continue to untangle whether and how social media algorithms systematically prioritize certain content and create or reinforce information silos. For example, research has generated inconclusive findings regarding the extent to which YouTube recommendations are politically biased. While some researchers found that users consistently migrate from politically milder content to more extreme far-right content on YouTube (Ribeiro et al., 2020), others found no evidence supporting the claim that YouTube recommendations systematically prioritize far-right content or fuel user radicalization over time (Hosseinmardi et al., 2021). In addition, there is little empirical attention to how the algorithms of online information platforms disseminate scientific information in society. Moreover, virtually no research has examined how algorithms may shape the science information experiences of diverse audiences who have distinct racial makeup and SES and are traditionally underserved by science communication and outreach efforts (Christopherson et al., 2021).

This dissertation therefore explores the extent to which algorithmically driven social media platforms recommend science content based on users' sociodemographic characteristics. Using YouTube as a case study, in Study A I explore how the platforms' algorithm might make content recommendations to diverse audiences based on their racial and socioeconomic profiles as they search for three science issues including HGE, AI, and COVID-19 vaccines on YouTube. Insights into these processes of algorithmic curation of science content provide evidence for the social picture of diverse audiences' science information experiences in new media environments. Further, they could facilitate efforts to pre-empt or counter structural barriers in our information environments that prevent the more equitable distribution of scientific information in society. If social media platforms such as YouTube are indeed exposing different parts of the audience to systematically different content, even when they are actively searching for the same science topic, then we are more likely than not to witness ever-increasing divergence in knowledge and perceptions of science across social segments. Before diving into these possibilities, however, it may be helpful to first consider the nature of inequalities in scientific understanding and what we know about how new media environments are shaping such inequalities.

New media, old phenomenon? Disparities in scientific understanding in the current information environment

The primary outcome variable of interest in this dissertation is knowledge about wicked science issues. As discussed earlier, public understanding of science matters for a multitude of reasons. From a societal perspective, it strengthens economies and economic competitiveness, leading to a higher standard of living (NASEM, 2016b); further, it is an important part of liberal education as science and technology are a defining feature of modern Western societies and culture (NASEM, 2016b). From a personal perspective, an understanding of science often helps individuals make informed decisions and better respond to challenges that arise in their personal and community circumstances, leading to richer and healthier lives (NASEM, 2016b). Moreover, from a democratic perspective, at the core of any deliberative democracy is the notion that members of society are willing to and capable of making informed decisions regarding issues of social importance (King, 1928), such as emerging science and technologies that can have highly disruptive impacts on many areas of society (Scheufele, 2022). In an era of post-normal science where science and technology issues often raise a host of ethical, social, and political questions that do not have purely technical answers, it is a democratic imperative to engage broad publics and stakeholder groups in the dialogue to ensure the quality of the resolution to these questions (Funtowicz & Ravetz, 1993). Scientific understanding matters in part because it provides the basis for one's ability to engage with science-related issues (NASEM, 2016b). As modern science and technologies are extremely complex, in order for informed public debate involving those issues to happen, it is necessary that citizens have certain scientific awareness and some baseline knowledge of the issues at hand (European Commission, 1995). Knowledge that cuts across multiple dimensions related to the science, application, history, and policy of a given issue matters because public debate involving modern science and technologies often engages the ethical, legal, and social implications of those issues (Sarewitz, 2015).

In this dissertation (particularly Study B and Study C), I focus on science knowledge, specifically knowledge about scientific facts related to a wicked science issue (e.g., HGE, AI, COVID-19), as the outcome variable of interest. Science knowledge is usually measured by true-or-false questions about scientific facts. In addition, perceived science knowledge, or how knowledgeable about science or familiar with science people *think themselves are* could also matter for important outcomes of democratic decision-making about science, such as public support for science issues (Akin et al., 2020; Ladwig et al., 2012). Perceived science knowledge is often measured by asking individuals to indicate how informed they think themselves are about science or a specific science issue. In this dissertation, I refer to science knowledge as *factual* science knowledge to distinguish it from *perceived* science knowledge.

Note that science knowledge represents only a small part of what Howell and Brossard term "civic science literacy," which refers to a broad understanding of "the many elements that shape the production of scientific knowledge, such as the people, institutions, training, resources, methods, and norms of science" (2021, p. 2). In other words, civic science literacy includes not only knowledge about scientific facts, but also an understanding of scientific processes, science's impacts on society, and, more importantly, the societal contexts that impact science itself (Howell & Brossard, 2021). In addition to civic science literacy, Howell and Brossard (2021) argue that science literacy comprises at least two other important dimensions: "digital media science literacy" and "cognitive science literacy." Digital media science literacy matters because most people today learn about science through mediated sources especially digital media sources

(Brossard, 2013; Scheufele, 2014). Specifically, digital media science literacy should encompass the ability to access (online) science information, an understanding of how science information travels through (online) media systems, and the ability to evaluate science information presented in mediated messages (Howell & Brossard, 2021). The third important dimension of science literacy—cognitive science literacy—refers to an awareness of one's own biases in processing and evaluating science information, such as one's beliefs, emotion, heuristics, and motivations (Howell & Brossard, 2021). Cognitive science literacy matters because it enables individuals to truly make informed decisions based on science information.

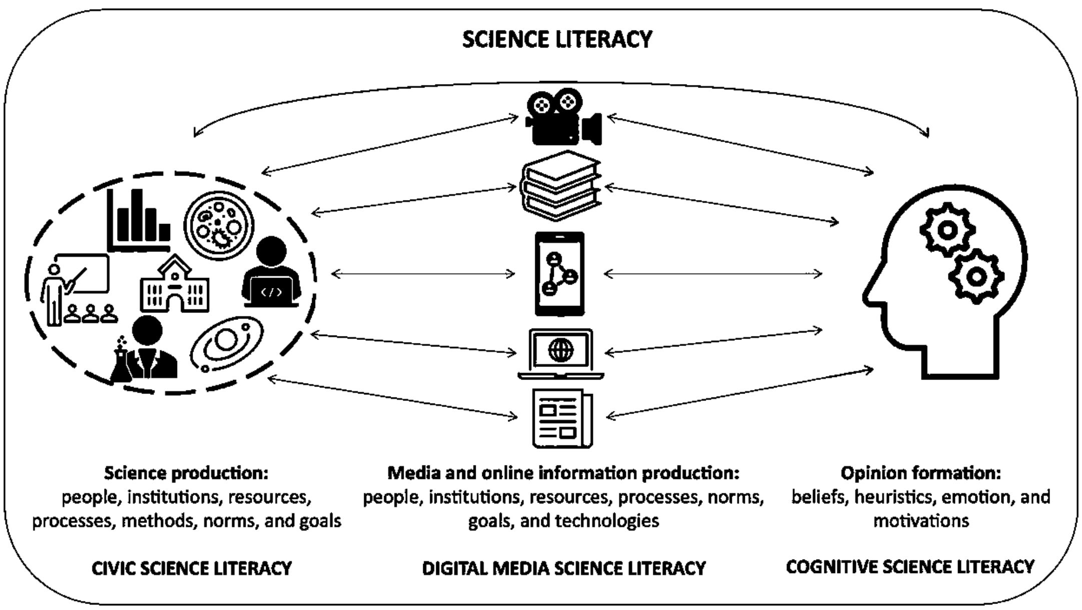


Figure 1.1. Dimensions of science literacy (adapted from Howell & Brossard, 2022).

To what extent science literacy may enable people to accomplish specific tasks, such as selecting information sources, judging expertise and information credibility, and learning from health and science messages, remain promising research directions (NASEM, 2016b). For example, while research has almost exclusively focused on the role of health literacy—motivation and capabilities of individuals to access, use, evaluate, and act on health information and to engage in health-related civic matters (NASEM, 2016b; Nutbeam, 2008)—in affecting people's use and interpretation of health information, science literacy also likely matters, especially when the communicated medical developments involve fast-changing science and high scientific uncertainty such as the COVID-19 vaccines (Scheufele et al., 2020).

Empirical evidence suggests that a sizable portion of the American public has limited knowledge about scientific methods and processes, limited interest in science, and low involvement in science activities (National Science Board, 2022). Perhaps what is more concerning is that such trends are especially pronounced among racial minorities and population segments that are less educationally and financially well-off (Kennedy & Atske, 2019; Kennedy & Hefferon, 2019). There are wide educational differences in science knowledge in the U.S., with more highly educated Americans scoring higher on science knowledge compared to Americans who are less educated, when knowledge is conceptualized as an understanding of scientific facts (related to life sciences, earth and other physical sciences) and processes, as well as numeracy and the ability to read charts (Kennedy & Hefferon, 2019; see Figure 1.2). With these same science knowledge measures, White Americans also tend to have substantively higher levels of science knowledge than Blacks and Hispanics, and such differences persist across domains of scientific facts as well as individuals' education levels (Kennedy & Atske, 2019; see Figures 1.3 and 1.4). Similar sociodemographic gaps are also identified for science interest. For example, White Americans and people who have college degrees report greater interest in following science news than their counterparts who are Black, Hispanic, or less educated (Saks & Tyson, 2022). Corroborating these trends are persistent disparities in racial diversity and inclusion in science, technology, engineering, and mathematics (STEM) jobs in the U.S. (Christopherson et al., 2021). As alluded to earlier, disparities in science knowledge across population segments prevent broader public engagement with science and more equitable and democratic decision-making on policy issues that involve science; therefore, for democratic reasons, it is important that all population segments have equitable access to science knowledge, even if science knowledge has limited impacts on individuals' perceptions of, attitudes toward, and behavior related to science (Allum et al., 2008; NASEM, 2016b).

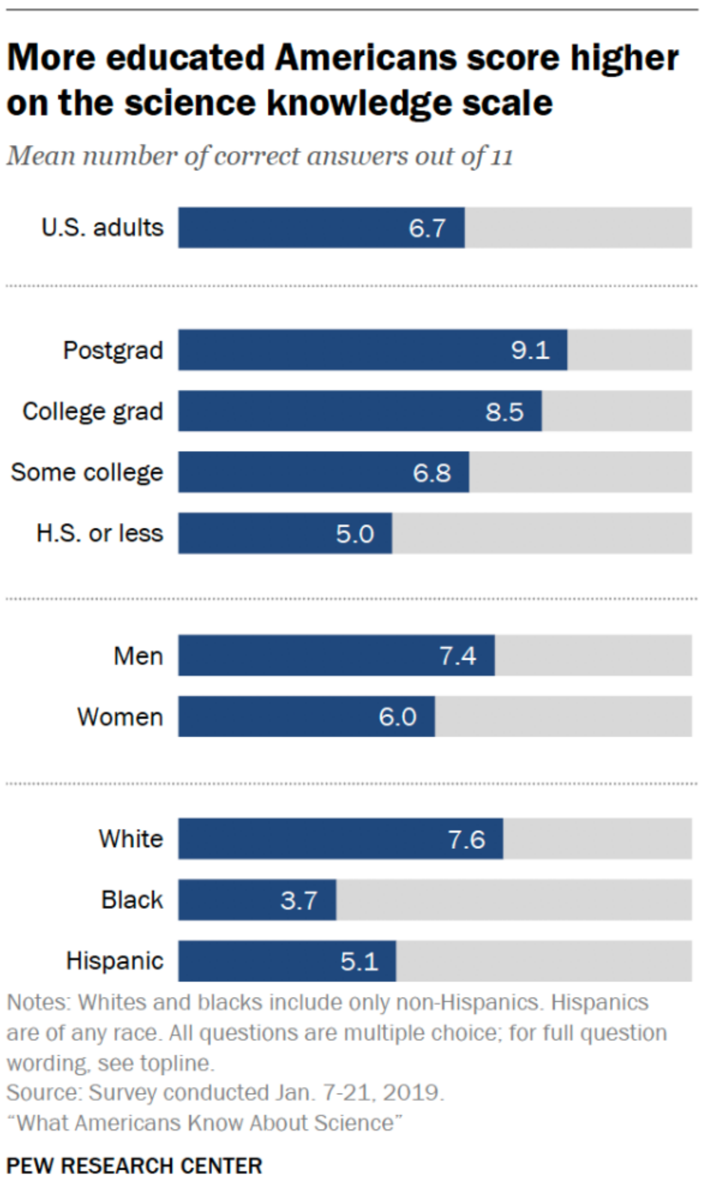


Figure 1.2. More educated Americans score higher on science knowledge measures (adapted from Kennedy & Hefferon, 2019).

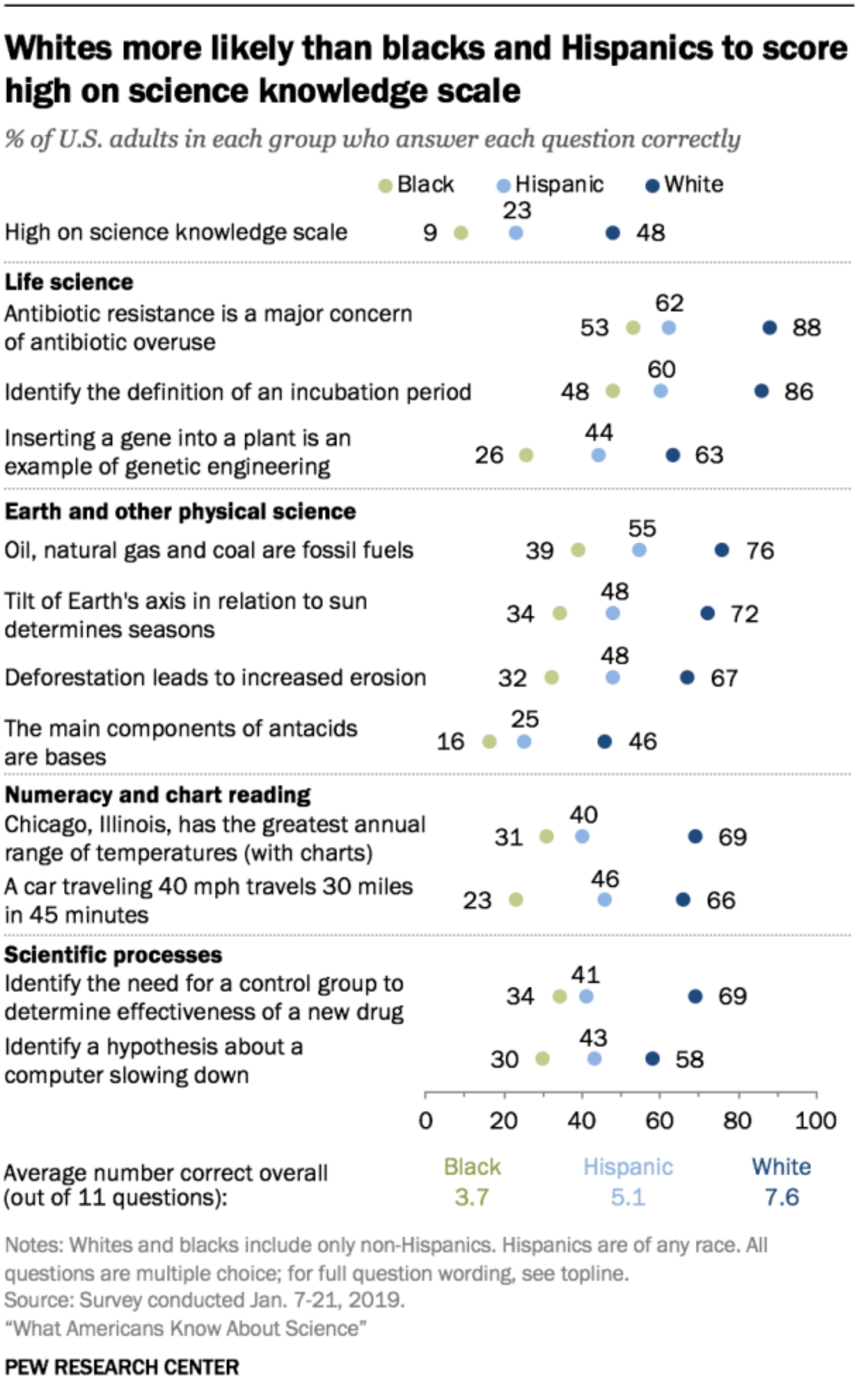


Figure 1.3. White Americans score higher on science knowledge measures than Black and Hispanic Americans (adapted from Kennedy & Atske, 2019).

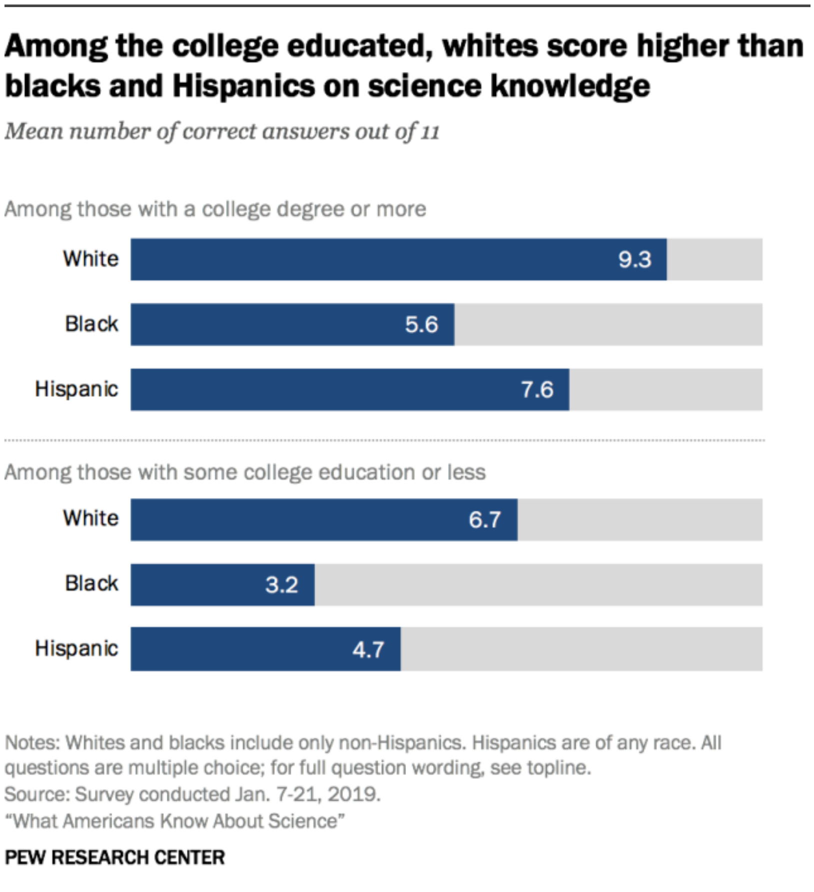


Figure 1.4. Among college-educated Americans, Whites score higher on science knowledge measures than Blacks and Hispanics (adapted from Kennedy & Atske, 2019).

Disparities in science knowledge between different social segments, like those discussed above, can be understood through the lens of digital divides and the knowledge gap hypothesis. Representing "who has access to online tools and information," digital divides exist along socioeconomic, racial, and other demographic lines (Howell & Brossard, 2021, p. 4). In the U.S., those who are White, male, wealthy, and well-educated are more likely to have access to and benefit from digital technologies and information (Christopherson et al., 2021; Kiser & Harrison, 2018; Voss, 2018). Further, even when given equal physical access to information technologies and resources, those who are already advantaged in society have different patterns of use and can often gain the most from informational resources (Howell & Brossard, 2021).

Another useful framework for understanding differential gains in knowledge from informational resources is the knowledge gap hypothesis (Tichenor et al., 1970), which states that members of society with different SES do not learn from media information equally fast. Specifically, high SES individuals are able to extract meaningful knowledge from media information more efficiently than low SES individuals. Therefore, as media information circulates in society, high SES audiences will experience greater knowledge growth than low SES audiences, widening existing gaps in knowledge between high and low SES social segments (Tichenor et al., 1970).

Among other factors, the nature of media systems that deliver information likely influences how knowledge gaps turn out. Early research on the effects of media systems on the formation of knowledge gaps typically focused on newspapers and televised news. Studies have generally found that newspaper reading is associated with widened knowledge gaps between high- and low-SES segments for a multitude of economic, motivational, and literacy reasons (e.g., Eveland & Scheufele, 2000). In addition, because many science issues tend to have short attention cycles in a newspaper (Anderson et al., 2013), it is possible that newspaper coverage of science issues wanes too soon before low education segments are able to catch up and close the knowledge gaps. In contrast, television viewing has been generally found to be linked to reduced SES-based knowledge gaps (Cacciatore et al., 2014; Chang et al., 2018). Several reasons may explain this. From a media content perspective, the audiovisual component of television content often helps contextualize and reinforce information, making it easier to grasp and retain information shown in television (Graber, 1990; Neuman et al., 1992). In addition, because television content aims to reach the widest audiences possible, it often requires less cognitive efforts from the audiences' end, making the content more accessible and comprehensible in general (Neuman et al., 1992). Relatedly, television content tends to be meager and superficial in its coverage of public issues, at least compared with newspapers, which imposes a ceiling effect on high SES audiences' learning as these audiences can only acquire limited new knowledge from the television (Ettema & Kline, 1977). From an audience perspective, research shows that individuals with lower education are particularly good at encoding, storing and retrieving television news information whereas individuals with higher education have better memory for print and web news (Grabe et al., 2009). It is also possible that high SES audiences are not as motivated to learn from television as low SES audiences, thus creating a self-imposed ceiling effect on their learning from television content (Ettema & Kline, 1977).

More recently, with the decline of traditional media infrastructures for science communication and the emergence of digital media as a primary source of science information for many people (Brossard, 2013; Scheufele, 2013), scholars have started to turn attention to various online media, examining how use of the Internet (Cacciatore et al., 2014; Lee, 2009; Shim, 2008), science blogs (Su et al., 2014), online newspapers (Chang et al., 2018; De Silva-Schmidt et al., 2022), and social media (Gerosa et al., 2021) might affect SES-based gaps in science knowledge. Of particular interest to this dissertation is the role of social media, which have become a prominent general news source especially among younger generations (Newman et al., 2022). Most social media users in the U.S. encounter science-related information on these platforms (Funk et al., 2017b). Before reviewing existing evidence of the relationship between social media use and science knowledge gaps, it may be useful to first consider the nature of social media.

Broadly speaking, social media refer to “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content” (Kaplan & Haenlein, 2010, p. 61). Within this broad conceptualization, social media may be further differentiated along two dimensions: a media-related dimension that concerns media richness or the level of social presence allowed to emerge between two interactants using a social media platform; and a social dimension pertaining to the amount of self-disclosure required and the type of self-presentation allowed on the platform (Kaplan & Haenlein, 2010).

With respect to the first dimension, a social media platform is considered to have greater media richness (i.e., transmitting more information within a given time frame) or enables greater social presence if it can achieve (a) a greater degree of multimodal communication—such as visual, acoustic, and physical contact; (b) a greater degree of communication immediacy via synchronous, as opposed to asynchronous, communications; and (c) a greater degree of communication intimacy via interpersonal, as opposed to mediated, communications (Kaplan & Haenlein, 2010).

With respect to the second dimension, because people have the inclination to manage the way others perceive them during any social interactions (Goffman, 1959), content generated and interaction engendered on social media will also involve self-presentation of the interactants involved, which is often done through self-disclosure, or the revelation of personal thoughts, feelings, experiences, and other information (Kaplan & Haenlein, 2010). The degree to which personal disclosure is permitted varies across social media platforms, as these platforms may be more or less focused on specific content domains or ruled by strict guidelines that force users to behave in a certain way (Kaplan & Haenlein, 2010).

Combining these two dimensions, Kaplan and Haenlein (2010) categorized social media into six groups, including blogs, collaborative projects, social networking sites, content communities, virtual social worlds, and virtual game worlds. Of increasing relevance is metaverse, which is characterized by high media richness/social presence and high self-presentation/self-disclosure. Figure 1.5 shows a classification of social media based on the two dimensions of media richness/social presence and self-presentation/self-disclosure. Although 13 years old, this classification is still relevant nowadays.

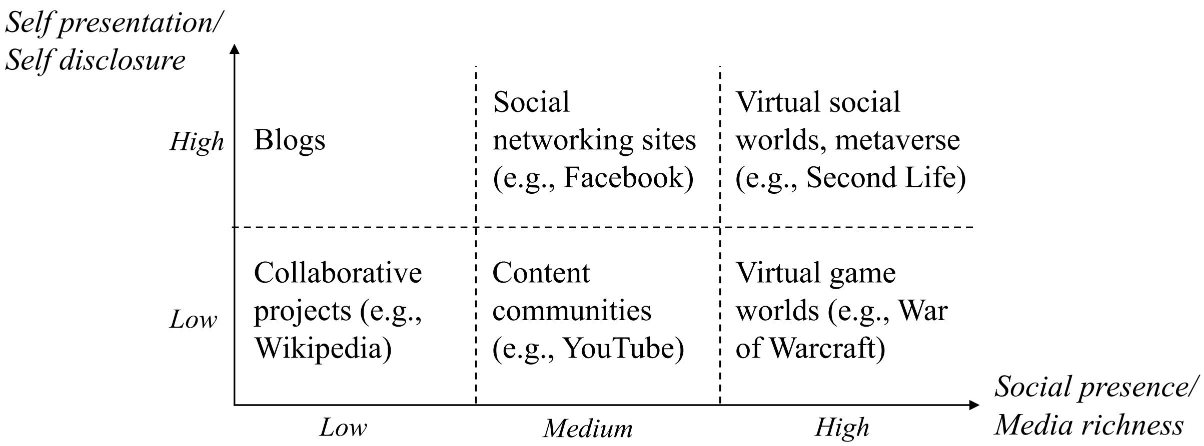


Figure 1.5. Social media classification (adapted from Kaplan & Haenlein, 2010).

Research examining how social media use might affect science knowledge has overall identified a negative relationship between social media use and knowledge, although the relationship may depend on how key variables are measured. In the U.S. context, a survey of a national sample of Americans found that those who frequently got their information about COVID-19 through social media reported lower levels of factual knowledge about COVID-19 (Gerosa et al., 2021). Another survey of U.S. adults found that frequent use of social media in general was linked to lower levels of factual knowledge about politics, although the same pattern was not identified for factual health knowledge (Li & Cho, 2021). Americans who frequently used Facebook in particular were less aware of gene editing; however, individuals who spent more time on Facebook on a daily basis reported higher gene editing awareness, as well as those whose Facebook networks were more diverse (Mueller-Herbst et al., 2020). Some of these patterns also hold in other national and cultural contexts. A national survey of South Koreans showed that those who paid more attention to science information on social media reported less factual science knowledge but at the same time perceived themselves to be more knowledgeable about science (i.e., higher perceived science knowledge; Chang et al., 2018).

Further, initial evidence appears to suggest that social media use has the potential to mitigate knowledge gaps formed on the basis of SES or education, at least in the context of political knowledge (Li & Cho, 2021); it is unclear, however, whether social media use would narrow science knowledge gaps among education and racial groups. Higher social media engagement (operationalized as sharing, liking, commenting, and reading comments on social media news posts) is associated with a smaller education-based gap in factual knowledge about politics, but not about health (Li & Cho, 2021). Additional studies also failed to show similar effects of social media use on education-based gaps in factual knowledge about science in general (Chang et al., 2018) or about COVID-19 specifically (Gerosa et al., 2021).

The structure of social media networks could also matter, as more diverse and denser social media networks seemed to contribute to a smaller gap in factual political, but not health, knowledge between education groups (Li & Cho, 2021). Finally, paying attention to science information on social media is linked to a narrower gap in perceived science knowledge between high and low education groups, indicating that social media use might enhance the perception of knowledge about various science issues especially among people with lower levels of educational attainment (Chang et al., 2018).

Taken together, this dissertation builds on this emerging line of research by asking and answering two main questions in Study B: What roles do use of social media platforms play in shaping science knowledge gaps between education segments and racial groups? Further, do these patterns vary depending on the specific science issues examined? Answers to these questions facilitate our understanding of disparities in science knowledge in an era of social media and wicked science. Further, they could inform efforts to accelerate greater use of equity-based communications strategies to fortify a more democratic civic science society. In the next section, I review and compare the three wicked science issues that make up the study contexts of this dissertation.

Contexts of study: Human gene editing, artificial intelligence, and COVID-19 as three wicked science issues

For study contexts, this dissertation will focus on three science issues: human gene editing (HGE), artificial intelligence (AI), and coronavirus disease 2019 (COVID-19) and its related vaccines. These represent great study context because, as I will discuss below, they are good exemplars of wicked science issues (Rittel & Webber, 1973). HGE refers to applying genome-editing technology—a method for making changes to an organism's genetic material by adding, replacing, or deleting deoxyribonucleic acid (DNA) from the genome sequence (NASEM, 2017b)—to humans. HGE can be distinguished by its purpose (for therapeutic purpose versus for enhancement purpose) and heritability (editing to somatic cells versus editing to heritable germline cells; NASEM, 2017b). Developed and applied across many fields from engineering to medicine to economics (Stone et al., 2016), AI broadly refers to using machines and computer systems to mimic human intelligence or perform tasks that would normally require human intelligence, making "predictions, recommendations, or decisions influencing real or virtual environments" for a given set of human-defined objectives (*National Artificial Intelligence Initiative Act*, 2020). Finally, COVID-19 is an infectious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which often causes respiratory symptoms such as fever or chills, sore throat, and difficulty breathing, among others (Centers for Disease Control and Prevention [CDC], 2023c). As part of the rapid response to the global COVID-19 crisis, different types of vaccines were developed to help people build protection from the disease (CDC, 2023c).

These three science issues are all likely to have profound impact on society (and already are) and bring about complex social, political, and ethical questions that do not have purely technical answers. They are examples of "wicked problems" as they (a) have implications for policy and people's lifestyle, (b) involve trade-offs between competing values held by different stakeholders, and (c) have no definitive best solution moving forward (Rittel & Webber, 1973). For instance, employers requiring workers to be vaccinated against COVID-19 at public and private institutions during the pandemic in 2021 raised intense political debate within the U.S., as some members of the public viewed such mandates as an infringement on personal liberties and religious freedom, while others uphold them as an imperative to avoid harms to others. Similarly, while HGE has been lauded for its potential to treat diseases that are otherwise difficult or impossible to cure and AI for its ability to enhance human decision-making, both technologies raise a broad range of concerns such as possibly worsening social equity, and to address those concerns requires more than scientific expertise alone.

In addition, these science issues can also be understood through the framework of "post-normal science." Post-normal science refers to areas of science that involve high systems uncertainties or high decision stakes (see Figure 1.6), where "facts are uncertain, values in dispute, stakes high and decisions urgent" (Funtowicz & Ravetz, 1993, p. 744). In other words, post-normal science raises difficulties not in the sense of discovering a scientific fact but in the sense of understanding and managing a fundamentally complex situation that often involves various external interests—benefits, costs, and value commitments—through various stakeholders (Funtowicz & Ravetz, 1993). Therefore, addressing the difficulties raised by post-normal science requires a different set of problem-solving strategies than those required by more traditional forms of science (e.g., applied science, professional consultancy). Specifically, an extended peer community—one that includes not just scientists and official experts but also anyone who has a stake in the issue, such as citizens, pressure groups, and investigative journalists—is essential for assuring the quality of scientific inputs into policy-making processes (Funtowicz & Ravetz, 1993).

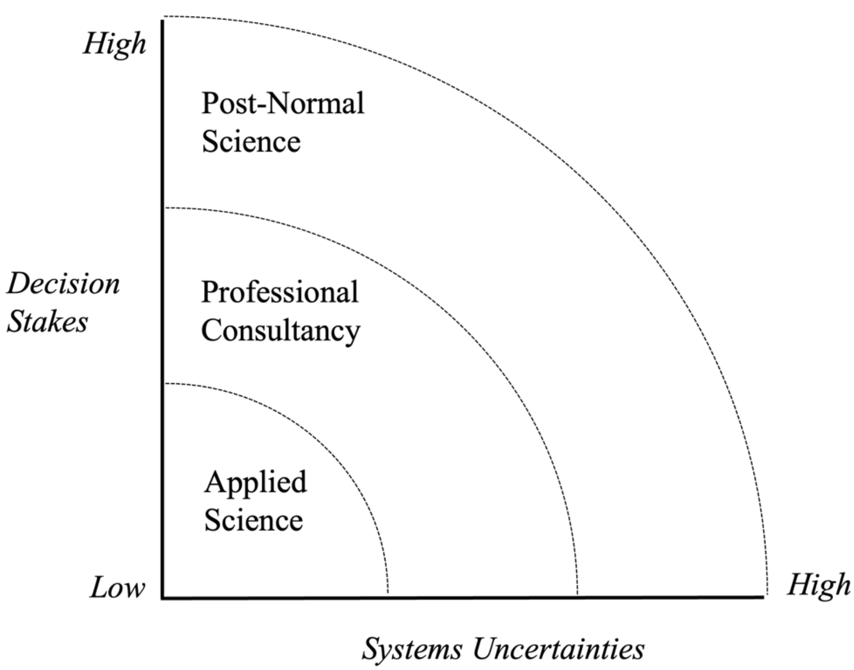


Figure 1.6. Post-normal science (adapted from Funtowicz & Ravetz, 1993).

Aside from their commonalities, the science issues examined here also differ in a number of ways. A useful framework for understanding these issues' differences is the issue attention cycle model (Downs, 1972), which states that many issue topics in public life follow a life cycle where public and media attention to these issues goes up and down (see Figure 1.7). An issue starts with receiving no or very little media attention and public concern. Then, in the alarmed discovery phase, there is a sudden increase in media and public attention to the issue, usually driven by a focusing event (e.g., government mandating COVID-19 vaccination). Next, in the mobilization phase, media coverage of the issue problems continues to grow and increasingly pays attention to the political conflicts surrounding the issue as more actors participate in the issue negotiation. After that, media and public attention starts to gradually decline as policy measures are being developed to solve the problems. News coverage at this stage tends to focus on the technical aspects of the issue (Nisbet & Huge, 2006). The issue cycle continues, and public and media attention tends to remain higher at the end of the cycle than before issue discovery. Figure 1.7 gives a broad representation of the position of each issue in the issue attention cycle.

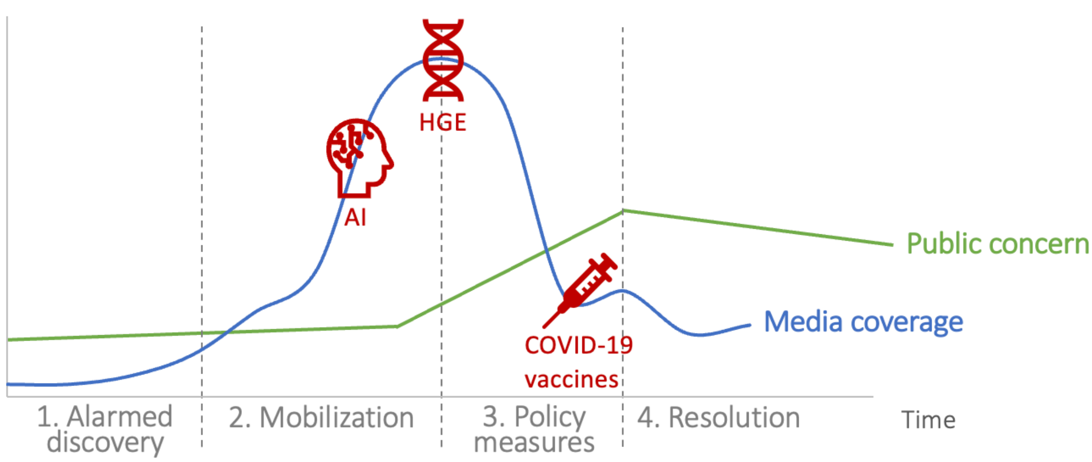


Figure 1.7. HGE, AI, COVID-19 vaccines and the issue attention cycle.

A cursory examination of the level of media and policy attention related to the three issues used as context in this dissertation, allow us to compare and contrast the three issues (see Table 1.1) in relation to the figure above.

Table 1.1. Total volume of U.S. news coverage, congressional hearings, and federal register on HGE, AI, and COVID-19 from January 2010 to May 2023.

	HGE	AI	COVID-19 in general	COVID-19 vaccines
News coverage	94	2,065	19,593	3,604
Congressional hearings	2,397	2,246	1,668	764
Federal register	1,280	522	12,254	810

Sources: Data on volume of news coverage come from Newspaper Source Plus index of *New York Times, Wall Street Journal*, *Washington Post*, and *USA Today*. Data on number of congressional hearings held and federal register come from ProQuest Congressional Record (U.S.). The Boolean search string is (human*) AND ((gene edit*) OR (genome edit*) OR (genetic engineer*)) for HGE; artificial intelligence for AI; covid* for COVID-19; and (covid*) AND (vaccin*) for COVID-19 vaccines.

Arguably, COVID-19 and its related vaccines are further into the policy measures stage along the issue attention cycle (see Figure 1.7). There are already many policy measures in place regarding COVID-19 and the related vaccines. Coverage of COVID-19 including the related vaccines was highly saturated in the news (Table 1.1) and has passed its peak, gradually declining. The pandemic also raised widespread public concern within the U.S., as nearly half (45%) of the U.S. adult population named COVID-19 as the nation's most important problem in April 2020, topping the list of a host of public issues (Brenan, 2021).

In comparison, HGE has just entered the policy measures stage in the issue attention cycle (Figure 1.7), with 1,280 federal regulatory documents being released and 2,397 congressional hearings held on HGE since 2010 (Table 1.1). The National Academies of Sciences, Engineering, and Medicine (NASEM) also issued two consensus reports on gene-editing technology—one on genetically engineered crops (NASEM, 2016a) and the other on human genome editing specifically (NASEM, 2017b). Media and public attention to HGE surged when the birth of the first gene-edited babies was announced in 2018 and when the developers of CRISPR-Cas9—a novel gene-editing technique—won the 2020 Nobel Prize in Chemistry.

AI is still at an early stage in the issue attention cycle relative to COVID-19 vaccines and HGE (Figure 1.7). Particularly, while AI has received some policy attention, the extent of regulatory developments on AI is not comparable to that on COVID-19 (and its related vaccines) or HGE (Table 1.1). Nonetheless, with the advent of ChatGPT and other highly disruptive AI tools, we should expect that media, public, and policy attention to AI technology will continue to grow.

Moreover, the examined science issues also differ in level of wickedness or post-normalness. First, these issues involve varying degrees of uncertainty related to the production, processes, and implications of the science (Funtowicz & Ravetz, 1993), with COVID-19 vaccines arguably involving lower scientific uncertainty than AI or HGE. While the COVID-19 vaccines are relatively new developments, the science behind the different types of COVID-19 vaccines has been around for decades. In the U.S., there are three types of COVID-19 vaccines for use, namely mRNA vaccines, protein subunit vaccines, and viral vector vaccines (CDC, 2023d). Researchers have studied the mRNA technology, the protein subunit technology, and the viral vector technology for decades and used them in vaccines for other infectious diseases prior to COVID-19 (Cid & Bolívar, 2021; Fang et al., 2022; Vrba et al., 2020). Decades' research on these vaccine technologies as well as on vaccines for other types of coronaviruses helped accelerate the initial development of the COVID-19 vaccines. Once developed in the laboratory, the COVID-19 vaccines went through all phases of clinical trials with tens of thousands of human subjects from different age, racial, and ethnic background; in addition, the vaccines have undergone “the most intense safety monitoring in U.S. history” after being approved and recommended for use (CDC, 2023d). Following all these established procedures for vaccine development and approval helped reduce the scientific uncertainty related to the safety and effectiveness of the COVID-19 vaccines, improving the quality of the scientific evidence. Further, because COVID-19 vaccines are a transient science issue and have a limited issue scope (e.g., scope of application and implications), the related scientific uncertainty is also constrained by the temporality and scope of the issue.

In contrast, HGE and AI are more wicked than COVID-19 vaccines. Whereas COVID-19 is much more transient, has a constrained issue scope, and involves high, tangible personal risks, HGE and especially AI have a very broad range of applications with long-term, far-reaching impacts at the societal level. What this means is that there is no immediate or ultimate test to a solution to the problems raised by HGE and AI, as there will be waves of unforeseeable consequences over an extended—even unbounded—period of time once a solution is implemented (Rittel & Webber, 1973). In addition, once a solution is implemented, it cannot be reversed and the repercussions from the action are often at very large scales (Rittel & Webber, 1973). Consider, for example, that once HGE is applied to making heritable changes in humans, the genetic changes will be passed down to future generations and there is no turning back. Moreover, HGE and AI have very high “completeness uncertainties” (Funtowicz & Ravetz, 1993, p. 744), meaning that it is almost impossible to enumerate all the significant consequences and relationships in the formation of the issue problems. These issues do not have an enumerable (or an exhaustively describable) set of potential solutions given their broad scope, nor is there a well-described set of permissible operations that may be incorporated into the solution plan (Rittel & Webber, 1973).

Second, the examined science issues also evoke different types of value-based divide and may produce social conflicts based on different systems of worldviews. While controversy surrounding COVID-19 and its related vaccines is primarily politically driven (Galston, 2021), controversy around HGE is less politically motivated and instead motivated by moral and religious concerns (NASEM, 2017b).

Dissertation overview

In this dissertation, I conduct three related but independent studies that answer the research question of how media- and message-level factors might affect different social segments' knowledge of wicked science issues in today's algorithmically infused media environment. My dissertation is organized as follows:

Chapter 2 presents study A, which uses a novel sock puppet algorithm audit design to explore the extent to which social media algorithms, such as the YouTube algorithm, recommend information based on users' racial and SES profiles as they search for science issues online, as well as how the algorithmic content recommendations might differ across audience groups. Findings from the study provide an empirical understanding of how social media infrastructures shape the distribution of science information in society and inform efforts to pre-empt or counter structural barriers in our information environments that prevent more equitable science communication outcomes.

Chapter 3 presents Study B, which uses three public opinion survey datasets to investigate how the use of algorithmically driven social media platforms including Facebook, Twitter, YouTube, Instagram, and TikTok may influence race- and education-based gaps in both factual and perceived knowledge of wicked science issues, including HGE, AI, and COVID-19 vaccines. Building on knowledge gap research, findings from this study facilitate our understanding of race- and education-based disparities in factual and perceived knowledge of science issues an era of social media and wicked science.

Chapter 4 presents Study C, which uses an online survey experiment to test the relative effectiveness of visual and narrative components of COVID-19 vaccine safety information on enhancing information understanding and elaboration among individuals with varying levels of science literacy. Finding from this study inform the design of equity-based science and health communication strategies for reaching a broad audience within society.

The final chapter summarizes the research findings of this dissertation and discusses its theoretical, methodological, and practical implications, with a focus on facilitating more equitable science communication in an era of algorithmic information curation.

Chapter 2

Biased algorithm? Exploring how YouTube recommends science videos to racially and socioeconomically diverse audiences (Study A)

As discussed in chapter 1, people today are getting information, including science information, primarily from online media, which are largely driven by AI algorithms that narrowcast information toward audience members based on their demographics, preferences, and a wealth of user histories and digital trace data (Brossard & Scheufele, 2022). Accordingly, researchers have begun to investigate the impacts of algorithmic information curation on people's information experiences, such as how the AI algorithms underpinning online media platforms might be biased. This body of research has generated mixed evidence at best, and to my knowledge virtually no research has examined how online media algorithms disseminate science information across diverse audiences with distinct racial makeup and SES, especially audiences that are traditionally underserved by science communication and outreach efforts. Therefore, using YouTube as a case study, in this chapter I explore the extent to which social media algorithms make content recommendations based on users' racial and socioeconomic profiles as they search science topics. I first consider the relevance of YouTube for communicating and consuming science content.

YouTube as a window for science communication

Founded in 2005 and purchased by Google in 2006, YouTube is the world's second-largest search engine (Davies, 2021) and second-most visited site after Google.com (Similarweb, 2023). The platform's exceptionally far reach makes it an impactful venue for communicating science. Globally, YouTube attracts 1.7 billion unique visitors and gets 14.3 billion visits every month (McLachlan, 2022). About 81% of all Internet users have used YouTube (McLachlan, 2022). Younger users are more active on YouTube. Specifically, about 26% of YouTube users are between ages 18-24, 30% are between 25-34—making the largest age group, and 18% are between 35-44 (Similarweb, 2023). Nearly 60% of YouTube users identify as male and 40% as female (Similarweb, 2023). Within the U.S., YouTube is the most widely used social media platform among the adult population, with 81% of U.S. adults reporting ever using the platform and 54% reporting visiting the site daily as of 2021 (Auxier & Anderson, 2021). Fully 95% of U.S. adults between ages 18-29 report using YouTube, along with 91% of those between 30-49 and 83% of those between 50-64 (Auxier & Anderson, 2021). Approximately 79% of White Americans and 84% of Black Americans report using YouTube. SES-wise, 70% of U.S. adults who have completed high school education or less report using YouTube, and 89% of college graduates say they use the platform (Auxier & Anderson, 2021). Among those earning less than $30k annually, 75% report using YouTube, and 90% of those with an annual income of $75k or more say they use YouTube (Auxier & Anderson, 2021). Given the extensive reach of YouTube, it is unsurprising that any content—including science content—is more likely to be found by Internet users if it has a YouTube presence.

Moreover, science and technology is one of the most prominent video categories on YouTube (Yang et al., 2022). Some of the most popular science-themed channels enjoy tens of millions of subscribers and billions of views (Feedspot, 2023). Relative to the flourishing of science content on YouTube, however, we know very little about the audiences that consume such content—their demographics, motivations, and experiences. Initial evidence from analysis of one of the fastest-growing YouTube science channels suggests that young males may be the most active and engaged audience of science videos on YouTube (Yang et al., 2022). A 2018 study surveying a convenience sample of Singaporeans found that people sought YouTube science videos out of enjoyment of science, their informational (versus entertainment) use of YouTube, and their perceptions of the extent to which people around them engage in or approve of watching YouTube science videos (Rosenthal, 2018).

Still, we are largely ignorant of how racially diverse and socioeconomically disadvantaged audiences—those traditionally underserved by science outreach efforts—consume science content on the YouTube platform. Further, because social media algorithms are designed to maximize and monetize user engagement by tailoring information to individuals' social context, preferences, and digital traces (Brossard & Scheufele, 2022), YouTube's recommendation algorithm may plausibly expose racially and socioeconomically diverse individuals to different subsets of (science-related) information based on what the algorithm deems personally relevant to the users, even when they are seeking information on the same topic. To my knowledge, no previous research has directly examined these phenomena. Nonetheless, before taking a closer look, it makes sense to first consider the nature of YouTube's content recommendations and the factors that are known to shape them.

Factors shaping YouTube's content recommendations

YouTube recommendations are a major way of getting content discovered, as 70% of the content watched on YouTube is recommended by YouTube's algorithm (Rodriguez, 2018). Content recommendations on YouTube can take different forms, such as (a) homepage video highlights, or a list of videos that show up on a user's homepage when the user first opens YouTube; (b) related video recommendations, or videos appearing in the "Up-Next" panel that are suggested based on what one is currently watching, other videos that are similar to the video currently playing, as well as what other users with similar tastes are watching; and (c) search results recommendations, or the list of videos returned after one types a search query into the YouTube search box and initiates the search. Query-based search results recommendations are the top source of viewership, accounting for over two-thirds of video views on YouTube (Zhou et al., 2010). Given the primary role of search results recommendations in driving viewership on YouTube, this study analyzes content that gets returned after users search for particular science topics on YouTube.

Research has identified a host of factors as sources of influence on how the YouTube algorithm recommends content (e.g., Abbas et al., 2017; Borghol et al., 2012; Covington et al., 2016; Davidson et al., 2010; Smith et al., 2018). Those factors can be roughly categorized into two interactive clusters: one cluster primarily relates to the supply side of video production (video factors) and the other cluster mainly concerns the demand side of video consumption (user factors). On the supply side, video characteristics such as video title, description, length, thumbnail, and tags may all influence how likely a video will be recommended to prospective viewers. Longer videos, for example, may be more likely to be recommended by the YouTube algorithm in part because they generate more watch time (Smith et al., 2018), thus conducive to YouTube's goal of keeping users on the platform watching videos (YouTube Creators Academy, 2017), even though longer videos overall attract fewer views and are watched for smaller percentages, meaning that users are more likely to quit longer videos sooner than if watching shorter videos (Yang et al., 2022). On average, videos are more likely to be viewed if their titles convey strong negative sentiment, their descriptions provide lots of information, they use more—up to 17—tags (Tafesse, 2020), or they include high-quality thumbnails (Hoiles et al., 2017), and videos with more views, along with more likes, shares, and comments, are more likely to be recommended by the YouTube algorithm than videos scoring low on those engagement metrics (Dean, 2017).

In addition, characteristics of the channel that warehouses a video could also affect the likelihood of a video being recommended to other YouTube users. The size and structure of a channel's network, for instance, significantly drive the popularity of the channel's videos (Yoganarasimhan, 2012). The number of subscribers to a channel and the number of videos published by a channel are both positively related to the viewership of a channel's videos, even after controlling for the channel's age (Lopezosa et al., 2020). Videos from channels with a regular upload schedule are also more likely to be recommended by the YouTube algorithm (Hoiles et al., 2017).

On the demand side—which is of particular interest to this study—are factors that are user-dependent. A user's personal activities online—such as channel subscriptions, watch and search histories, and engagement patterns—shape the steady stream of content spit out by the YouTube algorithm. For instance, clicking on a video could indicate a strong interest in the content; based on this information, YouTube may recommend further videos with similar content to the user (YouTube, n.d.b). Liking and sharing a video also signal user satisfaction with the video and YouTube draws on these user engagement activities to predict the likelihood that further videos will also be liked or shared by the user to optimize its content recommendations (YouTube, n.d.b). On the other hand, if the user dislikes a video, the YouTube algorithm will respond to that by recommending less of such content to the user (YouTube, n.d.b). The YouTube algorithm also leans heavily on a user's past watching and search histories in personalizing content recommendations. Different users could get completely different sets of search results even for the same query. For example, when two users search for "cricket," the one that watches a lot of sports videos would likely end up with video recommendations featuring the sport cricket whereas the other user may primarily get nature videos with crickets in them (YouTube, n.d.a). Finally, YouTube also takes into account contextual cues such as a user's geographic location and time of day when making personalized video recommendations. By leveraging this information, YouTube is able to recommend locally relevant content to the user (YouTube, n.d.b).

In sum, little is known about how the YouTube algorithm might recommend (science) videos to *diverse* audiences who come from distinct racial and socioeconomic backgrounds. Therefore, I pose the following research question:

> **RQ1:** As users search information on wicked science issues on YouTube, to what extent does YouTube recommend videos based on users' racial and SES profiles? In what ways do the recommended videos differ across those groups?

Methods

Sock puppet algorithm audit

To understand and measure the opaque algorithmic influences in the YouTube platform, this study used a sock puppet algorithm audit design (Hussein et al., 2020; Sandvig et al., 2014). A sock puppet algorithm audit involves researchers using computer programs to impersonate users, usually by generating programmatically construed traffic, and then having the impersonated users (also called "sock puppets" or "agents") interact with the digital platform to observe how the algorithm behaves accordingly (see Figure 2.1 for an illustration). In this study, a sock puppet was implemented as an automated web browser instance that interacted with YouTube by searching, watching videos, and recording video recommendations in YouTube.

The sock puppet audit approach has a number of methodological advantages. First, it prevents the issue of infringing user privacy as it does not involve tracking real users' digital or web behaviors that may reveal private information about the users. Second, a sock puppet audit design enables researchers to more effectively investigate sensitive topic domains such as misinformation consumption on digital platforms (e.g., Hussein et al., 2020; Srba et al., 2023); it also allows researchers to assign sock puppets to categories that are normatively important but difficult to talk about, such as groups living in poverty, sexual minorities, and groups with stigmatized health conditions (Sandvig et al., 2014). Last but not least, because a sock puppet audit design is a form of field experiment, it allows for strong causal inference by experimental manipulation and preserves external validity by occurring in the natural field.

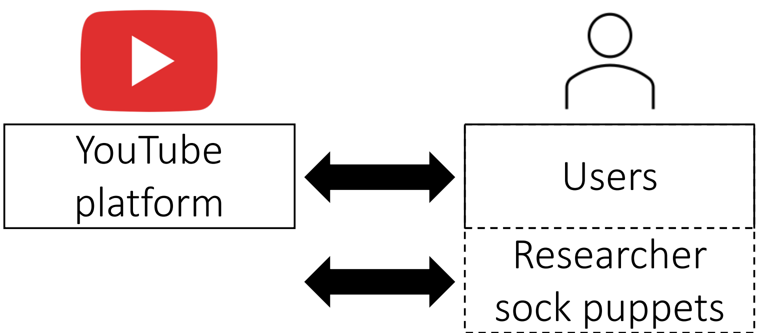


Figure 2.1. Sock puppet audit design for studying the YouTube algorithm (adapted from Sandvig et al., 2014).

The sock puppet experiment in this study used a 2 (race: Black vs. White) × 2 (SES: high vs. low) factorial design. Each sock puppet was trained to mimic one of four demographic groups: (a) high-SES White American, (b) high-SES Black American, (c) low-SES White American, and (d) low-SES Black American. The goal of the training process was to imitate the behaviors of real-world users from each of these four demographic groups. For the purpose of this study, I trained a total of $N = 840$ sock puppets, with 210 sock puppets being trained in each of the four demographic groups.

The final number of sock puppets used was much smaller than what was planned (20,000 sock puppets in each experimental group for a total of 80,000 sock puppets across all four groups). The initial plan was based on the assumption of the successful use of the Center for High Throughput Computing (CHTC) facilities at the University of Wisconsin-Madison to run the experiment because of the intense computing power and resources needed. However, despite multiple rounds of meetings and communication with the CHTC, the collaboration did not work out as planned due to technical difficulties in installing web browsers, Python, and other dependencies necessary for running the source code in the Docker environment on the centers' devices. Consequently, my plan had to be reconsidered since due to the much longer processing time than expected.

It should be noted that although the final number of sock puppets used in this study ($N = 840$) was smaller than originally planned, it still serves the purpose of this research well. To put that into perspective, prior research employing a similar sock puppet algorithm audit design has used from less than 10 (Hussein et al., 2020) to less than a hundred (Srba et al., 2023) to tens of thousands of sock puppets (Haroon et al., 2022).

The experiment was eventually run on the Google Cloud Platform's virtual machines (VM). All VMs were configured in an identical way, running from central U.S. and using 8 vCPU, 32 GB memory, 200 GB disk, and Ubuntu 18.04 LTS operating system. A total of 60 VMs were used to run the experiment in parallel. The complete training and testing session for a sock puppet took about 240 minutes; that is, one VM could run approximately 6 sock puppets per day. With 60 VMs running in parallel, it took about 3,360 minutes, or 2.3 days, to run the full experiment (i.e., all 840 sock puppets). The sock puppets were implemented using the Selenium library version 4.9.1 in Python. Other technical setup included using Python 3.8, Docker 20.10.24, Google Chrome browser version 114 with ChromeDriver version 114.0.5735.90, and the incognito mode of the Google Chrome browser to prevent browsing history or cookies being carried over across different sessions. All rules of using the YouTube Data Application Programming Interface (API; YouTube, 2022) were followed throughout the experiment.

Geographic location

Given that YouTube utilizes users' geographic location as an input feature in its recommendation algorithm (Covington et al., 2016), it is desirable to have all sock puppets connect to the Internet from the same geographic area to rule out the potential confounding influence of location. Hence, the location setting on Google Chrome was customized to be in Milwaukee, Wisconsin, U.S. for all sock puppets. The total population estimate for Milwaukee as of July 1, 2022 was 563,305, with a median household income of $45,318 and 24% of the population living in poverty (United States Census Bureau, 2022). About 85% of the population were high school graduate or higher, and 26% held a bachelor's degree or higher (United States Census Bureau, 2022). Nearly 89% of the households in Milwaukee had a computer and 80% of the households had access to broadband Internet subscription (United States Census Bureau, 2022).

The city of Milwaukee was primarily chosen because Whites (40%) and Blacks (39%) were the dominant racial categories in Milwaukee and comprised about equal shares of the total population in Milwaukee (United States Census Bureau, 2022). This means that, just based on geographic location, the chances of a sock puppet being recognized by the YouTube algorithm as either a White user or a Black user would likely be about equal and would be larger than the chance of the sock puppet being identified as a non-White and non-Black user.

To customize the sock puppets' geographic location, I first collected all 5-digit zip codes for the city of Milwaukee and their corresponding latitude and longitude pairs. Then, from this collection, a pair of latitude and longitude was randomly selected each time, with random noise between 2 miles to the north and 2 miles to the south (or ±0.03 degrees in latitude) added to the latitude and random noise between 155 feet to the west and 155 feet to the east (or ±0.0006 degrees in longitude) added to the longitude, to set the customized location of the Google Chrome browser for each sock puppet.

Training

In marketing research, psychographic data is often used in conjunction with demographics information to define the best way to target audiences. Psychographics focus on target audiences' interests, lifestyle, habits, attitudes, and behaviors (Meredith, 2021) and therefore provided useful training information for the puppets used in this study.

To conduct psychographic profiling on the four demographic groups of interest (i.e., high-SES White, high-SES Black, low-SES White, and low-SES Black Americans), I relied on the Simmons Insights database (MRI-Simmons, 2018), which provides access to the demographic and psychographic consumer data collected by MRI-Simmons's annual, nationally representative National Consumer Study (MRI-Simmons, n.d.). Since 1962, the National Consumer Study surveys over 25,000 American adults aged 18 or older every year to collect data on consumer buying habits and preferences, media usage, and demographic characteristics (MRI-Simmons, n.d.). The survey contains about 5,000 questions with 50,000 possible answers

(University of Wisconsin-Madison Libraries, n.d.). As alluded to above, the Simmons Insights database is often used to aid marketing strategy and advertising decisions by allowing researchers to generate customized crosstab data tables that can help determine the size and characteristics of targeted audiences and the media that should be used to reach a particular target group. I used the spring 2020 Simmons Connect Plus dataset that was fielded between April 22, 2019 and June 11, 2020, which was the most recent data available at the time of this research due to a 2-year delay in data availability. The total sample size for the spring 2020 survey is $N =$ 25,365, with 18,119 Whites and 3,140 Blacks, 2,695 not graduating high school and 8,659 with a bachelor's degree or higher, and 3,060 having a household income of less than $25,000 and 5,705 with a household income of $150,000 or more.

Empirical evidence indicates that Americans with distinct racial makeup and SES tend to engage in systematically different activities when online, such as by consuming different kinds of media products and searching for different topics online (e.g., Bacher-Hicks et al., 2021; Barthel et al., 2019; Leonhardt, 2014; Marshall & Naumann, 2018). Based on existing research and data, three broad categories of factors emerged as relevant for distinguishing the four demographic groups of interest, namely media use, lifestyles, and health.

With respect to media use, group differences exist in the consumption of digital news (national and local news) and entertainment media (e.g., movies, music). For instance, high-SES Americans are overall more likely to visit news websites than their low-SES counterparts (Barthel et al., 2019; Forman-Katz & Matsa, 2022). Further, at both the national and local (i.e., Milwaukee) levels, there are digital news media that are oriented toward and primarily consumed by a particular racial group (Pew Research Center, 2021); this is also true for movie viewership. In addition, White and Black Americans generally prefer different music genres, as Whites tend to prefer rock, country music, and pop music more than Blacks, whereas Blacks tend to prefer hip-hop, rap, and soul more than Whites (Marshall & Naumann, 2018). SES is also linked to preferences for different music genres (Snibbe & Markus, 2005). With respect to lifestyles, a decade of Google search data reveal that Americans living in richer, more educated counties are more likely to search for topics such as various foreign travel destinations, digital cameras, and new technology products than Americans living in poorer, less educated counties (Leonhardt, 2014). Americans with lower income also tend to have lower technology adoption (Vogels, 2021). Additionally, one's SES can also affect their online shopping behaviors (e.g., Huang et al., 2018), such as what stores they choose to shop from (McGuirt et al., 2022; Pechey & Monsivais, 2015). Finally, turning to health, decade-long Google search data show that Americans living in poorer, less educated places are more likely to search health problems (e.g., diabetes, blood pressure, weight-loss diets) online than Americans living in richer, more educated areas (Leonhardt, 2014). Indeed, systematic medical research indicates that lower SES is linked to increased risk of various health conditions, such as cardiovascular diseases (Schultz et al., 2018), type 2 diabetes (Agardh et al., 2011), and cognitive impairment and dementia (Wang et al., 2023), among others.

Given the above empirical evidence, I turned to the spring 2020 Simmons Connect Plus dataset to find relevant media consumption, lifestyles, and health factors distinguishing White versus Black Americans and low- versus high-SES Americans (Table 2.1). The factors in Table 2.1 were selected because the difference between White and Black or between high- and low-SES groups was the largest on these factors; in other words, these factors were the strongest indicators of racial or SES group membership. Specifically, for a group to be considered as having a characteristic, it needed to meet several criteria: (a) among members of the group (e.g., Blacks), the percentage of people who had the given characteristic (e.g., listening to neo soul) should be larger than the percentage of people who had the same characteristic among the comparison group (e.g., Whites); (b) among people who had the characteristic (e.g., listening to neo soul), the percentage of people from the target group (e.g., Blacks) should be larger than the percentage of people from the comparison group (e.g., Whites); and (c) compared to the total population, the target group was more likely to possess the given characteristic (as indicated by an "index" of over 100) and the comparison group was less likely to possess the characteristic (as indicated by an "index" of under 100), where "index" is a pre-calculated metric provided by the Simmons Insights database. Note that despite careful thought put into factor selection for the purpose of training sock puppets, these factors are by no means a comprehensive representation of all the ways sociodemographic audiences differ. Nonetheless, researchers could only realistically focus on a limited set of factors that elicit the largest group differences out of the virtually boundless possibilities.

Particularly, to identify digital news media that were owned by or oriented toward Black Americans, I turned to the Mapping Black Media project hosted by the Center for Community Media (CCM) at the City University of New York, which provides a directory of over 400 media outlets across the U.S. that primarily serve Black communities (Center for Community Media, 2020). From the directory, 56 digital media entries with a national audience and 1 digital media entry with a Milwaukee audience (carvdnstone.com) were identified. I then checked the 56 national digital media entries against survey questions on news website use from the spring 2020 Simmons Connect Plus dataset and excluded entries that Whites were more likely than Blacks to read and those whose readership was primarily Whites.

Table 2.1. Differentiating racial and SES groups by media consumption, lifestyles, and health factors.

	Race		SES	
	Whites	Blacks	High	Low
Media consumption				
National digital news	• yahoo.com	• thegrio.com		
Local digital news	• jsonline.com	• carvdnstone.com		
Movies	• Isn't It Romantic (2019) • Pet Sematary (2019)	• Little (2019) • Tyler Perry's A Madea Family Funeral (2019)		
Music	• 60s-70s pop classic rock (Beetles, Eagles) • Alternative rock (Green Day, Weezer) • Mainstream/pop country (Carrie Underwood, Taylor Swift) • Traditional country (Keith Urban, Martina McBride)	• Neo soul (Erykah Badu, Sade)	• Classical • Jazz	• Mexican (regional)
Lifestyles				
Foreign travel			• France • Italy • UK (England, Scotland, Wales) • India • Japan	
Still camera ownership			• Canon • Nikon • Sony	
Camcorder ownership			• Canon • Sony • JVC • GoPro	
Future tech purchase			• 4K ultra HD television set	

Online shopping sites	• Costco.com • Dickssportinggoods.com • Homedepot.com • Target.com • Google.com/shopping	• Fingerhut.com • Familydollar.com • Super-samples.com • Rainbowshops.com • Aarons.com
Health		
Personal ailment		• ADD/ADHD • Arthritis (osteoarthritis, rheumatoid, and any) • Diabetes type II • Heart attack/ stroke • Hiatal hernia
Caregiving		• Alzheimer’s disease

Sources: Data from spring 2020 Simmons Connect Plus (Simmons Insights), except for those in the “local digital news” row, which come from the Mapping Black Media project (https://www.journalism.cuny.edu/2020/11/mapping-black-media/).

The sock puppets were then trained to mimic users from each of the four sociodemographic groups of interest, by searching and watching content related to the factors identified in Table 2.1 on YouTube (see Table 2.2 for a summary of the specific training activities).

Table 2.2. Summary of training activities for the experimental conditions.

<table>
<tr><th></th><th>Whites</th><th>Blacks</th></tr>
<tr><td>High-SES</td><td>News:
YouTube watch a random set of 5 videos from @YahooFinance (each for 5 minutes max)
YouTube watch a random set of 5 videos from @jsonline (each for 5 minutes max)
Movies:
YouTube search Isn't It Romantic (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search Pet Sematary (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Music:
YouTube search (any one of) classic rock, alternative rock, country pop, and traditional country and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search (any one of) classical music and jazz and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Foreign travel destinations:
YouTube search (any one of) France, Italy, UK, India, and Japan and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Tech:
YouTube search (any one of) Canon, Nikon, Sony, and GoPro cameras and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search 4K ultra HD television set and watch a random set of 5 videos from top 20 results (each for 5 minutes max)</td><td>News:
YouTube watch a random set of 5 videos from @thegrio (each for 5 minutes max)
YouTube watch a random set of 5 videos from @carvdnstone (each for 5 minutes max)
Movies:
YouTube search Little (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search Tyler Perry's A Madea Family Funeral (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Music:
YouTube search neo soul and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search (any one of) classical music and jazz and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Foreign travel destinations:
YouTube search (any one of) France, Italy, UK, India, and Japan and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Tech:
YouTube search (any one of) Canon, Nikon, Sony, and GoPro cameras and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search 4K ultra HD television set and watch a random set of 5 videos from top 20 results (each for 5 minutes max)</td></tr>
<tr><td>Low-SES</td><td>Movies:
YouTube search Isn't It Romantic (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search Pet Sematary (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Music:</td><td>Movies:
YouTube search Little (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
YouTube search Tyler Perry's A Madea Family Funeral (2019) and watch a random set of 5 videos from top 20 results (each for 5 minutes max)
Music:</td></tr>
</table>

	• YouTube search (any one of) classic rock, alternative rock, country pop, and traditional country and watch a random set of 5 videos from top 20 results (each for 5 minutes max) • YouTube search Mexican (regional) and watch a random set of 5 videos from top 20 results (each for 5 minutes max) Health: • YouTube search (any four of) ADD/ADHD, arthritis, diabetes type II, heart attack, stroke, hiatal hernia, and Alzheimer's disease and watch a random set of 5 videos from top 20 results (each for 5 minutes max)	• YouTube search neo soul and watch a random set of 5 videos from top 20 results (each for 5 minutes max) • YouTube search Mexican (regional) and watch a random set of 5 videos from top 20 results (each for 5 minutes max) Health: • YouTube search (any four of) ADD/ADHD, arthritis, diabetes type II, heart attack, stroke, hiatal hernia, and Alzheimer's disease and watch a random set of 5 videos from top 20 results (each for 5 minutes max)

Note. Each cell contains $n = 210$ sock puppets.

Each sock puppet was set to watch each video for either its entirety or a maximum of 5 minutes, whichever smaller. The threshold of 5 minutes was chosen because empirical data indicated that the length of YouTube videos on average was about 11.7 minutes and that for a 7- to 14-minute-long YouTube video, users on average watched 50% of it (Narakeet, 2022; Wistia, 2023), meaning that the time an average user spent watching a typical YouTube video was around 5 minutes.

To develop search queries for training as well as testing activities, I adopted a systematic approach that was previously used in similar algorithm audit research (e.g., Hussein et al., 2020). Specifically, I collected relevant queries for each search activity from both Google Trends and YouTube's autocomplete suggestions to make sure that the queries were both popular among real-world users and relevant to the specific YouTube platform. I began by searching the seed queries (see Appendix A) in Google Trends as a "topic" for each search activity, to get the most popular related search queries over the past five years (April 2018-April 2023). Note that Google Trends allows researchers to search a query either as a "topic" or a "term." A "topic" covers a set of related queries that people use to search on a common theme (i.e., topic), whereas a "term" is a literate search query. This way, I was able to collect the most popular search queries that people used when searching in Google for information on a topic (e.g., the neo soul music genre). I also typed the seed queries into YouTube's search box to observe the autocomplete suggestions. YouTube's autocomplete feature suggested 14 popular queries for each seed query fed into the search box. The popular queries obtained from Google Trends and YouTube autocomplete suggestions made up the "sample related queries" column in Appendix A.

From this collection of sample related queries, I then selected the most popular and representative query for each search activity. Several rules were applied during this selection process: (a) queries that were potentially confusing or irrelevant to the search activity were excluded to ensure relevance and accuracy. For example, while "isn't it romantic elle fitzgerald" was a popular search query for the seed query "isn't it romantic," it referred to a song released in the 1950s by the jazz singer Elle Fitzgerald, instead of the 2019 movie. Therefore, this query was excluded from the set; (b) duplicate queries were removed and semantically similar queries were replaced with a single relevant query; (c) queries with a lower popularity index based on Google Trends were dropped; and (d) when everything else equal, the shorter query was selected because longer queries tended to be overly specific and lack generality or representativeness of the search topic at hand (Hussein et al., 2020). See Appendix A for the final queries used in each training and testing activity.

Finally, it is worth noting that researchers may also choose to open Google accounts and specify the age, gender, and name of those hypothetical users by setting up their user profiles in Google accounts (Hussein et al., 2020). Researchers may also reveal the artificial users' race by setting up a user profile picture. However, this approach may have limited utility for two reasons. First, prior research found that the specified user demographic information (e.g., age, gender) did not have a significant impact on the returned search results recommendations received by users who had brand new Google accounts without much past watching and searching histories built into the accounts (Hussein et al., 2020). In addition, research that implements sock puppets as Google accounts is limited in the number of sock puppets it can have because creating a Google account takes additional verification measures and cannot be automatically done on a large scale (Hussein et al., 2020).

Testing

During the testing phase, each trained sock puppet searched the three science issues of interest on YouTube, with the search order being randomized. The search queries used were: "genetic engineering human" for HGE, "ai" for AI, and "covid vaccine" for COVID-19 vaccines, respectively (see Appendix A). These search queries were selected using the same procedures described in the previous section. These queries were the most popular ones that Americans used when searching online for information related to the three science issues, according to Google Trends data from April 2018-April 2023. For example, the search query "ai" accounted for about 100% of web searches on the general topic of artificial intelligence. In other words, I selected the search queries based on what people actually used when searching the three science topics online, and therefore the sock puppets would likely receive video recommendations that people would actually encounter in real-life scenarios. The goal here was to maximize the ecological validity of the experiment, rather than search an exhaustive list of queries that would return nearly all YouTube videos related to HGE, AI, or COVID-19 vaccines.

Each sock puppet recorded the top 20 video recommendations returned by YouTube for each of the three science issue searches. For each video recommendation, I collected video-level information, including video title, video link, video upload time, video length, video view count, video likes, video comment count, video thumbnail link, video description, and video transcript; and channel-level information, including channel name, channel link, channel subscriber count, channel view count, channel video count, and channel view-subscriber ratio.

Analysis

To understand the science video recommendations, I focused on the semantic content of video transcripts. Future research will examine the semantic content of video titles and descriptions as well as the visual content of video thumbnails and frames.

For quality checking, I manually inspected the collected video transcripts and identified instances where the transcript was unavailable. For those videos, I went to their YouTube video page to verify whether the video was transcribable. Nearly all the videos whose transcript was unavailable were in a language other than English, which were excluded from later analysis. Videos whose transcript was not readily available due to YouTube setting were transcribed and included in subsequent analysis.

To analyze the topics covered by the video recommendations, I relied on structural topic modeling (STM; Roberts et al., 2019), a computational content analytical method for discovering topics—or the latent thematic structure of a collection of textual documents, as well as the relationship of those topics to document metadata (Roberts et al., 2019). STM can process large corpora of text efficiently compared to human coders and can aid the identification of potentially useful but empirically understudied or unknown organizations of text (Grimmer & Stewart, 2013). Moreover, STM combines human researchers' expertise with computational efficiency, especially during the phase when researchers need to interpret the meaning of the topics and validate the usefulness of those topics to the study at hand (Roberts et al., 2019). In science communication as well as other fields of social sciences, STM has been used to analyze a range of texts such as open-ended survey responses (Chen et al., 2020) and news media coverage (Ophir, 2018; Wirz et al., 2022).

I constructed structural topic models separately for the three science issues searched at the testing stage—HGE, AI, and COVID-19 vaccines. First, I constructed structural topic models on the transcripts of all video recommendations returned for a science issue across experimental groups. After that, within each science issue, I constructed structural topic models separately for each of the four experimental groups on videos that were recommended only to that group but not the others. This was to see whether videos recommended to a particular sociodemographic audience discuss science differently from videos recommended to another sociodemographic audience, and if so, in what ways. Future research will incorporate covariates (e.g., experimental conditions, video upload time) in STM analysis.

Before constructing the topic models, I first cleaned the data by applying the "textProcessor" function of the STM R package (version 1.3.6), which performs a series of text preprocessing steps including stemming words (reducing words to their root form), dropping punctuations, and removing stop words (Roberts et al., 2019). The data were then converted to a matrix format for STM analysis. To determine the number of topics, I used the "searchK" function and selected models with a relatively high held-out likelihood, high semantic coherence, and a small residual (see Appendix B for model comparisons). Following Roberts et al's (2020) guidelines, I compared models with 3 to 40 topics for the corpus of unique videos for each of the three science issues. A model of 4, 8, and 4 topics was selected for the HGE, AI, and COVID-19 vaccine unique video corpus, respectively, based on the size of the corpus (for example, there were more AI videos than HGE and COVID-19 vaccine videos).

Next, within each of the three science issues, I compared videos received only by one of the four experimental groups but not the others. For HGE videos, a total of 2 videos were received only by one of the four experimental groups (1 video received only by the high-SES White group, and 1 video received only by the high-SES Black group). For AI videos, a total of 42 videos were recommended only to one of the four groups (11 videos received only by high-SES Whites, 15 videos received only by high-SES Blacks, 8 videos received only by low-SES Whites, and 8 videos received only by low-SES Blacks). For COVID-19 vaccine videos, a total of 13 videos were recommended only to one of the four groups (3 videos received only by high-SES Whites, 6 videos received only by high-SES Blacks, 3 videos received only by low-SES Whites, and 1 video received only by low-SES Blacks). Because the sizes of the corpora were very small (all under 20 video transcripts), all models had 3 topics, following Roberts et al's (2020) guidelines.

I labeled each topic by examining frequent and exclusive words to each topic and reading example video transcripts of each topic. Finally, for each video transcript, STM provided theta values representing the probability that the transcript belonged to a certain topic (and all theta values for a single video added up to 1). Based on a video's highest theta value, I then assigned the according topic (i.e., most probable topic) to each video and counted the number of times each video appeared in the full corpus (including duplicate videos) to estimate the aggregate topic prevalence across all videos (see, e.g., Wirz et al., 2022).

Results

Diversity of science video recommendations

A full corpus of *N* = 50,400 videos (840 sock puppets, each collecting 20 videos for each of the three science issues) spanning 216 unique channels were collected at the end of the testing phase. Of these 50,400 videos, 268 are unique videos. As Table 2.3 shows, sock puppets representing high-SES Americans, especially high-SES Whites, encountered a greater diversity of YouTube videos (213 unique videos) and channels (173 unique channels) when searching for science issues, whereas sock puppets representing low-SES Americans, especially low-SES Blacks, encountered a less diverse collection of YouTube videos (187 unique videos) and channels (154 unique channels). In addition, among the three science issues examined, videos related to AI were the most diverse (155 unique videos) and came from the greatest variety of channels (128 unique channels), whereas videos related to HGE were the least diverse (40 unique videos) and came from a limited collection of YouTube channels (38 unique channels), which might have to do with the lack of HGE-related on YouTube in the first place.

Out of the 268 unique videos received by the trained sock puppets as they searched for the three science issues on YouTube, about 53% (or 143 videos) were recommended to all four experimental groups. Further, 80% (or 32 videos) of the 40 unique HGE video recommendations, about 43% (or 66 videos) of the 155 unique AI video recommendations, and about 62% (or 45 videos) of the 73 unique COVID-19 vaccine video recommendations were received by all four experimental groups. As Figures 2.2 through 2.4 show, the number of times a video showed up in the search results recommendations was largely consistent for sock puppets that were assigned different racial and SES categories. Take HGE for example, the first 15 unique HGE videos were recommended at similar frequencies to sock puppets representing different racial and SES categories (see Figure 2.2). Similar patterns were also detected for the issue of AI (Figure 2.3) and COVID-19 vaccines (Figure 2.4). Appendix C lists the video title corresponding to each of the unique video IDs shown in Figures 2.2 through 2.4.

Table 2.3. Summary of video search results recommendations collected during the testing phase, by experimental group and science issue.

	Total videos recommended	Total unique videos recommended	Total unique channels recommended
Group 1: High-SES Whites	*12,600*	*213*	*173*
HGE	4,200	37	35
AI	4,200	116	95
COVID-19 vaccines	4,200	60	43
Group 2: High-SES Blacks	*12,600*	*204*	*171*
HGE	4,200	37	35
AI	4,200	107	91
COVID-19 vaccines	4,200	60	45
Group 3: Low-SES Whites	*12,600*	*191*	*160*
HGE	4,200	36	34
AI	4,200	98	85
COVID-19 vaccines	4,200	57	41
Group 4: Low-SES Blacks	*12,600*	*187*	*154*
HGE	4,200	37	35
AI	4,200	96	80
COVID-19 vaccines	4,200	54	39
Total	***50,400***	***268***	***216***
HGE	***16,800***	***40***	***38***
AI	***16,800***	***155***	***128***
COVID-19 vaccines	***16,800***	***73***	***50***

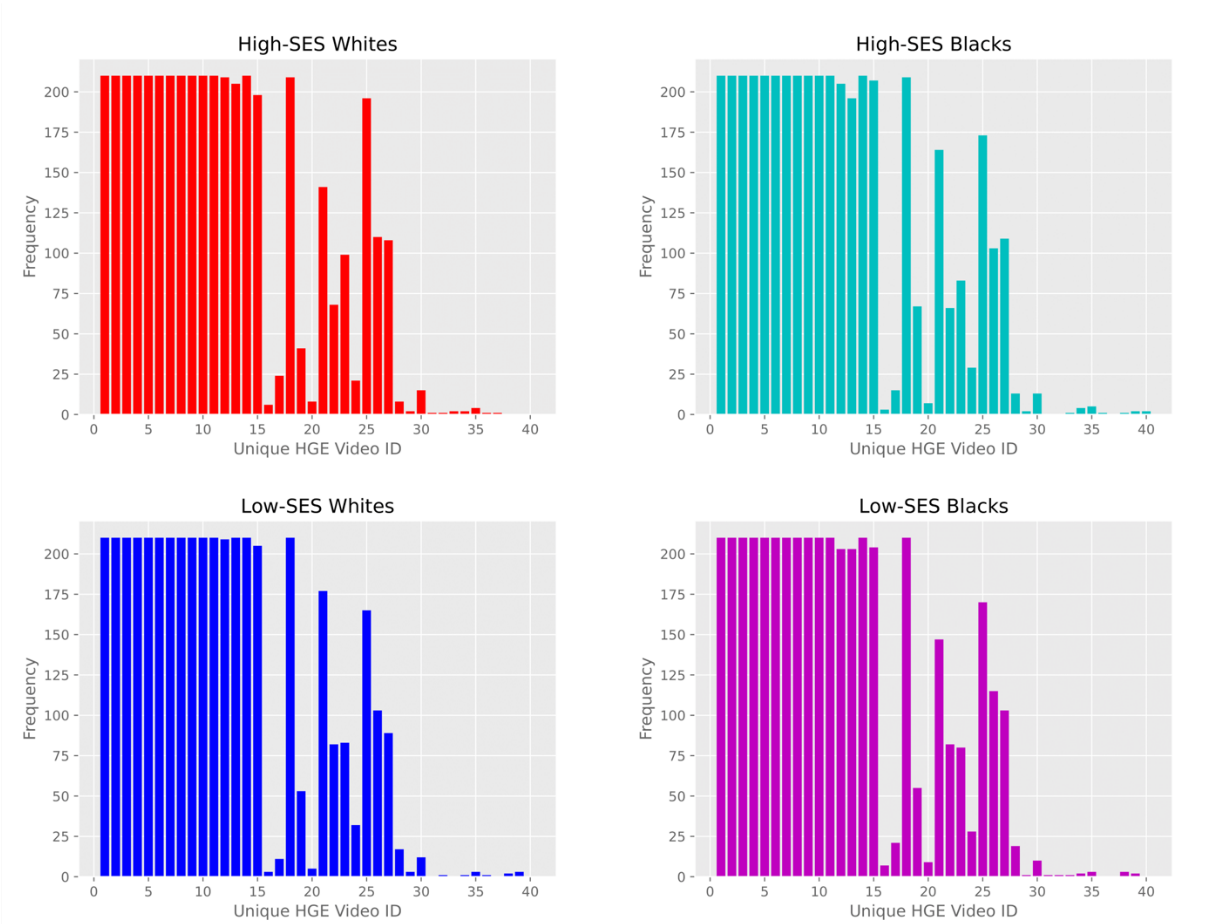


Figure 2.2. Frequency of unique HGE videos appearing in search results recommendations across experimental groups.

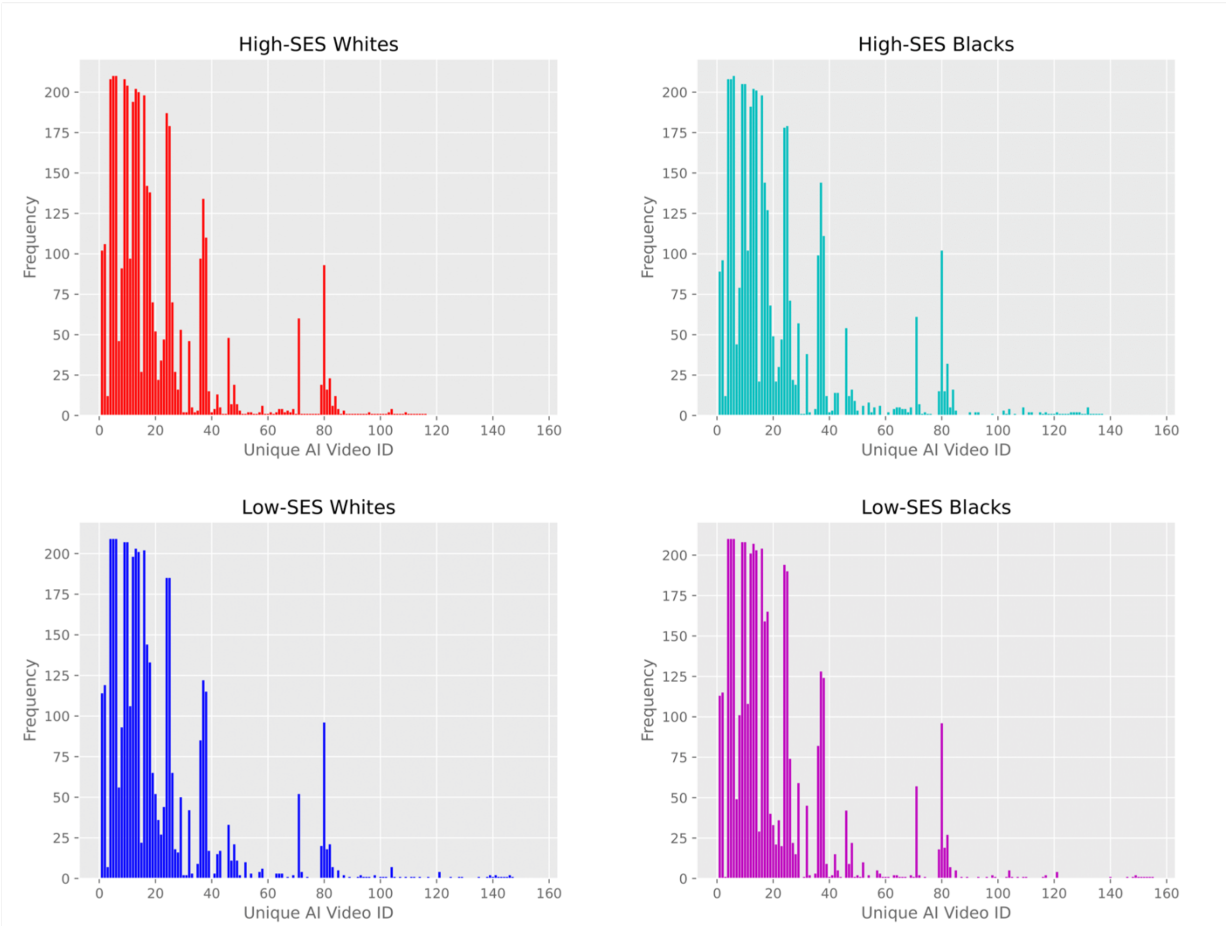


Figure 2.3. Frequency of unique AI videos appearing in search results recommendations across experimental groups.

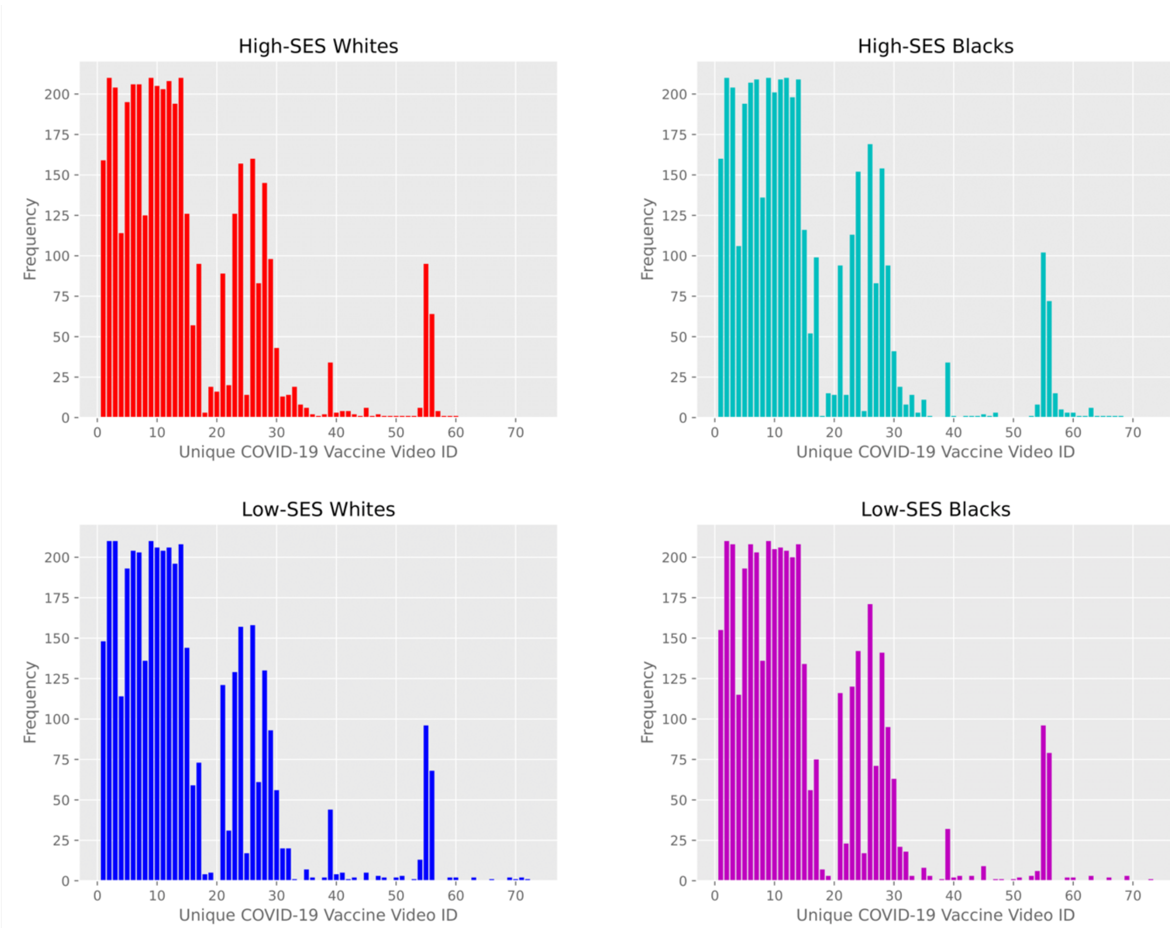


Figure 2.4. Frequency of unique COVID-19 vaccine videos appearing in search results recommendations across experimental groups.

Semantic content of science video recommendations

STM was applied to all videos whose transcript was available in English. A total of 50 unique videos were excluded from the unique video corpus due to not meeting this criterion, resulting in a final corpus of $N = 218$ unique videos (38 unique HGE videos, 115 unique AI videos, and 65 unique COVID-19 vaccine videos).

Before analyzing the difference between groups to explore my RQ1, I conducted a review of the content of the corpus of videos for HGE, AI and COVID-19 vaccines.

Four topics emerged from the HGE video corpus, with their total expected prevalence calculated (see Table 2.4). The most prevalent topic underpinning HGE videos focused on the *clinical applications of HGE* (76.16%), such as curing diseases and genetic disorders, therapeutic uses, and gene-editing babies, as advances in gene-editing technology (e.g., CRISPR) have made precise, targeted changes possible. In addition to HGE's clinical uses, searching HGE on YouTube also returned videos that discussed *gene editing non-human organisms* (20.73%), which included, for example, applying genetic engineering to modifying apples so that they become more resistant to browning, genetically modifying dairy cows and cow milk by introducing human genetic coding, creating human insulin by transferring the human insulin gene to microorganisms that are often cultured in fermenters, and genetically modifying salmon. A much less prevalent topic was *humanity in an era of HGE* (2.9%), which involved discussions of the link between homo sapiens, Neanderthals, and other primates, the relation of humanity and life on earth to extraterrestrial beings (aliens), and the potential risk of creating unequal social classes and a "slave society" with gene-editing technology. Finally, *other* (0.09%) topics of video recommendations related to HGE included creative technologies, collaborative research, and historical artifacts.

Table 2.4. Video search results recommendations: Overview of topics, keywords, and expected prevalence.

Issue	Topics	Exp. Prev.	Keywords
HGE	Clinical use of HGE	76.16%	edit, babi, target, therapi, cure, disord, clinic
	Gene editing non-human organisms	20.73%	appl, nucleus, cow, insulin, milk, salmon, ferment
	Humanity in an era of HGE	2.90%	primat, alien, homo, neanderth, slave, mind, earth
	Other	0.09%	creativ, model, guess, kind, carv, like, collabor
AI	AI players and products	33.38%	microsoft, report, chatbot, bing, compani, search, openai
	Risks of AI	24.54%	moloch, ration, machin, gpt-, max, relev, regul
	AI and music	16.17%	okay, song, bro, drake, yall, music, locat
	AI and art	10.96%	one, week, imag, tool, cool, digit, photo
	Showcasing AI use	9.85%	laugh, sakura, killer, swing, trade, gamer, shot
	Other	4.00%	elect, bard, cell, evolv, agi, smarter, align,
COVID-19 vaccines	Public health updates	34.40%	shot, world, booster, healthi, covert, omicron, sever
	Vaccine misinformation and controversies	33.40%	florida, mandat, money, mortal, conspiraci, claim, report
	Science of COVID-19 vaccines	17.65%	protein, spike, brain, bodi, cell, scienc, antibodi
	Other	5.45%	slide, next, provid, monoval, dose, addit, immunocompromis

Turning to video recommendations related to AI, the leading topic focused on *AI players and products* (33.38%), such as technology companies (e.g., Microsoft, OpenAI), AI-supported chatbots, and AI-supported search engines (see Table 2.4). Following that, *risks of AI* (24.54%), such as technology companies racing to develop more sophisticated (but not safer) AI due to Molochian commercial pressures, AI being intelligent but not truth-driven rational machines, AI replacing what people are able to do, and AI developments far outpacing regulatory efforts, were also a key feature of AI video recommendations. The next topic focused on *AI and music* (16.17%), which included discussions of AI automatically writing songs, AI making music that sound like real musicians, and virtual musical artist (e.g., Kizuna AI). A fourth topic was *AI and art* (10.96%), which involved discussing, for example, how AI has changed photography, AI-generated images and visual art, and AI tools for artists (e.g., AI video editing tools, AI imaging applications). Next, *showcasing AI use* (9.85%) involved the YouTuber showcasing specific AI-generated work or the YouTuber's own interactions with AI for particular use purposes, such as interacting with ChatGPT to write a romance story, interacting with ChatGPT to work out a video game strategy, interacting with simulated AI characters in a video game, showcasing an AI-generated golf video lesson, showcasing how to use ChatGPT to do cryptocurrency trading, and showcasing how to create a short film entirely using AI. Finally, *other* (4%) topics of AI video recommendations included AI's capabilities to learn from massive data, to predict election outcomes, to generate creative and humane-sounding speech, and to aid biological research, artificial general intelligence (AGI), and challenges associated with aligning AI with human values and objectives, among others.

Finally, with respect to the COVID-19 vaccine video corpus, four topics emerged including *public health updates* (34.4%), *vaccine misinformation and controversies* (33.4%), *the* *science of COVID-19 vaccines* (17.65%), and *other* (5.45%; see Table 2.4). The topic *public health updates* included discussions of latest public health guidelines on getting COVID-19 vaccines and booster shots, vaccination guidelines by public health authorities such as the World Health Organization, and vaccination against the omicron variant as well as severe infection. The next topic focused on *vaccine misinformation and controversies*, which included both videos attempting to debunk misinformation related to the COVID-19 vaccines and videos casting doubt on the vaccines, such as discussions of Florida health officials altering a study of COVID-19 vaccine safety to make the vaccines appear less safe, claims of a link between COVID-19 vaccine uptake and heightened mortality rate, and stories about how one became the target of a conspiracy theory. A third topic was the *science of COVID-19 vaccines*, which included, for example, how mRNA vaccines work, the spike proteins, how human cells function, how scientists developed the COVID-19 vaccines, and how the COVID-19 vaccines help the body create antibodies. Finally, *other* topics of video recommendations related to COVID-19 vaccines included effectiveness of monovalent COVID-19 vaccines, healthcare providers administrating COVID-19 vaccines, and immunocompromised populations.

Science video content by audience groups

STM was then applied to videos that were only recommended to one of the four experimental groups, to garner a further understanding of the topics that were uniquely encountered by different socioeconomic audiences.

Table 2.5. Topics and keywords of video search results recommendations by science issue and experimental group.

Issue	Group	Topics	Keywords
HGE	H, W	Benefits of HGE	genet, can, gene, edit, use, human, crispr
AI	H, W	Showcasing AI use	readi, wait, anywher, care, chat, swing, gamer
		AI for mental health	famili, friend, heart, whos, support, specif, anyway
	H, B	Showcasing AI use	fine, drake, covers, song, write, feel, fun
		AI for space exploration	comput, intellig, object, everi, space, minut, univers
		Problems with AI	gpt, eventu, open, generat, sourc, chat, safeti
	L, W	Showcasing AI use	content, generat, photo, imag, realist, reason, bro
	L, B	Risks of AI	predict, parti, data, super, algorithm, white, public
COVID-19 vaccines	H, W	Tips on helping children get vaccinated	can, honest, get, percent, got, today, children
		Vaccine misinformation and controversies	caus, heart, like, happen, death, realli, talk
		Explanation of data and statistics	peopl, time, look, now, data, give, that
	H, B	Public health updates	updat, covid-, heart, variant, blood, can, protect,
		Vaccine misinformation and controversies	pfizer, will, data, thank, pleas, question, vaccin
	L, W	Public health updates	year, know, virus, covid, booster, infect, immun
	L, B	Public health updates	vaccin, get, everi, kid, new, want, well

Note. "H, W" denotes sock puppets representing high-SES Whites, whereas "H, B," "L, W," and "L, B" denote sock puppets representing high-SES Blacks, low-SES Whites, and low-SES Blacks, respectively.

In terms of HGE-related video recommendations, the high-SES White group received additional content about the benefits of HGE that was not received by the other groups. Specifically:

Sock puppets representing **high-SES White Americans** were recommended video content that focused on the *benefits of HGE*, such as treating human diseases, creating a healthier and more sustainable food supply, and facilitating reforestation efforts. The one video received only by the high-SES Black group was excluded from STM analysis as its transcript was not available in English.

With respect to AI-related video recommendations, all groups except for the low-SES Black group received videos showcasing AI use, although the specific use cases might differ across groups. The two Black groups (high- and low-SES Blacks) also encountered videos discussing the problems or potential risks of AI. Specifically:

Sock puppets representing **high-SES Whites** encountered the topics of *showcasing AI use* and *AI for mental health*. Specific to the high-SES Whites group, *showcasing AI use* involved showcasing an AI-generated golf video lesson, interacting with simulated AI characters in a video game, and showcasing AI-generated commercials; *AI for mental health* involved talking about mental health support communities and using AI tools (e.g., Midjourney application) to visualize mental illness experiences (e.g., schizophrenia).

Meanwhile, sock puppets representing **high-SES Blacks** also encountered *showcasing AI use*, but for somewhat different use cases, such as generating rap music (songs that sounded like they were created by Drake, a Canadian Black artist), creating artificial videos showing basketball movement, as well as interacting with AI characters in video games. High-SES Black sock puppets also encountered the topic of *AI for space exploration*, which involved discussing how AI can be used to help humans explore the universe and the many objects in the universe, given AI's computing power and intelligence. In addition, high-SES Blacks also received video recommendations that focused on *problems with AI*, such as how AI chatbots such as ChatGPT generate and spread misinformation, and how social media AI features (e.g., Snapchat filter) do not work well for dark-skinned individuals.

Sock puppets representing **low-SES White** Americans received video recommendations that focused on showcasing AI use, for example, inviting viewers to tell AI-generated images from real images, showcasing AI-generated imagery of the YouTuber themselves, showcasing AI-generated short music videos, and showing AI-generated audiovisual materials of politicians (e.g., AI-generated negative campaign ad on Joe Biden, AI imitating Donald Trump, Barack Obama, and Joe Biden talking to each other playing Minecraft).

Sock puppets representing **low-SES Black Americans** encountered videos discussing the *risks of AI*, specifically, threats to user privacy as AI learns from massive social media user data, risk to democratic societies due to misuse of predictive AI by governments and politicians, and dangers of AI.

Turning to <u>video recommendations related to the COVID-19 vaccines</u>, public health updates were a consistent topic across most groups. The two High-SES groups received videos discussing COVID-19 vaccine misinformation and controversies, although the high-SES White group encountered misinformation-debunking videos whereas the high-SES Black group encountered misinformation-supporting videos. Specifically:

Sock puppets representing **high-SES White** Americans encountered videos talking about *tips for helping children get vaccinated*, such as parents being honest with their children about what is going to happen as they get vaccinated. They also received videos debunking *vaccine misinformation and controversies*, such as claims that COVID-19 vaccines cause heart failure and sudden death. In addition, high-SES White sock puppets encountered videos explaining COVID-19 vaccine data and statistics, for example, by discussing and giving more context to data published by official sources.

Sock puppets representing **high-SES Black** Americans received videos that focused on *public health updates*, such as updates on COVID-19 vaccine policy, safety, effectiveness, and programs, new variants of the novel coronavirus, the connection between COVID-19 and heart diseases, guidelines for holiday gatherings during the pandemic, as well as public service announcements (PSA) encouraging people to get an updated vaccine or booster; another topic they encountered was *vaccine misinformation and controversies*, which included discussions of Pfizer, the biopharmaceutical company behind the Pfizer-BioNTech COVID-19 Vaccines, rolling out the vaccines before having data on the vaccines' effects on stopping the transmission of the virus.

Sock puppets representing **low-SES White** Americans encountered the topic *public health updates*, specifically involving health experts (e.g., Dr. Anthony Fauci) discussing where the U.S. is on COVID-19 infections and deaths, lessons learned about COVID-19 and the related vaccines over the years, the effectiveness of vaccine boosters, and so forth.

Finally, the **low-SES Black** sock puppets were also recommended video content about *public health updates*, specifically announcing the kids' vaccination program in the U.S.

Discussion

The goal of this exploratory study was to assess to what extent algorithmically driven social media platforms, such as YouTube, make biased personalized content recommendations based on users' racial and socioeconomic profiles as users search for science information online. To this end, I used a novel computational method, the sock puppet experiment, to audit the recommendation algorithm of YouTube. Four groups of automated web browser instances (i.e., sock puppets) were trained to represent White and Black Americans with different SES by engaging with online activities and information characteristic of the racial and socioeconomic groups of interest. The trained sock puppets then searched for three wicked science issues on YouTube, including HGE, AI, and COVID-19 vaccines, and collected video recommendations returned by the YouTube algorithm. The collected video recommendations related to each of the three science issues were then analyzed for their diversity of sources and semantic content both on aggregate and across experimental groups.

Before elaborating on the research findings and their implications, it is necessary to first discuss some limitations of the current work. Specifically, the current experimental design is unavoidably a simplified representation of the social phenomena that it is interested in capturing, given all its complexities. For example, while this study focused on White and Black as two of the basic racial categories identified in the U.S. (United States Census Bureau, 2022), race is in effect a multifaceted social construct that involves not just a social definition (e.g., national origin, sociocultural group membership) but also biological, anthropological, as well as genetic definitions (United States Census Bureau, 2022). In addition, people can identify with multiple race categories at the same time. Race also intersects with other social identities (e.g., SES, age, and gender) and psychological and social influences in shaping individuals' preferences and behaviors. Similarly, SES is also a multidimensional construct and influences decisions, actions, and behaviors in synergy with other individual- and societal-level factors. In addition, despite the fact that individuals are embedded in their homophilic social networks comprising peers similar in race as well as other sociodemographic, behavioral, and intrapersonal characteristics (McPherson et al., 2001), no such social network data were available to train the sock puppets to better represent the racial categories. Together, racial and SES groups can express innumerable behavioral differences online and offline, and researchers are forced to focus on a limited set of factors that elicit the largest group differences when manipulating the racial and SES categories of automated agents in experimental research. Finally, the YouTube algorithm is also not fixed forever; instead, it is constantly evolving, just as any other digital media platforms do.

Even though researchers cannot possibly overcome all the limiting factors discussed above, research that investigates the algorithmic information curation processes and their impacts on society still needs to take place. On the one hand, AI algorithms power many media platforms, services, and tools that we use, drive the information that we consume, and shape the ways in which we communicate with each other (Wagner et al., 2021). On the other hand, how the algorithms work remains a black box to people who use and people who are affected by those platforms. Research that measures and evaluates the opaque yet pervasive algorithmic influences, albeit imperfect (due to many inherent social and technical challenges) and doomed to obsoletion (due to digital platforms constantly evolving), is not only valuable but also necessary for societies to function in a healthy and responsible way. Further, with the knowledge gained from this type of research, which facilitates our understanding of the processes and consequences of algorithmic information curation, societies are better prepared to respond to the challenges at the collective, policy level as well as at the individual level by helping citizens take informed actions, even if directly changing the algorithms or the social media platforms powered by the algorithms is infeasible.

With these considerations in mind, this study provides important information concerning how social media algorithms, such as the YouTube algorithm, recommend content to diverse racial and SES audiences when they search for science information online, as well as in what ways the information received by the different sociodemographic audiences might differ.

With respect to the diversity of content recommendations, audience members who have higher SES are more likely to receive a greater diversity of content recommendations from YouTube. As the current study finds, sock puppets representing high-SES Americans, especially high-SES White Americans, encountered a wider range of YouTube videos (213 unique videos) and channels (173 unique channels) when searching for the three science issues, whereas sock puppets representing low-SES Americans, especially low-SES Blacks, encountered a less diverse collection of YouTube videos (187 unique videos) and channels (154 unique channels). Then, the interesting question is, why are high-SES Americans getting a wider range of science information? While part of it may have to do with individual choices or voluntary information use habits, the current research shows that the algorithms of online media platforms may also play a role in exposing higher-SES audiences, as opposed to lower-SES audiences, to a wider range of science information sources.

In addition, the specific science issue context also appears to influence the diversity of content recommendations. Videos related to searches on the issue of AI were the most diverse (155 unique videos) and came from the most diverse channels (128 unique channels), whereas videos related to searches on the issues of COVID-19 vaccines and especially HGE were less diverse (COVID-19 vaccines: 73 unique videos; HGE: 40 unique videos) and came from a limited collection of channels (COVID-19 vaccines: 50 unique channels; HGE: 38 unique channels). However, the level of diversity of content recommendations for a science issue may be an artifact of the amount of available issue-relevant content on the YouTube platform in the first place.

Overall, a small majority (53%) of videos recommended by YouTube to the sock puppets as they searched for the three science issues were identical, regardless of the sock puppets' racial and SES categories. Again, issue contexts play a role here: for searches on AI, less than half (43%) of the video recommendations were identical across racial and SES audiences, whereas for searches on COVID-19 vaccines and HGE, an overwhelming majority (COVID-19 vaccines: 62%; HGE: 80%) of the video recommendations was identical across racial and SES audience groups. It therefore appears that when sociodemographically diverse audiences search for science issues on YouTube, they will be more likely to get different sets of information if the science issues are heavily discussed by a wide range of sources on YouTube than if the issues are only discussed by a limited set of sources on the platform.

Further, when searching for HGE on YouTube, users are likely to receive video recommendations that discuss the clinical applications of HGE, gene-editing non-human organisms, as well as humanity in an era of HGE, among other topics. High-SES White American audiences are particularly more likely to receive videos that focused on the benefits of HGE technology than low-SES or Black American audiences. When searching for AI, users are likely to encounter video recommendations that talk about AI players and products, risks of AI, AI related to music and art, as well as videos that present specific AI use cases or AI-generated work, among other topics. In addition, different racial and SES audiences may encounter different topics or similar topics but with different examples or emphasis. For instance, high-SES White audiences may encounter video recommendations that discuss the use of AI tools for mental health issues, whereas high-SES Black audiences may receive videos that talk about how the social media AI features (e.g., filters) do not work well for dark-skinned individuals. Finally, when searching for COVID-19 vaccines, users are likely to receive search results that discuss public health updates regarding the vaccines, vaccine-related misinformation and controversies (either debunking or supporting misinformation), the science of COVID-19 vaccines, among other topics. Audiences from different racial and SES groups may also encounter different topics or similar topics with different examples or emphasis in their search results recommendations.

More importantly, the findings suggest that social media algorithms, such as the YouTube algorithm, can indeed expose sociodemographically diverse audiences to different subsets of information even when people are actively searching for the same science issue, although the degree of information tailoring may depend on the search topic itself (e.g., how heavily the science issue is discussed by different sources on YouTube). In addition, higher-SES audiences, especially higher-SES White American audiences, are likely to receive a wider range of video and channel recommendations when searching for science issues than lower-SES audiences. Although low-SES audiences already tend to be informationally disadvantaged, the algorithmically infused media environment may expose them to a narrower range of science information than it does to their higher-SES, informationally resourceful peers, thus exacerbating disparities in science information and understanding.

While disparities in scientific understanding between high- and low-SES social segments may be attributed to differences in learning ability or to the human tendency to select attitude-confirming information (Tichenor et al., 1970), findings from this study suggest that such disparities may also be caused by algorithms that target human preferences and expose population segments to different subsets of information without their own free will. With the advent of ChatGPT and other large, multimodal language models that directly answer a user's searches and questions by instantly pulling and synthesizing information from numerous sources online, we should expect to see increasing power of AI algorithms in determining the information we see and meanwhile diminishing autonomy of the information consumer. The search results returned by search engines that incorporate AI language models will be more highly edited than ever before and, even if provided additional links, users are unlikely to check those links out because humans are cognitive misers who collect only minimum information to satisfy a decision-making need (Fiske & Taylor, 2008). Moreover, the AI language models may also fabricate utterly false search results or synthesize information in a biased way, requiring users to get into the habit of fact-checking (Jiang, 2023). As emerging AI technologies disrupt how we communicate and consume (science) information, it is necessary for future research to examine these platforms and tools so that we can pre-empt or counter structural barriers in our information environments that prevent equitable distribution of scientific information in society.

Chapter 3

Connecting social media use with education- and race-based gaps in factual and perceived knowledge across wicked science issues (Study B)

As discussed in chapter 1, disparities in science knowledge exist along socioeconomic, racial, as well as other demographic lines in the U.S. Such disparities may be understood through the lens of digital divides (Howell & Brossard, 2021) and the knowledge gap hypothesis (Tichenor et al., 1970), both of which point to the role of media and information in the formation of knowledge gaps between social segments. In today's world, where digital media are quickly outdating traditional media infrastructures as a primary source of science information for many people (Brossard, 2013; Scheufele, 2013), researchers have begun to examine how the use of the Internet (Cacciatore et al., 2014; Lee, 2009; Shim, 2008), science blogs (Su et al., 2014), online newspapers (Chang et al., 2018; De Silva-Schmidt et al., 2022), and social media (Gerosa et al., 2021) might affect SES-based gaps in science knowledge. Building on this emerging line of research, in this study I examine both factual knowledge and perceived knowledge about wicked science issues across sociodemographic segments to explore how individuals' social media use as well as science issue contexts might shape race- and education-based disparities in scientific understanding.

Outcome of interest: Science knowledge

As discussed in chapter 1, science knowledge—especially knowledge that cuts across multiple dimensions related to the science, application, history, and policy of given science issues—matters because it provides the basis for one's ability to engage effectively with science-related issues. Interestingly, how science knowledge is measured could affect to what extent knowledge gaps are identified. In their original thinking on the knowledge gap hypothesis, Tichenor et al. (1970) did not formally define knowledge and used both true/false factual statements and acceptance of stated beliefs to measure knowledge. Ensuing research has shown that factual knowledge about a novel science issue such as nanotechnology and perceived familiarity with nanotechnology are only slightly correlated with each other and are predicted differently by media use and cognitive processing variables; specifically, increased attention to Internet science information had a stronger association with perceived familiarity with nanotechnology than with factual nanotechnology knowledge (Ladwig et al., 2012). Further, SES- and education-based knowledge gaps are often observed and more pronounced when close-ended, factual-type knowledge about an issue is measured, as opposed to when open-ended, belief-type measures are used (Chang et al., 2018; Hwang & Jeong, 2009, 2010; Su et al., 2014).

As discussed in chapter 1, factual science knowledge is part of "civic science literacy," or "understanding of the many elements that shape the production of scientific knowledge" (Howell & Brossard, 2021, p. 2), and is hence a crucial basis for effective public engagement with science in society. However, how knowledgeable about science people think themselves are could also matter for important outcomes of democratic decision-making about science, such as public support for science issues (Akin et al., 2020). In addition, scholars have argued that facts are less definitive than often assumed and are subject to group and cognitive biases as well as contextual influences (Johnson, 1993). Which facts are relevant also depends on the specific audiences and situation of concern (Johnson, 1993). Therefore, some researchers have focused on belief-based knowledge measures instead of factual knowledge measures.

In this study, I examine both factual and perceived knowledge of wicked science issues. Moreover, knowledge gaps are identified not in the sense of overtime change (Tichenor et al., 1970), but with interactions between individuals' education level and (social) media use, wherein the strength of the positive relationship between issue knowledge and education is expected to vary depending on levels of media use in the presence of a knowledge gap, with the relationship being stronger at high levels of media use and weaker at low levels of media use (Eveland & Scheufele, 2000; Kwak, 1999; McLeod et al., 1979).

Social media use and science knowledge gaps

As I discussed in Chapter 1, although traditional mass print and broadcast media used to play a central role in disseminating news about scientific breakthroughs and bridging the science-public divides, they are giving way to online media as lay Americans are now turning to the Internet for information about science (Brossard, 2013). In particular, social media have become a prominent general news source (Newman et al., 2022), especially for younger generations (Newman et al., 2022). Most social media users in the U.S. encounter science-related information on these platforms (Funk et al., 2017b). (For a detailed discussion and classification of social media, see chapter 1 and Figure 1.5 on p. 17).

As discussed in chapter 1, research seems to suggest an overall negative relationship between social media use and factual science knowledge (Chang et al., 2018; Gerosa et al., 2021), with a few exceptions (e.g., Li & Cho, 2021; Su et al., 2014). Further, although there is some evidence indicating that increased attention to science information on social media overall was associated with a narrower gap in perceived science knowledge between high and low education groups (Chang et al., 2018), it is less clear whether social media use would similarly narrow factual science knowledge gaps among education and racial groups. Looking more closely at different types of social media platforms, few studies have examined how the use of specific social media platforms might shape (science) knowledge gaps across social segments, with the exceptions of Su et al. (2014) which looked at science blog use and SES-based gaps in nanotechnology knowledge and Yoo and Gil-de-Zúñiga (2014) which examined education-based gaps in political knowledge and social media use (blog, Twitter, Facebook). Particularly, Su et al. (2014) found that the SES-based gap in factual nanotechnology knowledge first closed as science blog use increased from no use at all to low use and then widened as science blog use went up from low use to high use. Because social media platforms differ in their media richness/social presence and the levels of self-presentation/self-disclosure they allow for (Kaplan & Haenlein, 2010), one should not assume that individuals' use of social media platforms will uniformly impact their science knowledge as well as disparities in scientific understanding within the broader society. However, the extent to which these effects may vary is an empirical question that requires further investigation.

In this study, I examine the relationship between science knowledge and the use of five social media platforms that are among the most widely used in the U.S. and worldwide for both general purposes and news use specifically: Facebook, YouTube, Instagram, Twitter, and TikTok (Newman et al., 2021). The five social media platforms differ in a number of ways (see Table 3.1).

Table 3.1. Launch year, proportion of sampled users using each social network for any purpose (for news purpose) in the last week, audience age distribution, primary platform use, and primary modality of five social media platforms.

	Launch year	Weekly use (weekly news use)*	Age distribution	Primary use	Primary modality
Facebook	2004	60% (30%)	Slightly older	News; ordinary people	Text-based
YouTube	2005	61% (19%)	Evenly split	News; personalities	Visual
Twitter	2006	21% (11%)	Skews younger	Mainstream news	Text-based
Instagram	2010	40% (12%)	Mostly under 45 years old	Personalities	Visual
TikTok	2016	16% (4%)	Mostly under 25 years old	Personalities	Visual

Sources: Data from Newman et al., 2021 (all 46 markets globally). *Data from Newman et al., 2022 (12 markets including UK, USA, Germany, France, Spain, Italy, Ireland, Denmark, Finland, Japan, Australia, and Brazil).

First, they have varied audience makeup. For example, Twitter users tend to be younger, have higher incomes and educational attainment, and identify themselves as Democrats, compared to the general U.S. adult population (Wojcik & Hughes, 2019), although this may have changed drastically since Elon Musk took over the platform in 2022 (Anderson, 2023). Second, the types of information sources that people pay attention to on these platforms also differ. For instance, users of Instagram and TikTok tend to pay attention to Internet personalities. However, this does not mean that serious issues such as COVID-19 are not discussed on these platforms. News stories blend in with videos and images shared by users of Instagram and TikTok and tend to be highly engaging to reach a wide audience (Newman et al., 2021). Third, these platforms also differ in their primary communication modality. Instagram, TikTok, and YouTube are primarily visual platforms where users create, share, and discover images, videos, visual stories, and the like, whereas Facebook and Twitter are more text-based (Pelled et al., 2017). Because audiences with lower levels of educational attainment tend to process audiovisual information better than text-based information (Grabe et al., 2009), differences in the primary communication modality of the social media platforms may imply differential effects on the knowledge gap phenomenon. In light of this discussion, I pose the following research questions:

RQ1a: Does use of (1) Facebook, (2) Twitter, (3) YouTube, (4) Instagram, and (5) TikTok have a positive, negative, or no relationship to factual knowledge?

RQ1b: Does use of (1) Facebook, (2) Twitter, (3) YouTube, (4) Instagram, and (5) TikTok have a positive, negative, or no relationship to perceived knowledge?

RQ2a: Does use of (1) Facebook, (2) Twitter, (3) YouTube, (4) Instagram, and (5) TikTok have a positive, negative, or no relationship to the gap in factual knowledge among education and racial groups?

RQ2b: Does use of (1) Facebook, (2) Twitter, (3) YouTube, (4) Instagram, and (5) TikTok have a positive, negative, or no relationship to the gap in perceived knowledge among education and racial groups?

The role of issue contexts in shaping science knowledge gaps

Besides media systems that disseminate information about public issues in society, characteristics of the issues themselves could also affect gaps in knowledge pertaining to these issues across social segments. Particularly, the more an issue appeals to the basic concerns within a society, the more likely that members of the society, from all walks of life, will overcome some of the personal and system barriers contributing to knowledge gaps and engage with the issue, ultimately equalizing knowledge distribution within the society (Donohue et al., 1975). Relatedly, the more an issue evokes social conflict within a society, the more likely that widespread concerns will occur within the society as a result of the conflict, and the more likely that both high and low SES segments will pay attention to the issue, diminishing the knowledge gap between these segments (Donohue et al., 1975). Indeed, controversy surrounding an issue can reduce knowledge gaps due to increased flows of information (Bauer & Bonfadelli, 2002). Additionally, complex issues could instigate wider SES-based knowledge gaps than relatively simple issues (Moore, 1987). As most science issues are "beyond hard" for the ordinary citizen to grasp (Xenos, 2017, p. 285), we should expect to see persistent gaps in knowledge about many scientific issues among sociodemographic segments. Echoing calls for more research that compares across science topics, we examine how knowledge gaps might be conditioned on science issue contexts.

As discussed in chapter 1, the three wicked science issues examined in this study—HGE, AI, and COVID-19 (as well as its related vaccines)—differ meaningfully in a number of ways. Overall, because COVID-19 (a) involves high, tangible personal risks and hence appeals to the basic concerns of members of the public much more than HGE and AI; (b) receives much more media attention and public discussion in society than HGE and AI; and (c) raises controversy and political conflict (especially in the U.S.) more than HGE and AI, I expect that the factual knowledge gap pertaining to COVID-19 across social segments will be smaller than those of HGE or AI:

RQ3: How does the gap in perceived knowledge among education and racial groups differ across wicked science issues?

H1: The gap in factual knowledge among education and racial groups will be smaller for COVID-19 than for human gene editing or AI.

Methods

To answer the research questions and test the hypothesis, I used three datasets examining U.S. public opinion on HGE, AI, and COVID-19, respectively. For the HGE dataset, a nationally representative online survey with $N = 1{,}600$ U.S. adults aged 18 years and older was conducted by YouGov in December 2016 and January 2017. The completion rate was 41.7% (American Association for Public Opinion Research [AAPOR], 2016; Callegaro & DiSogra, 2009). To ensure representativeness across sociodemographic characteristics, YouGov matched respondents to a sampling frame based on gender, age, race, education, political ideology, party identification, and political interest. The sampling frame was constructed using stratified sampling from the Census Bureau's 2010 American Community Survey. Matched cases were weighted to the sampling frame based on propensity scores.

For the AI dataset, data were collected through a nationally representative web-survey with U.S. adults aged 18 years and older, conducted by YouGov from February to March 2020. The survey sample was randomly selected from YouGov's U.S. panel, which had 2 million respondents. The final sample size was $N = 2{,}700$ with a completion rate of 41.3%, defined as the percentage of panel members invited to the study who provided a usable response (AAPOR, 2016). YouGov then used propensity score matching techniques for adjustment to make the sample representative of the U.S. population in terms of sociodemographic characteristics including age, education, gender, race, party identification, and political ideology.

For the COVID-19 dataset, I relied on data from an online survey of a national sample of 1,306 U.S. adults aged 18 years and older who had experience using Instagram. The survey was conducted by Forthright in March 2022. For the purpose of comparing results across datasets, I constructed weights for the COVID-19 sample and matched it with the 2020 AI sample—which served as the reference sample—in terms of sociodemographic characteristics including age, education, gender, race, Hispanic ethnicity, party identification, and political ideology. Table 3.2 shows descriptive statistics of demographic variables for the three datasets.

Table 3.2. Descriptive statistics of demographics of the HGE, AI, and COVID-19 samples.

	2016-17 HGE (*N* = 1,600)	2020 AI (*N* = 2,700)	2022 COVID-19 (*N* = 1,306)
Age	*M* = 46.7, *SD* = 16.7	*M* = 48.9, *SD* = 17.6	*M* = 48.8, *SD* = 16.9
Gender			
Male	48.5%	48.7%	43.5%
Female	51.5%	51.3%	56.5%
Race			
White	66.9%	63.5%	62.8%
Black	11.7%	12.0%	21.8%
Other	21.4%	24.5%	15.4%
Hispanic or Latino	14.3%	16.0%	13.1%
Political ideology (7 = very conservative)	*M* = 4.1, *SD* = 1.5	*M* = 4.1, *SD* = 1.7	*M* = 4.1, *SD* = 1.6
Party identification			
Democrat	40.7%	46.7%	44.0%
Independent	19.7%	16.3%	18.8%
Republican	35.9%	37.0%	37.2%
Education			
No college degree	41.3%	38.9%	29.8%
Some college or 2-year college degree	32.0%	31.9%	42.2%
4-year college degree or higher	26.7%	29.2%	28.0%

Measures

Factual science knowledge

In the HGE dataset, factual knowledge of HGE was assessed with four true/false statements using 4-point scales (1 = "definitely true," 4 = "definitely false"), including (a) "Over time, human DNA has picked up pieces of DNA from different species and viruses that naturally mixed in with human DNA" (T), (b) "Personal behavior or environmental factors cannot change human DNA" (F), (c) "To date, no scientists have started human gene editing trials" (F), and (d) "According to scientists, human beings developed from earlier species of animals" (T). Correct answers were coded as "1." False answers and "don't know" were coded as "0." Correct answers were summed up for each respondent to form a single variable measuring HGE factual knowledge (with a range of 0-4, *M* = 1.8, *SD* = 1.3).

In the AI dataset, factual knowledge of AI was measured with nine true/false statements using 4-point scales (1 = "definitely false," 4 = "definitely true"): (a) "AI research began in the early 2000s" (F), (b) "Programmers of AI know exactly how their algorithms adapt to new information" (F), (c) "Self-driving cars are currently being road tested in all 50 states" (T), (d) "The news you see on Facebook news feeds is curated by AI" (T), (e) "Federal law prohibits financial institutions from using AI in lending decisions" (F), (f) "President Trump signed an executive order to increase research and development on AI technology" (T), (g) "Tech companies use AI to combat online misinformation in U.S. elections" (T), (h) "When AI is used to make hiring decisions it is always free of bias" (F), and (i) "Police use of AI can result in systematic targeting of specific neighborhoods" (T). Correct answers were coded as "1." False answers and "don't know" were coded as "0." Correct answers were summed up for each respondent to form a single variable measuring AI factual knowledge (with a range of 0-9, *M* = 3.9, *SD* = 2.3).

Using the same 4-point scale, factual knowledge of COVID-19 was assessed with three true/false statements: (a) "COVID-19 can be treated with antibiotics" (F), (b) "COVID-19 is a respiratory syndrome caused by SARS-CoV-2 virus infection" (T), and (c) "COVID-19 is spread through droplet transmission only" (F). Correct answers were coded as "1." False answers and "don't know" were coded as "0." Correct answers were summed up for each respondent to form a single variable measuring COVID-19 factual knowledge (with a range of 0-3, *M* = 1.8, *SD* = .9).

Perceived science knowledge

In the HGE dataset, perceived knowledge of HGE was measured on a 5-point scale item (1 = "not at all informed," 5 = "very informed") that asked respondents how informed they would say they are about human gene editing (*M* = 1.9, *SD* = .9). Similarly, perceived knowledge in the AI dataset was measured by averaging five 5-point scale items (1 = "not at all informed," 5 = "very informed") asking respondents how informed they are about (a) the science behind AI, (b) concrete uses or applications of AI, (c) impacts of AI on society, (d) regulatory or legal questions emerging from AI applications, and (e) what kinds of information companies collect about ordinary citizens, respectively (*M* = 2.4, *SD* = .9, Cronbach's *alpha* = .92). Perceived knowledge of COVID-19 was measured on an item using the same scale that asked respondents how informed they are about COVID-19 (*M* = 4.1, *SD* = .8).

Social media use

In the HGE and COVID-19 datasets, use of specific social media platforms was measured by a series of 7-point scale items (1 = "less than once a month," 7 = "multiple times per day") asking respondents how often they use Facebook (HGE: *M* = 6.5, *SD* = 2.1; COVID-19: *M* = 5.3, *SD* = 2.3), Twitter (HGE: *M* = 2.9, *SD* = 2.5; COVID-19: *M* = 2.7, *SD* = 2.7), YouTube (HGE: *M* = 5.1, *SD* = 2.1; COVID-19: *M* = 5.1, *SD* = 2.2), Instagram (HGE: *M* = 2.8, *SD* = 2.5; COVID-19: *M* = 3.9, *SD* = 2.6), and TikTok (not asked in the HGE dataset; COVID-19: *M* = 2.2, *SD* = 2.7), with those who answered "never" recorded as "0." In the AI dataset, use of specific social media platforms was measured on a 6-point scale item (1 = "less than 10 minutes a day," 7 = "more than 3 hours a day") that asked respondents in the past week about how much time on average they have spent each day on Facebook (*M* = 3.4, *SD* = 1.9), Twitter (*M* = 1.9, *SD* = 1.5), YouTube (*M* = 3.3, *SD* = 1.9), Instagram (*M* = 2.0, *SD* = 1.6), and TikTok (*M* = 1.3, *SD* = 1.0), with those answering "none" recorded as "0."

Demographics

Age was measured as a continuous variable (HGE: *M* = 46.7, *SD* = 16.7; AI: *M* = 48.9, *SD* = 17.6; COVID-19: *M* = 48.8, *SD* = 16.9). Gender was a dichotomous variable with male coded as "1" and female coded as "2" (HGE: 48.5% males; AI: 48.7% males; COVID-19: 43.5% males). Race was measured by a dichotomous variable with White coded as "1" and Black and other racial minorities coded as "2" (HGE: 66.9% White; AI: 63.5% White; COVID-19: 62.8% White). In line with Tichenor et al. (1970) and other knowledge gap research, education was employed as an indicator of SES. Education was an ordinal variable with six levels ranging from "no high school diploma" (coded as "1") to "post graduate degree" (coded as "6"). Across all three samples, the median value for education was "3" or "attended some college" (HGE: *SD* = 1.5; AI: *SD* = 1.5; COVID-19: *SD* = 1.4).

Analysis

To examine the role of social media use in shaping factual and perceived science knowledge, as well as gaps in knowledge across sociodemographic segments (RQ1 through RQ2), I conducted hierarchical ordinary least squares (OLS) regression analyses, entering independent variables in blocks into the regression based on their presumed causal order. The final two blocks of the regression contained interaction terms. To prevent multicollinearity between interaction terms and their constitutive components, main effect variables were standardized before being multiplied together to create the interaction terms (Cohen & Cohen, 1983).

To examine the role of issue contexts in shaping science knowledge gaps (H1 and RQ3), I compared the zero-order correlations between knowledge and sociodemographic characteristics (i.e., education and race) across the three issues, following Tichenor et al's approach (1970). A stronger, positive relationship between knowledge and education, for example, would indicate a larger gap in knowledge between education groups with more highly educated individuals possessing more knowledge than their less educated counterparts, whereas a weaker relationship between knowledge and education would indicate a smaller education-based knowledge gap (Tichenor et al., 1970). Fisher's z-transformation for correlation coefficients was used to test the difference between two correlations (Fisher, 1921).

Results

Table 3.3 shows the results from the OLS regression models predicting factual and perceived knowledge of the three science issues. The overall pattern across the six models is largely consistent, although we outline some differences below.

Beginning with the demographic predictors, men were more likely than women to score higher on both factual and perceived knowledge of HGE and AI (but not COVID-19). Whites were more likely to possess more factual knowledge of HGE and AI than non-Whites, whereas racial minorities tended to report higher perceived knowledge of HGE than Whites. Individuals with higher education were more likely to score higher on both factual and perceived knowledge across all three issues. Finally, younger adults were more likely to express higher levels of perceived knowledge of HGE and lower levels of perceived knowledge of COVID-19.

Table 3.3. Regression models predicting factual knowledge and perceived knowledge of HGE, AI, and COVID-19.

	HGE		AI		COVID-19	
	Factual knowledge	Perceived knowledge	Factual knowledge	Perceived knowledge	Factual knowledge	Perceived knowledge
Block 1: Demographics						
Age	-0.01	-0.10**	0.03	-0.03	0.00	0.16***
Gender (female)	-0.06*	-0.14***	-0.10***	-0.14***	0.05	-0.00
Race (non-White)	-0.19***	0.09**	-0.08***	0.01	-0.01	-0.02
Education	0.24***	0.11***	0.25***	0.21***	0.11***	0.08**
Incremental R^2	*11.8%****	*7.4%****	*8.6%****	*9.7%****	*1.5%****	*2.1%****
Block 2: Social media use						
Facebook use	-0.06*	-0.04	-0.01	0.02	-0.01	-0.03
Twitter use	0.07*	0.18***	0.11***	0.11***	0.07*	0.10**
YouTube use	0.05	0.04	0.03	0.14***	0.02	0.09**
Instagram use	0.05	-0.07*	-0.05	0.06**	-0.01	0.10**
TikTok use			-0.04	0.12***	-0.01	0.00
Incremental R^2	*1.5%****	*3.2%****	*1.0%****	*7.9%****	*0.5%*	*3.4%****
Block 3: 2-way interactions						
Education × Facebook use	-0.03	0.00	0.04*	-0.04*	0.01	-0.02
Education × Twitter use	-0.00	-0.03	-0.01	-0.03	-0.02	0.04
Education × YouTube use	0.03	-0.02	0.04	-0.01	0.06*	0.00
Education × Instagram use	0.01	0.00	-0.01	-0.06***	0.04	-0.04
Education × TikTok use			-0.02	0.00	0.03	-0.03
Education × Race	0.02	0.01	0.00	-0.01	0.06*	-0.01
Race × Facebook use	0.02	0.00	0.00	0.01	0.02	0.06*
Race × Twitter use	-0.01	0.07**	-0.01	-0.01	0.00	-0.00
Race × YouTube use	0.04	-0.05	-0.02	-0.02	0.00	0.08**
Race × Instagram use	-0.02	0.03	0.00	-0.01	0.07**	0.06*
Race × TikTok use			0.04*	-0.02	-0.02	0.00
Incremental R^2	*0.6%*	*1.0%*	*0.7%**	*0.7%**	*1.8%**	*1.6%**
Block 4: 3-way interactions						
Education × Race × Facebook use	0.02	0.01	-0.01	-0.05**	0.00	0.00
Education × Race × Twitter use	0.07**	0.04	-0.02	-0.02	-0.00	0.00
Education × Race × YouTube use	-0.02	0.04	0.00	-0.05*	0.07*	-0.02
Education × Race × Instagram use	0.03	0.02	-0.01	-0.05*	0.03	-0.04
Education × Race × TikTok use			-0.02	-0.02	-0.01	-0.02
Incremental R^2	*0.7%**	*0.2%*	*0.1%*	*0.4%**	*0.6%*	*0.2%*
Total R^2	***14.6%***	***11.7%***	***10.5%***	***18.6%***	***4.4%***	***7.2%***

Note: Cell entries are final standardized regression coefficients, except for those in Blocks 3 and 4, which are before-entry standardized regression coefficients that represent the impact of a given variable with all previous blocks controlled for.
The sample size is $N = 1,600$ for the HGE dataset, $N = 2,700$ for the AI dataset, and $N = 1,306$ for the COVID-19 dataset.
*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.

RQ1a and RQ1b asked how use of specific social media platforms might be associated with factual and perceived knowledge levels, respectively. Across the regressions, Twitter use was positively related to both factual and perceived knowledge of all three science issues. Overall, social media use seemed to be associated with increased perceived knowledge. Specifically, those who used Twitter, YouTube, and Instagram more often were more likely to express higher perceived knowledge of AI and COVID-19. Individuals who used TikTok more often were also more likely to report higher perceived knowledge of AI. However, those who used Instagram more often were more likely to express lower perceived knowledge of HGE. In addition, Facebook use was negatively related to factual knowledge of HGE but had virtually no relationship to perceived knowledge of HGE as well as both knowledge measures of the other two science issues.

RQ2a and RQ2b asked how use of specific social media platforms might impact the gap in factual and perceived knowledge, respectively, among education and racial groups. Figures 3.1 and 3.3 show the two-way interaction effects between education and social media use and between race and social media use on levels of factual and perceived knowledge, respectively. Moreover, the intersectionality of education and race could also matter for how knowledge gaps turn out. To test these possibilities, I ran a series of three-way interactions between education, race, and social media use on factual knowledge (Figures 3.2) and perceived knowledge (Figure 3.4).

In response to RQ2a (factual knowledge as DV), the two-way interactions between education and social media use were significant for Facebook use on AI factual knowledge ($\beta = .04$, $p \leq .05$) and for YouTube use on COVID-19 factual knowledge ($\beta = .06$, $p \leq .05$).

As Figure 3.1 depicts, as Facebook use increased, the gap in factual knowledge of AI widened between high and low education groups, particularly driven by a downtick in knowledge among low-education respondents (Figure 3.1A).

Increased YouTube use can similarly lead to an increased gap in factual knowledge of COVID-19 between high and low education groups, as high-education respondents gained more factual knowledge and low-education respondents reported less knowledge with increased YouTube use (Figure 3.1C).

The two-way interactions between race and social media use were significant for TikTok use on AI factual knowledge ($\beta = .04$, $p \leq .05$) and for Instagram use on COVID-19 factual knowledge ($\beta = .07$, $p \leq .01$).

As TikTok use increased, the gap in factual knowledge of AI between Whites and racial minorities decreased, primarily driven by a decrease in Whites' knowledge (Figure 3.1B).

As Instagram use increased, the gap in factual knowledge of COVID-19 between Whites and racial minorities also decreased, with Whites expressing less knowledge and non-Whites more knowledge with increased Instagram use (Figure 3.1D).

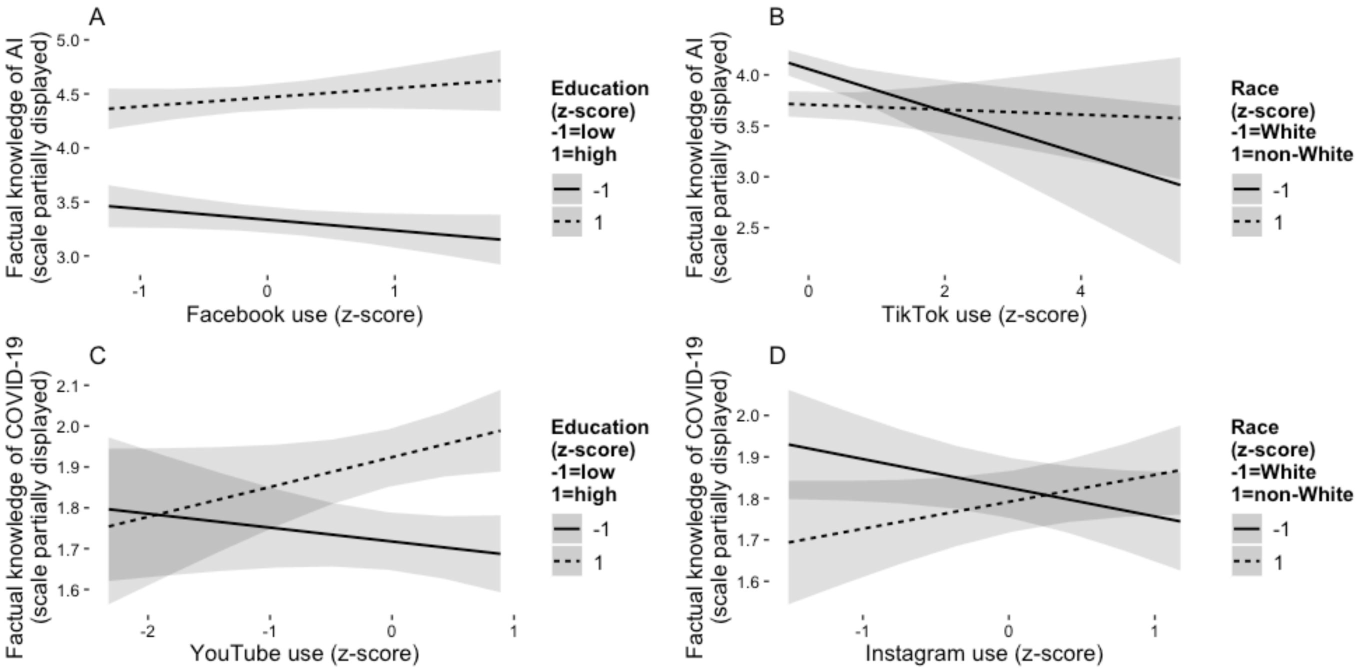


Figure 3.1. Two-way interactions between *education* and *social media use* and between *race* and *social media use* on *factual knowledge* of science issues. Error bars are 95% confidence intervals.

Further, regarding the intersectionality of education and race, **the three-way interactions between education, race, and social media use** were significant for Twitter use on HGE factual knowledge ($\beta = .07$, $p \leq .01$) and for YouTube use on COVID-19 factual knowledge ($\beta = .04$, $p \leq .05$).

As Figure 3.2A shows, among White respondents, increased <u>Twitter</u> use can lead to a diminished gap in factual knowledge of HGE between high and low education groups as low-education White respondents caught up with their high-education counterparts in knowledge with increased Twitter use. Among non-White respondents, increased Twitter use was associated with a wider education-based gap in HGE factual knowledge, as high-education non-White respondents gained knowledge at a significantly faster rate than their low-education counterparts.

Among White respondents, increased YouTube use had virtually no impact on the education-based gap in COVID-19 factual knowledge, whereas among non-White respondents increased YouTube use seemed to increase the knowledge gap between high and low education groups as those with higher education gaining more knowledge and those with lower education reporting less knowledge (Figure 3.2B).

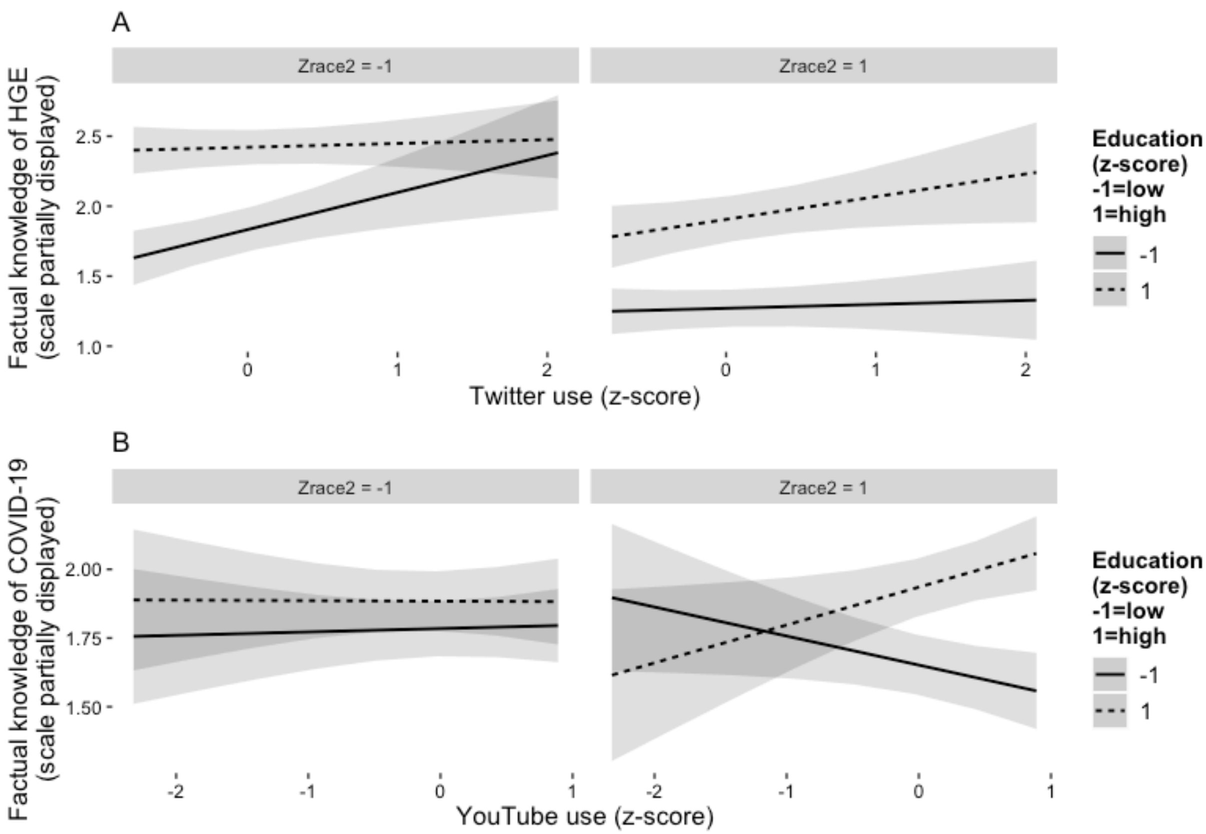


Figure 3.2. Three-way interactions between *education*, *race*, and *social media use* on *factual knowledge* of science issues. Error bars are 95% confidence intervals.

In response to RQ2b (perceived knowledge as DV), the two-way interactions between education and social media use were significant for Facebook (β = -.04, $p \leq .05$) and Instagram (β = -.06, $p \leq .001$) use on AI perceived knowledge.

As Figure 3.3 depicts, with increased Facebook use, the gap in perceived knowledge of AI narrowed between high and low education groups (Figure 3.3A).

Similarly, increased Instagram use can lead to a decreased education-based gap in AI perceived knowledge, primarily due to low-education respondents reporting more perceived knowledge as they used Instagram more frequently (Figure 3.3B).

The two-way interactions between race and social media use were significant for Facebook ($\beta = .06$, $p \leq .05$), Instagram ($\beta = .06$, $p \leq .05$), and YouTube ($\beta = .08$, $p \leq .01$) use on COVID-19 perceived knowledge and for Twitter use on he perceived knowledge ($\beta = .07$, $p \leq .01$).

As Facebook use increased, the gap in perceived knowledge of COVID-19 between White and non-White respondents decreased (Figure 3.3C).

Similarly, the gap in perceived knowledge of COVID-19 between Whites and racial minorities diminished as Instagram (Figure 3.3D) and YouTube (Figure 3.3E) use increased, with non-Whites growing COVID-19 perceived knowledge at a significantly faster rate than Whites.

Twitter use seemed to widen the race-based gap in perceived knowledge of HGE, with non-Whites growing perceived knowledge significantly faster than Whites (Figure 3.3F).

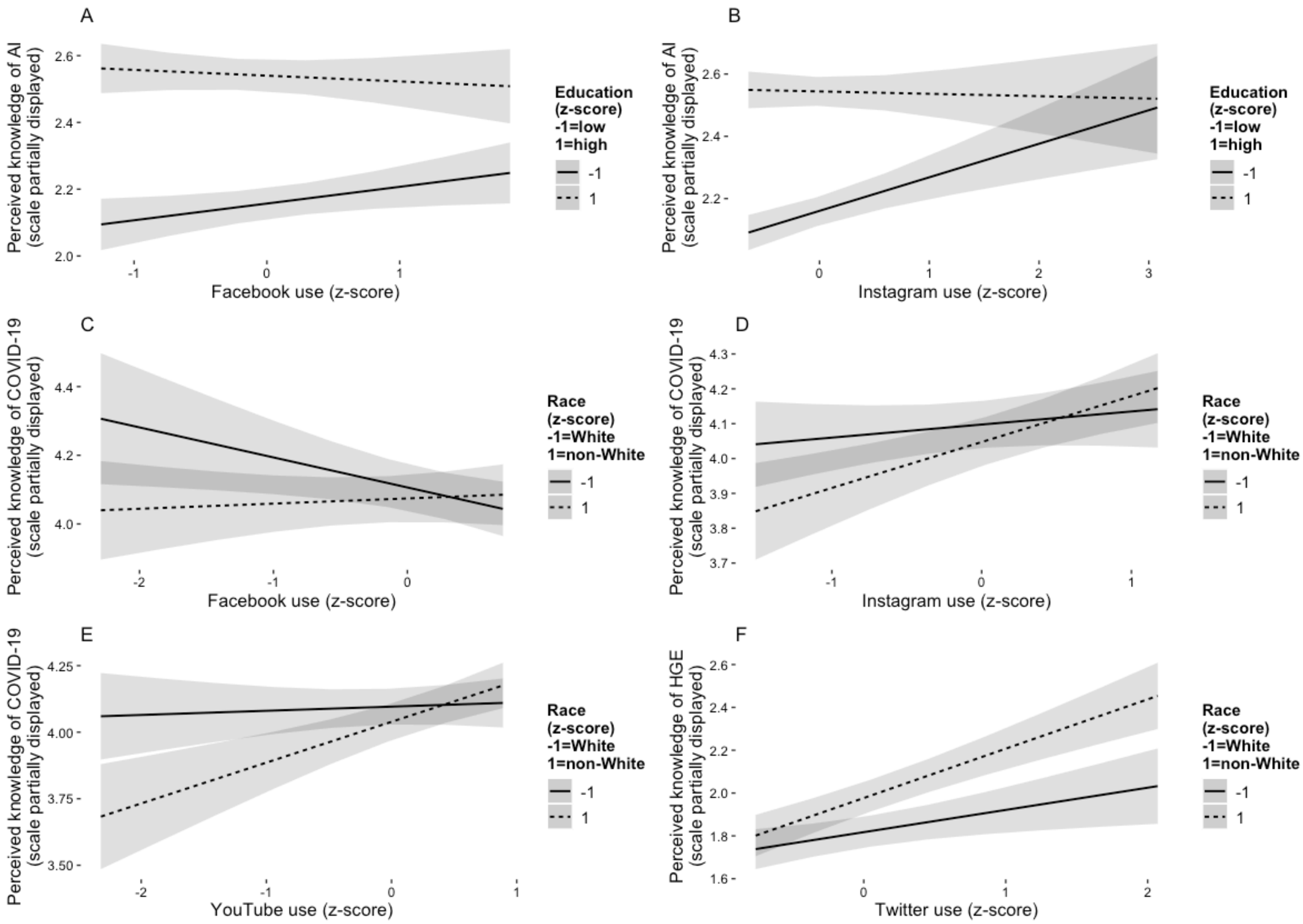


Figure 3.3. Two-way interactions between *education* and *social media use* and between *race* and *social media use* on *perceived knowledge* of science issues. Error bars are 95% confidence intervals.

Further, **the three-way interactions between education, race, and social media use** were significant for Facebook (β = -.05, $p \leq .01$), YouTube (β = -.05, $p \leq .05$), and Instagram (β = -.05, $p \leq .05$) use on perceived knowledge of AI.

As Figure 3.4 shows, increased Facebook use had limited impact on the education-based gap in AI perceived knowledge among White respondents but was associated with a diminished education-based gap among non-White respondents whereby low-education non-Whites caught up with their high-education counterparts in AI perceived knowledge with increased Facebook use (Figure 3.4A).

Increased YouTube use widened the education-based gap in AI perceived knowledge among Whites but narrowed the education-based gap among non-Whites (Figure 3.4B).

Finally, increased Instagram use slightly narrowed the education-based gap in AI perceived knowledge among White respondents but greatly diminished the gap among non-Whites; among non-Whites, low-education respondents increased their AI perceived knowledge with more Instagram use whereas high-education respondents decreased their AI perceived knowledge (Figure 3.4C).

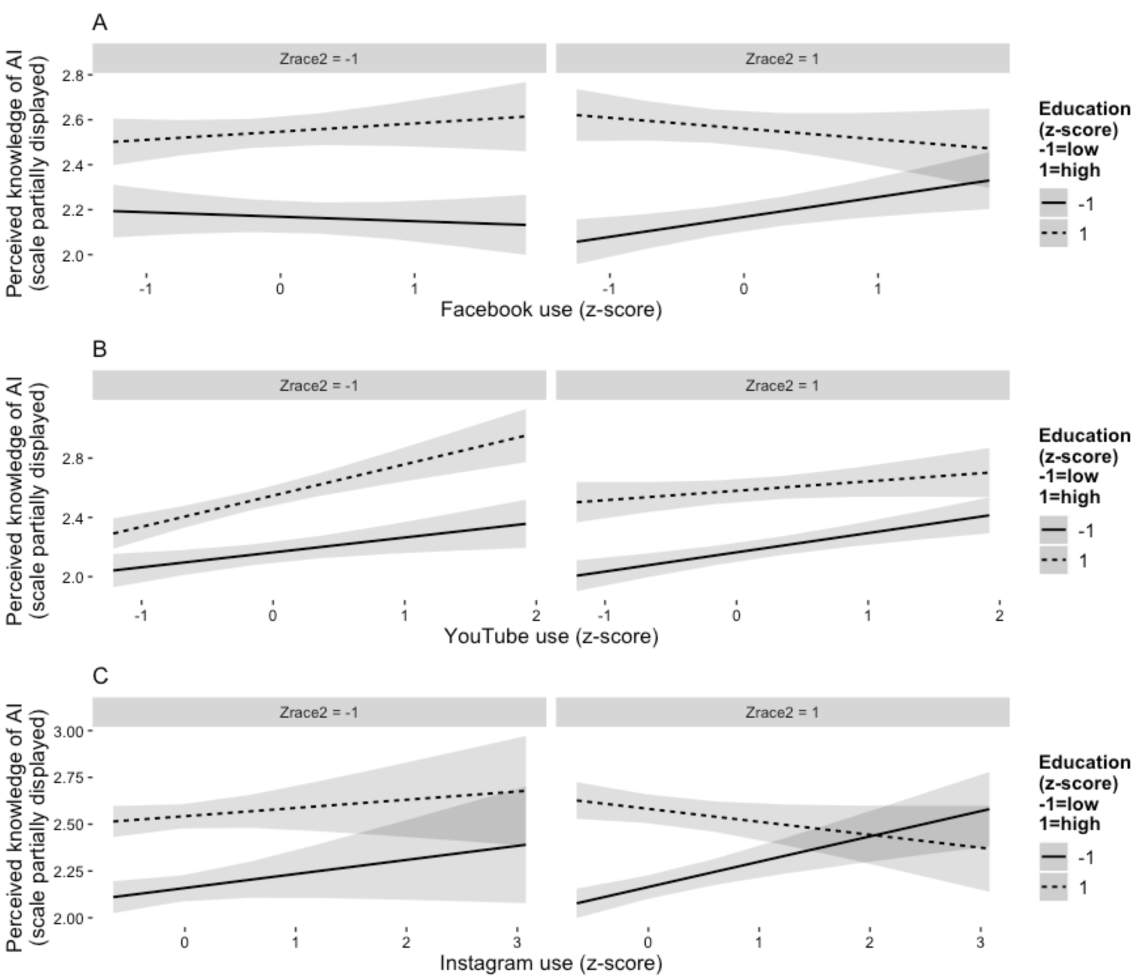


Figure 3.4. Three-way interactions between *education*, *race*, and *social media use* on *perceived knowledge* of science issues. Error bars are 95% confidence intervals.

Turning to the role of issue contexts, **RQ3 asked how the gap in perceived knowledge among those groups might differ across the three science issues**. Relatedly**, H1 hypothesized that the gap in factual knowledge among education and racial groups would be smaller for COVID-19 than for HGE or AI**.

Supporting H1, a smaller gap in factual knowledge between high and low education groups was identified for the issue of COVID-19 than for HGE and AI, as the correlations between factual knowledge and education in the HGE dataset (Fisher's $z = 4.81$, $p \leq .001$) and in the AI dataset (Fisher's $z = 4.61$, $p \leq .001$) were significantly stronger than that in the COVID-19 dataset. A smaller gap in factual knowledge between Whites and racial minorities was also identified for the issue of COVID-19 than for HGE and AI, as factual knowledge was significantly more strongly correlated with race in the contexts of HGE (Fisher's $z = -5.59$, $p \leq .001$) and AI (Fisher's $z = -3.48$, $p \leq .001$) than COVID-19. Additionally, factual knowledge was significantly more correlated with education for HGE than AI (Fisher's $z = -2.89$, $p \leq .01$).

Addressing RQ3, the correlation between perceived knowledge and education was significantly stronger in the AI dataset than in the HGE (Fisher's $z = -2.73$, $p \leq .01$) and COVID-19 (Fisher's $z = 2.13$, $p \leq .05$) datasets; the correlation between perceived knowledge and race was significantly stronger in the HGE dataset than in the AI (Fisher's $z = 2.81$, $p \leq .01$) and COVID-19 (Fisher's $z = 4.11$, $p \leq .001$) datasets.

Table 3.4. Zero-order correlations between (a) knowledge and education and (c) knowledge and race and Fisher's z-test statistics comparing the correlations between (b) knowledge and education and (d) knowledge and race across three science issues.

	Factual knowledge			Perceived knowledge		
	HGE	AI	COVID-19	HGE	AI	COVID-19
(a) Pearson's r between education and knowledge	0.29***	0.26***	0.11***	0.10***	0.19***	0.12***
(b) Fisher's z-test						
Factual knowledge HGE		0.97	4.81***			
Factual knowledge AI			4.61***			
Perceived knowledge HGE					-2.73**	-0.52
Perceived knowledge AI						2.13*
(c) Point-biserial r_{pb} between race (non-White) and knowledge	-0.20***	-0.11***	0.01	0.13***	0.05*	-0.02
(d) Fisher's z-test						
Factual knowledge HGE		-2.89**	-5.59***			
Factual knowledge AI			-3.48***			
Perceived knowledge HGE					2.81**	4.11***
Perceived knowledge AI						1.93

Note. Table entries in row (a) are zero-order Pearson's product-moment correlations between education and knowledge; in rows (b) are Fisher's z-test statistics for comparisons of correlations from (a); in row (c) are zero-order point-biserial correlations between race and knowledge; in rows (d) are Fisher's z-test statistics for comparisons of correlations from (c).
*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.

Discussion

This study examined the influences of education, race, use of specific social media platforms, and issue contexts on both factual and perceived knowledge gaps across three wicked science issues including HGE, AI, and COVID-19. We found that increased social media use overall predicted larger factual science knowledge gaps and smaller perceived science knowledge gaps between high and low education groups. Compared with more highly educated Americans, those with less education are less likely to gain factual science knowledge from increased social media use while they are more likely to gain perceived science knowledge. Racial minorities are more likely to gain both factual and perceived science knowledge with increased social media use than White Americans. We also found that increased social media use was linked to wider factual science knowledge gaps and narrower perceived science knowledge gaps among racial minorities than among Whites.

These findings increase our understanding of disparities in scientific understanding in an era of social media and post-normal science and could inform efforts to accelerate greater use of equity-based communications strategies to fortify a more democratic civic science society. However, before elaborating on our findings and their implications, we address some limitations of the current work.

Limitations

The first limitation concerns sampling. Specifically, our COVID-19 sample was limited to U.S. adults who had experience using Instagram (with or without an account). As of 2022 when the data were collected, 48.6% of Americans used Instagram (NapoleonCat, 2022). Compared to the overall U.S. adult population, Instagram users tended to skew younger, female, and more highly education (NapoleonCat, 2023a, 2023b; Statista, 2022). To counter this, we weighted the COVID-19 sample to match it with the AI sample—the largest nationally representative sample among the three datasets—in terms of key demographic variables, which increased the representativeness of the COVID-19 sample and enhanced the comparability of our results across the three issues. It is noteworthy, however, that weighting beyond demographic variables to account for other variables in which the samples could potentially differ is desirable when the goal is to draw comparisons across datasets.

A second limitation relates to the measurement of some of our variables. We used single-item measures to assess use of specific social media platforms, as well as perceived knowledge of HGE and COVID-19, which could have reduced the sensitivity of these survey instruments due to increased random measurement error. However, had multi-item measures been used, the observed relationships between social media use and perceived and factual science knowledge would likely have been even stronger. In addition, the number of true/false statements measuring factual knowledge differed across the three issues examined, which might have limited the variance observed for issues whose factual knowledge measures consisted of fewer items (e.g., HGE and COVID-19). Finally, our social media use measures also focused on overall use frequency. Future research may use multi-item measures to tap into the specific activities that people engage in when they are on those social media platforms.

A third issue concerns causality. The analysis reported here assumes, with theoretical justifications, that social media use influences how individuals develop science knowledge, but we cannot rule out alternative causal orders with cross-sectional survey data. For example, a reciprocal relationship between social media use and science knowledge is possible whereby one's social media use affects knowledge levels and knowledge in turn drives social media use patterns.

Theoretical implications

With these limitations in mind, our findings have important theoretical and practical implications. First, although research on social media and science knowledge gaps has for the most part examined aggregate social media use, our findings suggest that there is value in differentiating specific social media platforms because use of different platforms may not shape knowledge (gaps) in the same way. For example, whereas frequent Twitter users consistently reported higher factual knowledge on all three science issues even after controlling for demographics, frequent Facebook users expressed less factual knowledge about HGE and frequent users of YouTube, Instagram, and TikTok did not necessarily possess higher or lower factual science knowledge than people who used those platforms less often. Such differences in platform use's influence on factual science knowledge may in part be attributed to differences in the types of information sources that people pay attention to on these platforms. While users of Twitter tend to seek out mainstream news when they are on the platform, potentially contributing to the growth in their factual knowledge, users of Instagram, TikTok, YouTube, and Facebook tend to pay attention to a mix of Internet personalities, ordinary people, and news (Newman et al., 2021), which may not be the types of information sources that are most conducive to factual science knowledge acquisition. Future research may examine how social media platforms' features, attributes, and use patterns could shape science knowledge gaps to inform our understanding of and efforts to mitigate such inequalities.

Second, consistent with Ladwig et al (2012) who found that factual nanotechnology knowledge and perceived nanotechnology knowledge were only slightly correlated with each other and were predicted differently by media use and cognitive processing variables, our results also show that the identification of knowledge gaps depends on how knowledge is conceptualized and measured. Scholars should be careful to not conflate factual knowledge measures with self-reported knowledge measures when examining science knowledge gaps. For instance, whereas social media use seemed to overall widen education-based gaps in factual science knowledge, it appeared to narrow education-based gaps in perceived knowledge. When high-education individuals used YouTube and Facebook more, they gained more factual knowledge about COVID-19 and AI, respectively; in contrast, when low-education individuals used these platforms more, they did not acquire as much knowledge as their high-education counterparts (in fact, they showed a downtick in factual knowledge with increased social media use), subsequently widening the education-based gaps in factual knowledge. However, as low-education individuals used Facebook and Instagram more, their perceived knowledge of AI increased at a significantly faster rate than that of their high-education counterparts, leading to reduced perceived knowledge gaps between education groups.

Third, the identification of knowledge gaps could also depend on the science issue at hand. Previous research has suggested that knowledge gaps are contingent upon issue characteristics such as issue complexity (and thus knowledge complexity), controversy, issue appeal to a social system, and media publicity (Bauer & Bonfadelli, 2002; Donohue et al., 1975; Moore, 1987). Consistent with theoretical expectations, we found smaller education- and race-based factual knowledge gaps for COVID-19 than for AI and HGE, arguably because COVID-19 generated much more media and policy attention, public concern, and political controversy, all of which led to increased information flows across communities and all walks of life. Additionally, these science issues differ in their scope, temporality, and risks involved. While COVID-19 has a constrained issue scope, HGE and especially AI have a very broad range of applications touching many aspects of society. Whereas COVID-19 is relatively transient and involves highly tangible, personal risks to the self, the impacts of HGE and AI are more long-term, latent, and far-reaching at the societal level. These and additional issue characteristics may also affect people's concerns, salience, and self-efficacy regarding science issues and the way they acquire knowledge (Ettema et al., 1983; Shim, 2008). More systematic research is needed to examine what and how science issue characteristics might matter for knowledge gaps.

Practical implications

Our findings also have implications for equitable science communication practices. When it comes to science knowledge acquisition, social media use could differentially benefit different sociodemographic segments. People with lower levels of educational attainment are less likely to gain factual knowledge about AI and COVID-19 from increased social media use—specifically, Facebook and YouTube use—than their more highly educated counterparts. Because less highly educated audiences tend to select entertainment over information-oriented content when online (Bonfadelli, 2002), it is possible that people who are less highly educated do not pay as much attention to science topics as their more highly educated counterparts when using social media. While increased social media use facilitates acquisition of factual science knowledge among more highly educated individuals but not among less highly educated individuals, it is more likely to enhance perceived science knowledge among those with less education than among those with higher education. Taken together, less highly educated people are more likely to gain confidence in their own science knowledge with increased social media use when in fact they do not necessarily acquire knowledge. In other words, among less educated segments, increased social media use could contribute to one's "illusion of knowledge," or the tendency to overestimate one's knowledge about various issues (Rock et al., 2005). Our findings also align with Chang et al.'s (2018) finding that low-education individuals are more likely to perceive higher science knowledge than high-education individuals as aggregate social media use for science information increases.

Race-based gaps in science knowledge matters even after accounting for educational influences on science knowledge, partly because individuals are deeply embedded in racially homophilic social networks that narrow one's information diets and interpretation of new information (McPherson et al., 2001). We found that Whites are less likely to acquire factual science knowledge from increased social media use than racial minorities. Specifically, increased TikTok use was linked to decreased factual knowledge of AI among Whites, but the same trend was not identified for non-Whites. Similarly, increased Instagram use was linked to decreased factual knowledge of COVID-19 among Whites whereas among non-Whites factual knowledge of COVID-19 actually increased. These patterns of findings warrant further research attention. For example, do Whites use Instagram and TikTok in ways different from non-Whites, and if so, how? Are Whites more susceptible to science misinformation on social media than non-Whites? Future research should continue to unpack these possibilities. Meanwhile, in accordance with factual knowledge acquisition, racial minorities are also more likely than Whites to gain perceived science knowledge with increased use of Facebook, Twitter, YouTube, and Instagram.

Moreover, when looking at the intersectionality of education and race more closely, we identify overall more severe adverse effects of social media use on factual science knowledge gaps formed on the basis of education among racial minorities than among Whites. Specifically, less highly educated Whites are able to catch up with more highly educated Whites in their HGE factual knowledge as they used Twitter more, leading to a reduced knowledge gap between the two groups. However, the same is not true for racial minorities. Less highly educated non-Whites learn virtually nothing more about HGE when increasing their Twitter use, whereas more highly educated non-Whites acquire significantly more HGE knowledge than their less educated counterparts when Twitter use increases, widening the knowledge gap between the groups. Consistent with these patterns, increased YouTube use had virtually no impact on the gap in COVID-19 factual knowledge between high- and low-education Whites, but it seems to increase the knowledge gap between high- and low-education non-Whites, with more highly educated non-Whites acquiring COVID-19 factual knowledge at a faster rate than less highly educated non-Whites. Finally, social media use is more effective at reducing education-based gaps in perceived science knowledge among racial minorities than among Whites.

These findings together suggest that knowledge gaps can be multifaceted phenomena that warrant attention to the intersectionality of sociodemographic influences. More importantly, special efforts should be made to support less highly educated and low-SES racial minorities so that they could equally benefit from advancements in information technologies as these technologies (including social media) have become increasingly integrated with our life. Efforts toward developing science literacy (Howell & Brossard, 2021) as well as social media literacy (Cho et al., 2022) among low-education minority segments may prove especially fruitful for reducing disparities in scientific understanding in the American society.

Chapter 4

Examining the effectiveness of visual and narrative messaging in mitigating disparities in message elaboration and knowledge of COVID-19 vaccine safety (Study C)

The successful containment of public health crises in modern days largely depends on timely and effective communication between scientific experts, public health authorities at the local and national levels, media professionals, and various publics. Research examining ways to promote public understanding and acceptance of health communication messages has often focused on the role of health literacy. Representing the motivation and capabilities of individuals to access, use, evaluate, and act on health information and to engage in health-related civic matters (NASEM, 2016b; Nutbeam, 2008), health literacy has the potential to facilitate acquisition of health knowledge, use of health services, and medication adherence, among other health outcomes, although the effect of health literacy on decisions, actions, and behaviors is often limited and contingent upon additional factors such as cultural norms, self-efficacy, reasoning skills, trust, and values (for a review, see NASEM, 2016b).

Whereas much research has examined the role of health literacy in enhancing health information interpretation, science literacy could also matter, especially in the context of COVID-19 vaccines that involve rapidly evolving scientific developments and have raised heated political debates (Scheufele et al., 2020). The relationship between science literacy and health information processing and understanding is less clear but nonetheless remains a promising area of research (NASEM, 2016b). Coping with a fast-changing landscape of scientific facts, findings, and uncertainties surrounding COVID-19 vaccines that is coupled with challenges in the information environment, individuals need to possess a basic level of science literacy—or understanding of how science is produced, how science information travels through media systems, and how people make sense of science information when encountering it (Howell & Brossard, 2021)—to more appropriately interpret, evaluate, and integrate COVID-19 vaccine information. However, to what extent science literacy affects processing of COVID-19 vaccine information remains an empirical question. In this study, we examine how science literacy may shape individuals' information elaboration and knowledge acquisition regarding COVID-19 vaccine safety.

In addition, when communicating the availability, safety, and efficacy of the COVID-19 vaccines, practitioners have utilized a variety of messaging strategies—such as information visualizations, storytelling, and comics—to reach diverse publics (see, e.g., CDC, 2023a, 2023b). Despite their various formats, such commonly used messaging strategies can be generally categorized by modalities (i.e., visual vs. textual) and rhetorical modes (i.e., descriptive facts vs. narratives). While communicating health-related information through visual and narrative messaging has received increasing scholarly attention in recent years, little is known about whether the potential effectiveness of those health messages should be primarily attributed to the presence of visuals or the integration of narratives. Here, we examine the relative effectiveness of visual and narrative components of COVID-19 vaccine safety information to inform future communication efforts around health science developments.

More importantly, social justice concerns surged during the COVID-19 pandemic as it had inequitable impacts on people's lives. Subpopulations of society have not benefited to the same degree from medical developments such as the COVID-19 vaccines, due to differences in access, trust, history, and contexts (e.g., Kolbe, 2021; Sokale et al., 2022). Across social segments, disparities also exist in people's access to, use of, and ability to learn from media information regarding the COVID-19 vaccines as well as other public issues (Viswanath et al., 2020). In the face of COVID-19 driven inequities and information inequities, developing communications strategies that can facilitate more democratic public communication around the current pandemic and future health crises becomes an urgent ethical imperative. Addressing such needs, our study tests the effects of visual and narrative information about COVID-19 vaccine safety on population segments' interpretation of the information, with a focus on mitigating the gap in information elaboration and knowledge acquisition between highly science literate Americans and Americans who are less science literate.

Narrative as a rhetorical device for enhancing information understanding and elaboration

Public communication on health science developments has traditionally relied on literacy-based approaches emphasizing scientific rigor and objective, unbiased scientific facts (Yang & Hobbs, 2020). The focus on descriptive facts and arguments, also known as logical-scientific information (Dahlstrom, 2014), can fall subject to the fallacy of the knowledge deficit model, which assumes that communicating scientific facts alone would suffice to encourage attitudinal and behavioral changes among public audiences (Simis et al., 2016), a notion proven to be oversimplified according to empirical research (Brossard et al., 2009; Sturgis & Allum, 2004). Further, fact-based scientific information can be inaccessible, irrelevant, and uninteresting to lay audiences, especially those with limited financial and cognitive resources (Humm et al., 2020). In other words, communicating via fact-based scientific information may exacerbate existing disparities in knowledge and other relevant outcomes across population segments (Dahlstrom, 2014; Tichenor et al., 1970). To offset these drawbacks, practitioners have increasingly adopted alternative messaging strategies including narrative and visual messaging in their public communication about COVID-19 (e.g., Li & Molder, 2021; Li et al., 2022).

Unlike descriptive facts, narratives follow a pre-identified "structure that describes the cause-and-effect relationships between events that take place over a particular time period that impact particular characters" (Dahlstrom, 2014, p. 13614). Whereas fact-based scientific information seeks to capture the general truths about the world and is judged on its alignment with external reality, narratives scrutinize human experiences to construct meanings and realities and are judged on the verisimilitude of the specific situations they present (Dahlstrom, 2021; Fisher, 1989).

When used to convey health and science, narratives seem to offer various benefits. Research has generally found that narratives can facilitate understanding and health and science knowledge acquisition more than descriptive facts and arguments (Glaser et al., 2009; Mazor et al., 2007; Zabrucky & Moore, 1999). When information is presented in narrative formats, as compared to expository formats, a better understanding as well as recall of the information is achieved in a shorter amount of time (Zabrucky & Moore, 1999). In addition, research on the personalization effect shows that presenting words in conversational style rather than formal style encourages more active cognitive processing and deeper learning (Mayer, 2003). Educational content that is more integral to the causal structure of story events is processed in a more thorough and deeper fashion, leading to increased information understanding and acceptance (Dahlstrom, 2010, 2012; Fisch, 2000). Further, employing narrative formats can reduce the gap in health and science information comprehension among individuals with differing levels of reading skills, whereas the gap in information understanding due to disparate reading abilities persist when information is conveyed in expository formats (Michielutte et al., 1992; Zabrucky & Moore, 1999). Narrative messaging seems especially helpful for increasing health and science information understanding and cognitive processing among audiences with low literacy skills (Denton et al., 2015; Michielutte et al., 1992; Zabrucky & Moore, 1999). These properties make narratives particularly suited for communicating health science information to lay publics, especially audiences who are at a disadvantage with regard to cognitive resources. In light of this discussion, we hypothesize:

H1: Individuals exposed to narrative information on COVID-19 vaccine safety will report higher factual knowledge about COVID-19 vaccine safety than individuals exposed to non-narrative information.

H2: Individuals exposed to narrative information on COVID-19 vaccine safety will report higher message elaboration than individuals exposed to non-narrative information.

H3: The difference in factual knowledge about COVID-19 vaccine safety between individuals with higher and lower science literacy will be less pronounced when individuals are exposed to narrative information than if they are exposed to non-narrative information.

H4: The difference in message elaboration between individuals with higher and lower science literacy will be less pronounced when individuals are exposed to narrative information than if they are exposed to non-narrative information.

The effects of visual messaging on information understanding and elaboration

Visuals can help communicate health and science in a way that words alone cannot. Complementing texts with visual displays can offer a number of cognitive and perceptual benefits, such as increasing working memory capacity, reducing cognitive load, and providing contexts, which lead to enhanced information recall, processing, and comprehension (Cook, 2006; Ginther, 2002; Lidwell et al., 2010; Mayer, 2001). In addition, research examining the effects of pro-environmental messages shows that infographics are more effective than text-based factual information in increasing audiences' message elaboration (Lazard & Atkinson, 2015). Visual messaging also has the potential to reduce literacy-based disparities in information processing and knowledge acquisition, as research indicates that visual messaging is more effective at improving information recall and understanding among people who have limited prior knowledge and literacy skills (ChanLin, 1998; Garcia-Retamero & Galesic, 2010; Meppelink, Smit, et al., 2015; Meppelink, van Weert, et al., 2015). Based on this analysis, we propose:

H5: Individuals exposed to visual information on COVID-19 vaccine safety will report higher factual knowledge about COVID-19 vaccine safety than individuals exposed to non-visual information.

H6: Individuals exposed to visual information on COVID-19 vaccine safety will report higher message elaboration than individuals exposed to non-visual information.

H7: The difference in factual knowledge about COVID-19 vaccine safety between individuals with higher and lower science literacy will be less pronounced when individuals are exposed to visual information than if they are exposed to non-visual information.

H8: The difference in message elaboration between individuals with higher and lower science literacy will be less pronounced when individuals are exposed to visual information than if they are exposed to non-visual information.

It is less clear whether visual narratives—a combination of visual and narrative messaging, such as comics, storyboards, and picture books—will be more effective than visual messages alone or narrative messages alone at enhancing message elaboration and knowledge acquisition and at mitigating literacy-based inequalities in these outcomes. For example, research suggests that comics—a form of visual narratives—can facilitate message elaboration and knowledge acquisition due to advantages in their information modality and rhetorical mode. In terms of modality, comics can promote information processing and understanding in various health contexts more effectively than text-based materials (e.g., A.Gillies et al., 1990; Delp & Jones, 1996). The visual aspect about comics also make scientific concepts more concrete and approachable (Jee & Anggoro, 2012). In addition, the close juxtaposition of texts and images in comics may make the medium particularly conducive to comprehension of complex information and deeper learning (Ginns, 2006; Mayer, 2003). In terms of rhetorical mode, the narrative component of comics may foster comprehension of scientific information by supplying familiar structure to such information (Graesser et al., 2002; Martin & Brouwer, 1991) and improve information encoding and processing via relatable characters (Jee & Anggoro, 2012). Further, research indicates that the positive effects of comics on information processing can be especially pronounced among audiences with limited literacy skills or prior knowledge, less education, and low interest in scientific topics (Crawford, 2004; Delp & Jones, 1996; Hosler & Boomer, 2011). However, no empirical research has systematically examined whether the combined effect of visual modality and narrative mode will necessarily be greater than that of visual modality and of narrative mode alone. Therefore, we pose the following research questions:

RQ1: Does the incorporation of a visual enhance the effectiveness of narrative on reducing the difference in factual knowledge about COVID-19 vaccine safety between individuals with higher and lower science literacy?

RQ2: Does the incorporation of a visual enhance the effectiveness of narrative on reducing the difference in message elaboration between individuals with higher and lower science literacy?

Science literacy and health information processing

Although research has largely ignored the role of science literacy in shaping how people process health information and acquire health knowledge, we have some reason to believe that science literacy can improve knowledge acquisition from and cognitive elaboration of health information, such as in the case of the COVID-19 vaccines. To begin with, an understanding of science (e.g., scientific facts, processes, and practices) can be highly beneficial in evaluating scientific evidence related to medical advancements that may involve high uncertainty (Sharon & Baram-Tsabari, 2020). This understanding also helps individuals interact with sources of scientific expertise, including that in the health realm (Feinstein, 2011). For example, research has shown that some level of science literacy can mitigate the effects of health-related misinformation (He et al., 2021). In addition, individuals with higher science literacy tend to process science and technology information in a deeper manner (Yang et al., 2023). We therefore examine the following hypotheses:

H9: Levels of science literacy will be positively related to factual knowledge about COVID-19 vaccine safety.

H10: Levels of science literacy will be positively related to elaboration of COVID-19 vaccine safety information.

Methods

Participants

Participants (*N* = 328 U.S. adults aged 18 or older) were recruited through online panels provided by Forthright between January 7 and 21, 2022. Forthright is an online research panel available through the Bovitz, Inc., a marketing services and strategy agency. Panelists are recruited with both online and offline channels, including digital networks and mail campaigns via address-based probability sampling methods. Participants received $2 in cash as compensation upon completion of surveys.

As we purposefully oversampled Black Americans, approximately half of the sample self-identified as Black American (50.3%) and the other half self-identified as White (49.7%). The sample's average age was 43 years, with 62.8% being females, and 35.1% had finished high school or less. Because we focused on vaccine-hesitant populations, individuals who already received all the shots needed to be fully vaccinated were excluded from the study. Among the participants, 15.9% had received only one shot of COVID-19 vaccine, and the rest had received no shots at all. When asked about the primary reason for not getting fully vaccinated, over one-third of participants reported having "concerns about the long-term side effects of COVID-19 vaccines;" 13.3% were concerned about the immediate side effects of COVID-19 vaccines. Concerns about "the fact that they are the fastest developed vaccines" and lack of trust in governmental agencies were also frequently cited themes.

Procedures

The study followed a 2 information modality (visual vs. non-visual/textual) × 2 rhetorical mode (narrative vs. non-narrative/logical-scientific) design for the message treatment to investigate the influence of communication formats on individuals' factual knowledge about COVID-19 vaccine safety and their message elaboration. Once consenting to the study, participants were randomly assigned to one of four message conditions. When starting the online survey, participants first answered a series of questions regarding their experiences with the pandemic and COVID-19 vaccines, their perceived norms to get vaccinated, their attitudes toward COVID-19 vaccines, their science literacy levels, as well as trust in governmental agencies that address the COVID pandemic. Factual knowledge about COVID-19 vaccine safety and message elaboration were measured after exposure to the assigned stimuli with question order being randomized (measures are described below). Finally, participants' general interest in reading comics, demographics, and political ideology were measured at the end of the questionnaire. The survey questionnaire was developed by the researchers and pre-tested with a group of undergraduate students ($N = 346$) at a large public university in the American Midwest.

Stimuli

We created four message conditions, including a visual narrative (comic) condition, a visual non-narrative (infographic) condition, a non-visual narrative (written story) condition, and a non-visual non-narrative (written facts) condition (see Appendix D). Stimuli for all conditions were embedded as images and displayed in the online questionnaire. To create the comics, the researchers first staged and photographed two college students on a university campus. The photos were then organized into a five-panel comic strip and used as a template for a hired illustrator to create the final drawn comics. The cartoon characters in the comics were inspired by the "counselors" concept in the 2020 Pixar movie *Soul*. Comprised of a living line, these characters were abstract, free of racial identity, yet entertaining and recognizable. The researchers also drafted the dialogue between two characters, interweaving scientific information regarding the COVID-19 vaccine safety posted on the CDC's website. Key information included: (1) the vaccines were developed using "science that has been around for decades;" (2) the vaccines were "tested in tens of thousands of people" before being administered to the public; (3) "at this point, nearly 70% of Americans already took it;" and (4) "serious side effects are rare." The wording of the key facts was kept consistent across all stimuli.

Measures

Dependent variables

We assessed factual knowledge about COVID-19 vaccine safety by asking four questions with a four-point scale (1 = "definitely false," 2 = "likely false," 3 = "likely true," 4 = "definitely true"). The items were "The science used for developing COVID-19 vaccines was created during the pandemic" (false), "Nearly 50% of Americans have received at least one dose of COVID-19 vaccines" (false), "The COVID-19 vaccines were tested on tens of thousands of people before they became available to the public" (true), and "Severe side effects, such as blood clots, are common after COVID-19 vaccination" (false). We recoded participants' answer to each question into a dichotomous variable (1= correct, 0 = incorrect). The sum of correct (range 0-4) was used as an index to measure factual knowledge about COVID-19 vaccine safety ($M = 2.0$, $SD = 1.2$).

Message elaboration of COVID-19 vaccine safety information was assessed by averaging five 7-point scale items (1 = "strongly disagree," 5 = "strongly agree") that asked participants while viewing the assigned stimuli, to what extent they were (a) attentive to the ideas, (b) deep in thought about the message, (c) unconcerned with the ideas (reverse coded), (d) distracted to other thoughts not related to the message (reverse coded), and (e) searching their mind in response to the ideas (M = 4.5, SD = 1.0, Cronbach's *alpha* = .64 for the comic condition, .78 for the infographic condition, .74 for the written story condition, and .65 for the written facts condition).

Moderating variables

Science literacy was measured by nine standardized questions contained in the National Science Foundation (NSF) science literacy scale (National Science Board, 2020). These questions measured both factual knowledge (e.g., "The center of the Earth is very hot," "All radioactivity is man-made") and process knowledge (e.g., the likelihood of getting inherited illness from parents). The questions were asked using a true/false scale; correct answers were counted and used as an indicator of science literacy (M = 4.5, SD = 2.3). A median split was then performed to separate individuals with high science literacy and individuals with low science literacy in the analysis (DeCoster et al., 2011).

Covariates

We run Analyses of Variance (ANOVA) for a list of variables that might potentially relate to the dependent variables, to assess potential differences between experimental groups. Results suggested that unequal distributions existed in education, income, and preexisting knowledge of COVID-19 vaccines between participants with higher and lower science literacy. We therefore included these variables as covariates to rule out any possibility that the detected group mean differences may not be attributed to the treatment.

Preexisting knowledge of COVID-19 vaccines was assessed with five true/false statements using 4-point scales (1 = "definitely false," 2 = "likely false," 3 = "likely true," 4 = "definitely true"). Specifically, these items include: (a) "Antibiotics can be used to kill the coronavirus" (false), (b) "You may get COVID-19 by receiving a vaccine shot" (false), (c) "People are considered fully vaccinated two weeks after their final dose of COVID-19 vaccines" (true), (d) "Having a food allergy means you should not get a COVID-19 vaccine" (false), and (e) "It is still unclear how long COVID-19 vaccines protect people" (true). Correct answers were coded as 1. False answers and "don't know" were coded as 0. Correct answers were summed up for each participant to form a single variable measuring preexisting COVID-19 vaccine knowledge, with a range of 0-5 ($M = 2.6$, $SD = 1.4$).

We also included demographic variables including education and income as covariates as there appeared to be unequal distribution in education and income between groups with different levels of science literacy. Education was measured by asking participants the highest degree received (61.9% some college or more). Income was measured by asking participants their annual household income (47.6% with $35,000 or more).

Analysis

We conducted two three-way Analyses of Covariance (ANCOVA) to calculate covariate-adjusted means using independent sample t-tests with Bonferroni correction for post-hoc comparisons. The dependent variables for the two ANCOVA models were factual knowledge about COVID-19 vaccine safety and message elaboration, respectively, while the fixed factors were information modality, information rhetorical mode, and one's level of science literacy.

Results

Summary statistics on key participant characteristics, the moderating variable, and the dependent variables can be found in Appendix E.

The role of rhetorical mode

H1 predicted that individuals who viewed a narrative message would report higher factual knowledge about COVID-19 vaccine safety than individuals who viewed a non-narrative message. The ANCOVA results (see Table 4.1 showed that the main effect of information rhetorical mode on factual knowledge was not statistically significant ($F_{1, 308}$ = .76, p = .38). Therefore, H1 was not supported. In terms of message elaboration, H2 hypothesized that individuals who viewed a narrative message would report greater message elaboration than individuals who viewed a non-narrative message. The ANCOVA results (see Table 4.2) showed that the main effect of information rhetorical mode on message elaboration was also not statistically significant ($F_{1, 308}$ = 1.57, p = .21). H2 was hence not supported.

Table 4.1. Analysis of covariance for factual knowledge about COVID-19 vaccine safety by treatment with covariates.

Source	Sum of Square	*df*	Mean Square	*F*	*p*	η^2
(Intercept)	43.55	1	43.55	34.83	.000	.10
Visual	.00	1	.00	.00	.999	.00
Narrative	.95	1	.95	.76	.384	.00
Science literacy	8.86	1	8.86	7.09	.008	.02
Education	.22	1	.22	.18	.676	.00
Income	1.87	1	1.87	1.49	.223	.01
Preexisting COVID-19 vaccine knowledge	32.84	1	32.84	26.27	.000	.08
Visual × Narrative	1.06	1	1.06	.85	.358	.00
Visual × Science literacy	4.93	1	4.93	3.95	.048	.01
Narrative × Science literacy	7.17	1	7.17	5.74	.017	.02
Visual × Narrative × Science literacy	1.47	1	1.47	1.17	.280	.00
Error	385.05	308	1.25			

Further, H3 predicted a less salient gap in factual knowledge about COVID-19 vaccine safety between individuals with higher and lower science literacy if they viewed narrative information on COVID-19 vaccine safety than if they viewed non-narrative information.

ANCOVA analysis (see Table 4.1) showed that the interaction term (narrative × science literacy) on factual knowledge about COVID-19 vaccine safety was significant ($F_{1, 308} = 5.74$, $p < .05$, $\eta^2 = .02$). Post-hoc comparison using the t-test with Bonferroni correction indicated that after viewing a non-narrative message about COVID-19 vaccine safety, participants with higher science literacy reported significantly higher factual knowledge about COVID-19 vaccine safety than those with lower science literacy; the difference in factual knowledge about COVID-19 vaccine safety between participants with higher and lower science literacy was statistically insignificant when participants instead viewed a narrative message about the topic (see Figure 4.1). Therefore, H3 was supported.

Table 4.2. Analysis of covariance for message elaboration of COVID-19 vaccine safety information by treatment with covariates.

Source	Sum of Square	*df*	Mean Square	*F*	*p*	η^2
(Intercept)	685.50	1	685.50	648.75	.000	.68
Visual	.24	1	.24	.227	.634	.00
Narrative	1.66	1	1.66	1.57	.212	.01
Science literacy	5.31	1	5.31	5.02	.026	.02
Education	.30	1	.30	.29	.593	.00
Income	.25	1	.25	.23	.630	.11
Preexisting COVID-19 vaccine knowledge	4.05	1	4.05	3.84	.051	.01
Visual × Narrative	.01	1	.01	.01	.912	.00
Visual × Science literacy	.82	1	.82	.77	.380	.00
Narrative × Science literacy	4.79	1	4.79	4.54	.034	.02
Visual × Narrative × Science literacy	.25	1	.25	.24	.625	.00
Error	325.45	308	1.06			

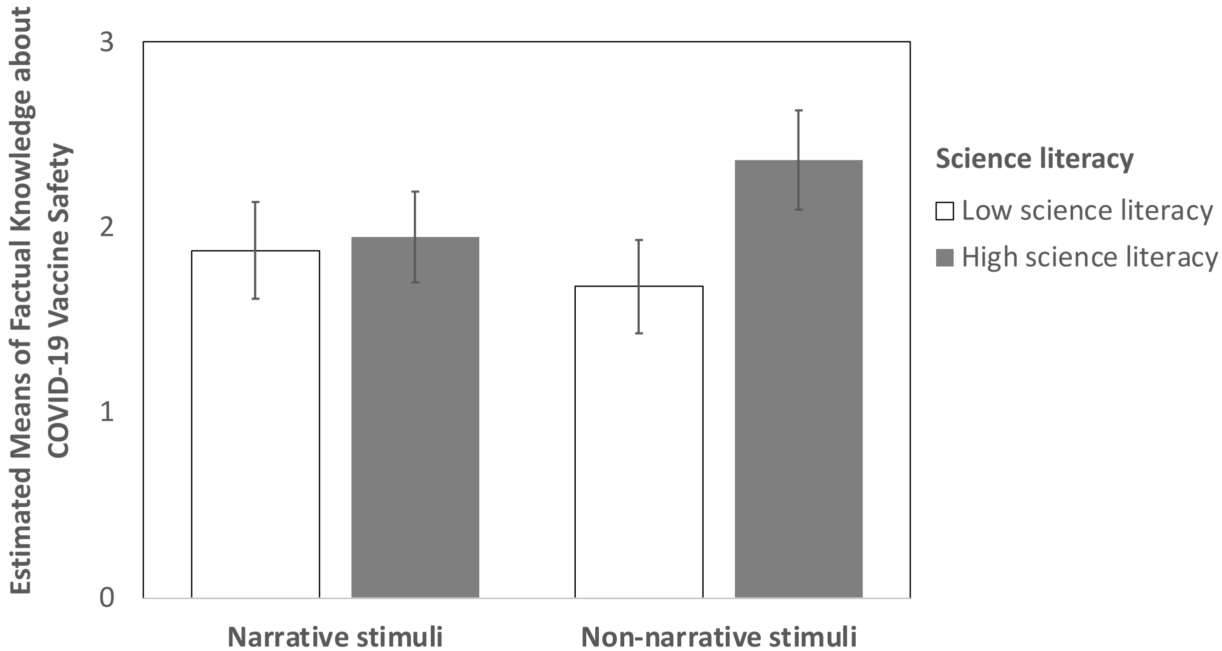


Figure 4.1. Difference in factual knowledge about COVID-19 vaccine safety between high- and low-science literacy groups by information rhetorical mode. Error bars are 95% confidence intervals.

H4 predicted that the difference in message elaboration between individuals with higher and lower science literacy would be less pronounced for those viewing narrative information on COVID-19 vaccine safety than for those viewing non-narrative information. ANCOVA results (see Table 4.2) showed that the interaction term (narrative × science literacy) was significant when predicting message elaboration ($F_{1, 308} = 4.54$, $p < .05$, $\eta^2 = .02$). Post-hoc comparison using the t-test with Bonferroni correction indicated that higher science literate participants reported significantly greater message elaboration than their less science literate counterparts when viewing a non-narrative message; however, such difference in message elaboration between participants with higher and lower science literacy became insignificant when they viewed a narrative message (see Figure 4.2). Therefore, H4 was supported.

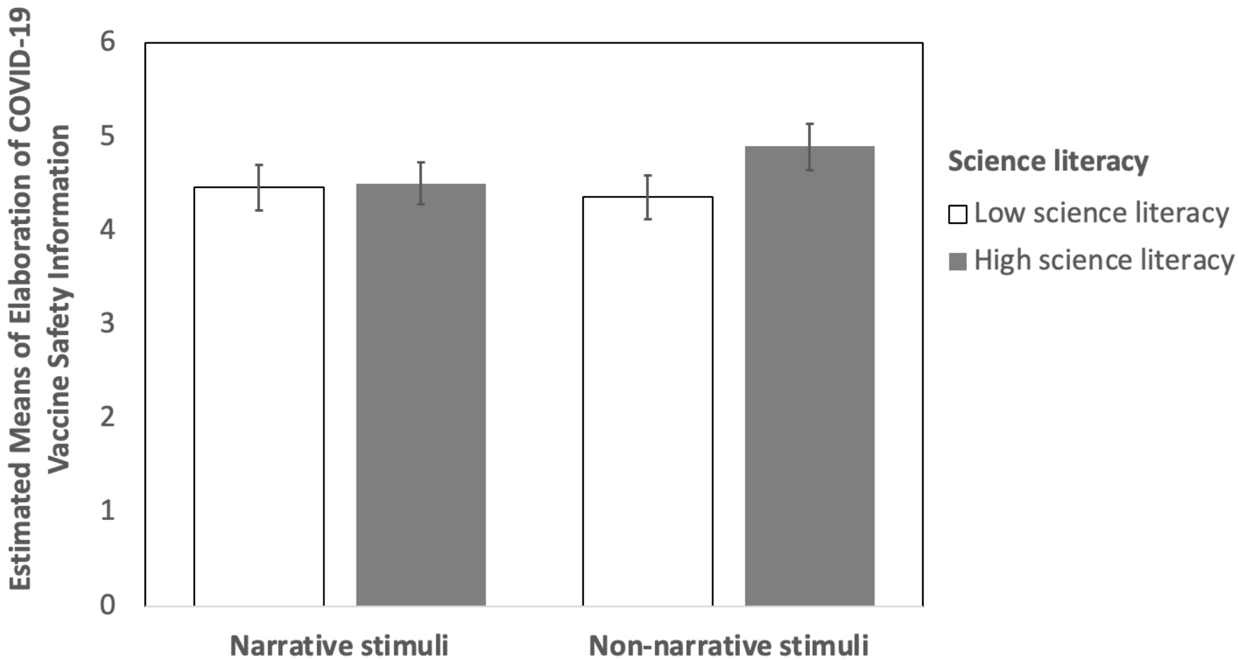


Figure 4.2. Difference in message elaboration of COVID-19 vaccine safety information between high- and low-science literacy groups by information rhetorical mode. Error bars are 95% confidence intervals.

The role of information modality

H5 hypothesized that individuals who viewed a visual message about COVID-19 vaccine safety would report higher factual knowledge about COVID-19 vaccine safety than individuals who viewed a non-visual message. The ANCOVA results (see Table 4.1) showed that the main effect of information modality on factual knowledge about COVID-19 vaccine safety was not statistically significant ($F_{1, 308} = .00$, $p = 1.00$). Therefore, H5 was not supported. In terms of message elaboration, H6 predicted that individuals who viewed a visual message about COVID-19 vaccine safety would report greater message elaboration than individuals who viewed a non-visual message. The ANCOVA results (see Table 4.2) showed that the main effect of information modality on message elaboration was also not statistically significant ($F_{1, 308} = .23$, $p = .63$). H6 was hence not supported.

Further, H7 predicted that the difference in factual knowledge about COVID-19 vaccine safety between individuals with higher science literacy and those with lower science literacy would be less pronounced when individuals were exposed to a visual message than if they were exposed to a non-visual message. An ANCOVA analysis (see Table 4.1) showed that the interaction term (visual × science literacy) was significant when predicting factual knowledge about COVID-19 vaccine safety ($F_{1, 308} = 3.95$, $p < .05$, $\eta^2 = .01$). Specifically, a post-hoc comparison using the t-test with Bonferroni correction indicated that after viewing a non-visual message about COVID-19 vaccine safety, participants with higher science literacy reported significantly higher factual knowledge about COVID-19 vaccine safety than participants with lower science literacy; however, such difference in factual knowledge between more and less science literate participants became insignificant when participants were instead exposed to a visual message about the topic (see Figure 4.3). Therefore, H7 was supported.

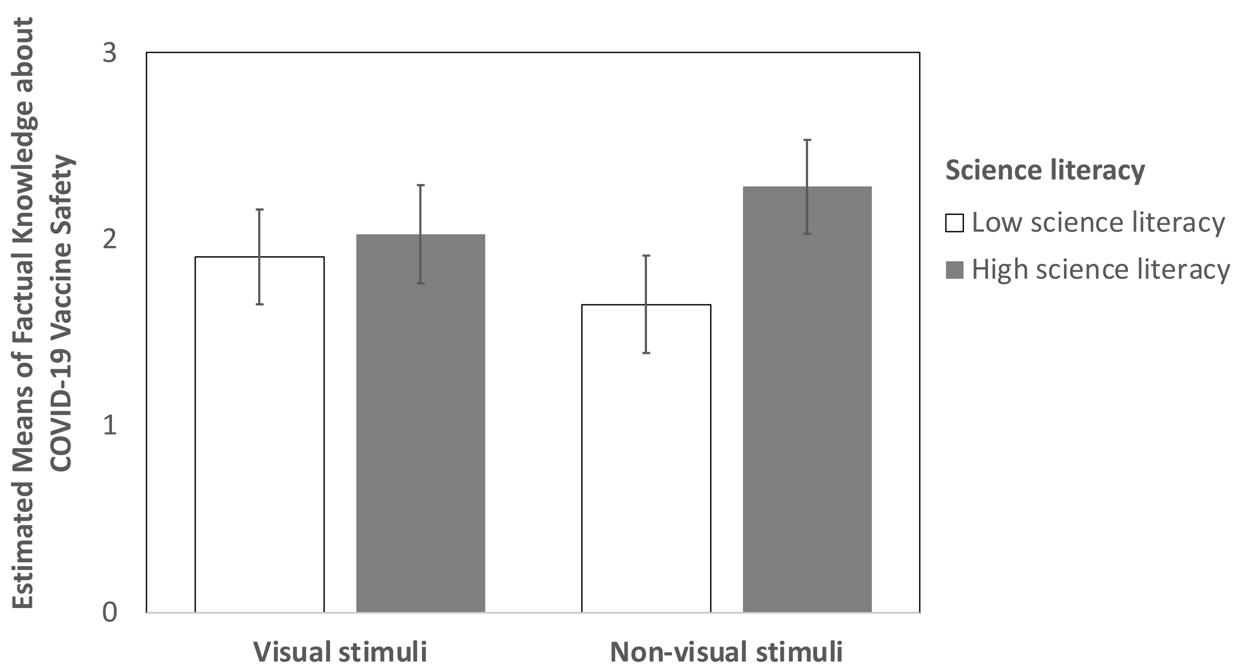


Figure 4.3. Difference in factual knowledge about COVID-19 vaccine safety between high- and low-science literacy groups by information modality. Error bars are 95% confidence intervals.

In addition, H8 predicted that the difference in message elaboration between individuals with higher and lower science literacy would be less pronounced for those viewing a visual message about COVID-19 vaccine safety than for those viewing a non-visual message. The ANCOVA results (see Table 4.2) showed that the interaction term (visual × science literacy) on message elaboration was not statistically significant ($F_{1, 308} = .77$, $p = .38$), indicating that exposure to a visual message did not significantly differentially impact message elaboration between participants with higher and lower levels of science literacy. Therefore, H8 was not supported.

Moreover, RQ1 asked whether the incorporation of visual elements would enhance the effectiveness of narrative on reducing the difference in factual knowledge about COVID-19 vaccine safety between individuals with higher and lower science literacy and RQ2 asked whether the incorporation of visual would enhance the effectiveness of narrative on reducing the difference in message elaboration between individuals with higher and lower science literacy. The three-way interaction (visual × narrative × science literacy) was also not statistically significant when predicting factual knowledge about COVID-19 vaccine safety ($F_{1, 308} = 1.17$, $p = .28$; see Table 4.1) and message elaboration ($F_{1, 308} = .24$, $p = .63$; see Table 4.2). Post-hoc comparison using the t-test with Bonferroni correction indicated that more highly science literate participants reported significantly higher factual knowledge about COVID-19 vaccine safety as well as message elaboration than less science literate participants only after viewing a non-visual non-narrative message, whereas such gaps were not statistically significant when participants received a message that contained either or both of visual and narrative. Overall, the insignificant three-way interactions suggested that the incorporation of visual elements did not significantly enhance the effectiveness of narrative on narrowing the gaps in factual knowledge about

COVID-19 vaccine safety and message elaboration between participants with higher and lower science literacy.

Main effects of science literacy

H9 hypothesized that science literacy would be positively related to individuals' factual knowledge about COVID-19 vaccine safety. Results from the ANCOVA analysis with factual knowledge about COVID-19 vaccine safety as the dependent variable (Table 4.1) showed that science literacy had a significant main effect on factual knowledge about COVID-19 vaccine safety ($F1, 308 = 7.09$, $p < .01$, $\eta2 = .02$). Participants with higher science literacy reported significantly higher factual knowledge about COVID-19 vaccine safety than those with lower science literacy. Therefore, H9 was supported.

Finally, H10 hypothesized that science literacy would be positively related to individuals' elaboration of message about COVID-19 vaccine safety. An ANCOVA analysis with message elaboration being the dependent variable (Table 4.2) indicated that science literacy had a significant main effect on message elaboration ($F1, 308 = 5.02$, $p < .05$, $\eta2 = .02$), with highly science literate participants reporting significantly greater message elaboration than their less science literate counterparts. H10 was hence supported.

Discussion

During the pandemic, visual storytelling was often used to explain how the COVID-19 vaccines work and why they are considerably safe and effective despite their rapid development. Although visual narratives have become increasingly popular in communicating about the COVID-19 vaccines and other health and science topics, empirical studies testing their effects—

especially quantitative analyses on large, adult samples—are rare (for an exception, see Shanahan et al., 2023). Further, to our knowledge no research has examined whether the potential effectiveness of visual storytelling should be attributed primarily to the presence of visuals or the integration of narratives. In addition, because visual narratives tend to be more accessible, visually engaging, and have stronger emotional appeal, they may help disadvantaged populations—such as those with limited literacy skills—more easily process the shown information, which could potentially lessen disparities in health and science knowledge across social segments of the American population. This possibility, however, remains largely underexplored. Moreover, much of prior research examining the effect of visual narratives on information processing and understanding took place in educational settings where students were exposed to visual narratives about mostly uncontroversial technical topics such as the formation of lightning or chemical structures and reactions, whereas we examine a highly debated topic in the American life that is relevant to broad segments of the population. Finally, research on health information processing and understanding has almost exclusively focused on the role of health literacy, whereas science literacy also likely matters, especially when the communicated medical developments involve fast-changing science and high scientific uncertainty.

Echoing calls for scholars and practitioners to make communication of health and science more equitable and inclusive, our study addresses important gaps in empirical research by conducting an experiment with a national sample of vaccine-hesitant White and Black American adults ($N = 328$) who received a series of customized messages depicting COVID-19 vaccine safety that varied in their levels of visual and narrative integration, including a comic strip, a written story, an infographic, and a short paragraph of written facts. The primary goals of our study were to (1) examine the relationship between science literacy and health information processing and understanding; (2) determine whether the primary factor contributing to the potential effectiveness of visual narratives in enhancing science information understanding is the presence of visuals (information modality) or the integration of narratives (rhetorical mode); and (3) investigate the potential of using visual and narrative messaging to catalyze more equitable outcomes of public health communication.

Before elaborating on our findings, we note several potential limiting factors related to the interpretation of our results. First, the data were collected amid the Omicron peak in the U.S.; the country reported 1.35 million COVID-19 cases in a day on January 10, 2022—three days into the data collection (Shumaker, 2022). The escalating situation might have increased participants' concern over COVID-19 and potentially made them more attentive to any form of information about vaccine safety. In other words, although we did not find a significant main effect of visual modality or narrative mode on people's factual knowledge about COVID-19 vaccine safety and message elaboration, it might be due to the heightened concerns and interest in the provided information.

In addition, much of prior research that identified a positive effect of visual narratives on information processing and understanding took place in educational settings and examined mostly uncontroversial technical topics, whereas we examine the COVID-19 vaccines, a highly debated topic in the American public life. The differences between our study and previous research in study sample and issue context may partly explain why we did not identify statistically significant effects of visual narratives. Future studies will need to replicate the study with a different context and population to examine the relative effectiveness of visual and narrative messaging, as well as their combinations.

Second, because we strived to preserve information consistency across experimental conditions while maximizing stimulus potency, it might have constrained the types of messaging strategies that we could incorporate into our stimuli. For example, research on the effects of visual messaging on health information processing has shown that the use of visual metaphors can lead to greater information elaboration (Lazard et al., 2016). The inclusion of awe-inspiring pictures in visual messages may also facilitate information engagement and understanding, given the centrality of awe-related imagery in science communication (Silva Luna et al., 2022). Whether a visual message is static or dynamic and its level of interactivity may also influence information processing and understanding (Mayer, 2003). Likewise, the extent to which informational content is integral to the causal structure of a narrative could matter for audience interpretation of the information (Dahlstrom, 2010). Longer narratives are also more likely to lead to greater information engagement (Slater & Rouner, 2006). While the purpose of our study was not to test all these visual and narrative features, future research would benefit from examining these and additional visual and narrative messaging strategies to determine under what conditions visual and narrative information promotes message elaboration and knowledge acquisitions across social segments.

Finally, we performed a median split on the science literacy variable to use it as a moderator of message effects on factual knowledge about COVID-19 vaccine safety and message elaboration. Future research may use science literacy as a continuous variable and analyze the message effects using regression analysis (DeCoster et al., 2011).

With these considerations in mind, our findings have important theoretical and practical implications for public communication on health topics. Whereas previous research has largely overlooked the role of science literacy in shaping people's interpretation of health information, our study illustrates the value of attending to science literacy when communicating novel health science developments. Specifically, our results show that science literacy is positively related to both message elaboration and factual knowledge about COVID-19 vaccine safety, as individuals who possessed higher science literacy reported greater message elaboration and factual knowledge about COVID-19 vaccine safety. Further, our study represents a first step toward understanding to what extent science literacy may enable people to accomplish specific tasks, such as learning from health and science messages (NASEM, 2016b). More research is needed to examine how science literacy may be instrumental to other health-related information processes and related outcomes, such as selecting health information sources, judging expertise and information credibility, and modifying health-related motivations and behaviors.

In terms of messaging strategies, we found no significant main effects of visual modality and narrative mode on individuals' message elaboration and factual knowledge about COVID-19 vaccine safety. The positive effects of visual and narrative messaging on message elaboration and knowledge acquisition—as identified in some of the previous research—may be conditional. Although exposure to our visual and narrative stimuli did not significantly enhance message elaboration and factual knowledge about COVID-19 vaccine safety among the overall sample, we did identify differential effects of visual and narrative messaging on individuals with different levels of science literacy. Specifically, after viewing a text-based COVID-19 vaccine safety message without any visual displays, less science literate Americans reported significantly lower factual knowledge about COVID-19 vaccine safety than more highly science literate Americans. However, among people who viewed a similar message incorporating visual displays, the difference in factual knowledge about COVID-19 vaccine safety between those with higher science literacy and those with lower science literacy significantly diminished. We also found a similar effect for narrative messaging. After exposure to a fact-based scientific message about COVID-19 vaccine safety that only contained descriptive facts, less science literate Americans reported a significantly lower factual knowledge about COVID-19 vaccine safety than more highly science literate Americans, whereas such science literacy-based gap in factual knowledge about COVID-19 vaccine safety did not exist among people who viewed a narrative message conveying the same information through characters and storytelling.

In addition to factual knowledge about COVID-19 vaccine safety, narrative messaging also appeared to mitigate differences in message elaboration between individuals with higher and lower science literacy. After viewing a non-narrative message about COVID-19 vaccine safety, highly science literate Americans processed the information in significantly greater depth than their less science literate counterparts. However, after viewing a narrative message on the same topic, people at different science literacy levels processed the narrative information to similar degrees. Further, although both visual and narrative messaging strategies appeared to be effective at narrowing the gap in factual knowledge about COVID-19 vaccine safety among individuals at different science literacy levels, combining the two strategies did not significantly further reduce the gap in message elaboration and factual knowledge about COVID-19 vaccine safety between more and less science literate groups.

The effect of visual and narrative messaging on narrowing the gap in factual knowledge about COVID-19 vaccine safety between science literacy groups suggests that there is value in incorporating either of these strategies into equity-based health and science messaging design. Integrating visual design into health messages or conveying health information through stories may be equally helpful for reaching diverse population segments—particularly subpopulations that lack science literacy and other cognitive resources—and promoting more equitable health communication outcomes. However, our findings indicate that the added value of combining the two strategies to form visual narratives about health topics may be limited, at least for the purpose of increasing factual science knowledge. In addition, the effect of narrative messaging on narrowing the gap in science information elaboration between science literacy groups—which was not found for visual messaging—indicates that visual and narrative messaging may affect science knowledge acquisition through distinct mechanisms. Whereas narrative messaging may foster science knowledge via extended reflection on and deeper processing of the information at hand, visual messaging does not seem to require the same amount of cognitive effort to produce improved knowledge. Because deeper information processing can lead to more stable persuasive outcomes (Eagly & Chaiken, 1993), our results may mean that narrative messaging will have longer-lasting effects on people than visual messaging, when all else equal. Alternatively, visuals may increase information accessibility via their ability to attract and hold viewer attention (Smerecnik et al., 2010), increase information memorability (Lidwell et al., 2010), and contextualize the presented information (Ginther, 2002). Moreover, because how people process health information and how well they understand such information can have implications for their subsequent health-related perceptions, attitudes, and behaviors, it would be worth examining how visual and narrative messaging strategies may affect long-term and short-term health-related perceptual and behavioral outcomes, as well as the disparities in these outcomes due to various structural and personal barriers. Overall, our study addresses the need to accelerate equity-centered public health communication related research, which will continue to matter even after the current crisis ends.

Chapter 5

Conclusion

Preference-based AI algorithms that exploit human cognitive and emotional weaknesses have become an indispensable underpinning of our information environments today. By tailoring information to people's preferences, biases, and contexts, these algorithms in online media could potentially amplify filter bubbles and echo chambers. Today, it is easier than ever for people to select their own content and avoid content they do not want. Consequently, people who are less interested in science will have less opportunity of exposure to science information and may therefore become more and more disconnected from science over time. What this means for the distribution of scientific information in society remains empirically understudied; further, given calls for more equitable and inclusive science communication research and practice, it is important to understand how the algorithmically infused media environment shapes the science information experience of diverse audiences, especially those who are traditionally underrepresented in and underserved by science. Additionally, algorithmic information curation may impact downstream outcomes at both individual and societal levels, such as individual cognitive and behavioral outcomes (e.g., knowledge of emerging science and technologies) and societal democratic decision-making related to science.

This dissertation addresses these challenges by (a) exploring how algorithmically driven social media platforms such as YouTube recommend science content based on users' racial and SES profiles, (b) examining how the use of different social media platforms shapes disparities in understanding of wicked science issues across racial and SES segments of the U.S. population, and (c) testing how message characteristics such as information modality and rhetorical mode could be leveraged against inequalities in scientific understanding.

Overview of findings

Broadly speaking, the purpose of my dissertation is to explore the media- and message-level factors that might affect diverse social segments' knowledge of wicked science issues in today's algorithmically infused media environment. Using three science issue (HGE, AI, and COVID-19 vaccines) as contexts of inquiry, my dissertation: (a) explores how a specific social media (YouTube) algorithm recommends science information to diverse audiences; (b) examines the relationship between the use of different social media platforms and race- and education-based science knowledge gaps across wicked science issues; and (c) investigates the effects of information modality and rhetorical mode on gaps in science information understanding and elaboration.

Study A uses a sock puppet experiment to audit the recommendation algorithm of YouTube. Results from the experiment indicate that as people search for science issues on YouTube, those who have higher SES are more likely to receive a greater diversity of search results recommendations from YouTube than audiences with lower SES. Sock puppets representing high-SES Americans, especially high-SES White Americans, encountered a wider range of YouTube videos (213 unique videos) and channels (173 unique channels) when searching for the three science issues, whereas sock puppets representing low-SES Americans, especially low-SES Blacks, encountered a less diverse collection of YouTube videos (187 unique videos) and channels (154 unique channels). Overall, a small majority (53%) of videos recommended by YouTube to the sock puppets as they searched for the three science issues were identical, regardless of the sock puppets' racial and SES categories; issue contexts appear to play a role here: for searches on AI, less than half (43%) of the video recommendations were identical across racial and SES audiences, whereas for searches on COVID-19 vaccines and HGE, an overwhelming majority (COVID-19 vaccines: 62%; HGE: 80%) of the video recommendations was identical across racial and SES audience groups.

In addition, by examining the semantic content of search results recommendations, Study A reveals the prevalent topics covered by the recommended content for each of the three science issues and across audience groups. Different racial and SES audiences may encounter different topics or similar topics but with different examples or emphasis in their search results recommendations. Specifically, for HGE-related video recommendations, sock puppets representing high-SES White Americans received additional content about the benefits of HGE that was not received by the other groups. For AI-related video recommendations, all sociodemographic groups except for the low-SES Black sock puppets received videos showcasing AI use, although the specific use cases might differ across groups. In addition, sock puppets representing Blacks (both high- and low-SES Blacks) also encountered additional videos discussing the problems or potential risks of AI that the White sock puppets did not receive. Finally, for video recommendations related to COVID-19 vaccines, public health updates were a consistent topic across most sociodemographic groups. Sock puppets representing High-SES Americans (both Whites and Blacks with high SES) encountered additional videos discussing COVID-19 vaccine misinformation and controversies that the low-SES sock puppets did not encounter, although the high-SES White sock puppets encountered misinformation-debunking videos whereas the high-SES Black sock puppets encountered misinformation-supporting videos.

Study B investigates how the use of algorithmically driven social media platforms including Facebook, Twitter, YouTube, Instagram, and TikTok may influence race- and education-based gaps in both factual and perceived knowledge of wicked science issues. Knowledge gaps were identified not in the sense of overtime change (Tichenor et al., 1970), but with interactions between education and social media use, wherein the strength of the positive relationship between issue knowledge and education is expected to vary depending on levels of social media use if a knowledge gap is present, with the relationship being stronger at high levels of media use and weaker at low levels of media use (Eveland & Scheufele, 2000; Kwak, 1999; McLeod et al., 1979).

Results from Study B indicate that social media use is associated with increased perceived, although not necessarily factual, knowledge of the wicked science issues under study. In terms of education-based factual knowledge gaps, increased Facebook use appeared to be associated with a widened gap in factual knowledge of AI between high and low education groups, particularly driven by a downtick in knowledge among low-education respondents; increased YouTube use was similarly related to an increased educated-based gap in factual knowledge of COVID-19, as high-education respondents gained more factual knowledge and low-education respondents reported less knowledge with increased YouTube use. In terms of race-based factual knowledge gaps, increased TikTok use was linked to a reduced gap in factual knowledge of AI between Whites and racial minorities, primarily driven by a decrease in Whites' knowledge. Additionally, as Instagram use increased, the race-based gap in factual knowledge of COVID-19 also decreased, with Whites expressing less knowledge and non-Whites more knowledge with increased Instagram use.

In terms of education-based perceived knowledge gaps, increased Facebook use was associated with a narrowed gap in perceived knowledge of AI between high and low education groups. Similarly, increased Instagram use was related to a decreased education-based gap in AI perceived knowledge, primarily due to low-education respondents reporting more perceived knowledge as they used Instagram more frequently. In terms of race-based perceived knowledge gaps, increased Facebook use was associated with a reduced gap in perceived knowledge of COVID-19 between Whites and non-Whites. Similarly, the race-based gap in perceived knowledge of COVID-19 narrowed as Instagram and YouTube use increased, with non-Whites growing COVID-19 perceived knowledge at a significantly faster rate than Whites. Twitter use seemed to be an exception as it widened the race-based gap in perceived knowledge of HGE, with non-Whites growing perceived knowledge significantly faster than Whites.

Moreover, I examined the intersectionality between race and education by running a series of three-way interactions between race, education, and social media platform use on both factual and perceived science knowledge. An important finding is that increased social media use overall predicted larger factual science knowledge gaps and smaller perceived science knowledge gaps between high and low education groups. Compared with more highly educated Americans, those with less education are less likely to gain factual science knowledge from increased social media use while they are more likely to gain perceived science knowledge. Racial minorities are more likely to gain both factual and perceived science knowledge with increased social media use than White Americans. Further, increased social media use was overall linked to wider factual science knowledge gaps and narrower perceived science knowledge gaps among racial minorities than among Whites.

Finally, consistent with theoretical expectations, the race- and education-based gaps in factual knowledge of COVID-19 were smaller than those of HGE and AI. The educated-based gap in perceived knowledge of AI was larger than those of HGE and COVID-19, whereas the race-based gap in perceived knowledge of HGE was larger than those of AI and COVID-19.

Addressing disparities in scientific understanding requires changes at various system levels. Focusing on equity-based message-level strategies, Study C investigates how information modality and rhetorical mode could affect message elaboration and factual knowledge about COVID-19 vaccine safety among individuals with varying levels of science literacy. We found no significant main effects of visual modality and narrative mode on individuals' message elaboration and factual knowledge about COVID-19 vaccine safety. However, we did identify differential effects of visual and narrative messaging on individuals with different levels of science literacy. Specifically, after viewing a text-based COVID-19 vaccine safety message without any visual displays, less science literate Americans reported significantly lower factual knowledge about COVID-19 vaccine safety than more highly science literate Americans. Among people who viewed a similar message incorporating visual displays, the difference in information understanding between those with higher science literacy and those with lower science literacy significantly diminished.

Narrative messaging also has a similar effect on narrowing the gaps formed on the basis of science literacy. After exposure to a fact-based scientific message about COVID-19 vaccine safety that only contained descriptive facts, less science literate Americans reported a significantly lower factual knowledge about COVID-19 vaccine safety than more highly science literate Americans, whereas such science literacy-based gap in factual knowledge about COVID-19 vaccine safety did not exist among people who viewed a narrative message conveying the same information through characters and storytelling. In addition to factual science knowledge, narrative messaging also appeared to mitigate the difference in message elaboration between individuals with higher and lower science literacy. After viewing a non-narrative message about COVID-19 vaccine safety, highly science literate Americans processed the information in significantly greater depth than their less science literate counterparts. However, after viewing a narrative message on the same topic, people at different science literacy levels processed the narrative information to similar degrees.

Further, although both visual and narrative messaging strategies appeared to be effective at narrowing the gap in factual knowledge about COVID-19 vaccine safety among individuals at different science literacy levels, combining the two strategies did not significantly further reduce the gaps in message elaboration and factual knowledge about COVID-19 vaccine safety between more and less science literate groups. Finally, science literacy is positively related to both message elaboration and factual knowledge about COVID-19 vaccine safety, as individuals who possessed higher science literacy reported greater message elaboration and factual knowledge about COVID-19 vaccine safety.

Theoretical implications

First, this dissertation provides an empirical understanding of how social media distribute science information in society. Findings from this dissertation suggest that social media algorithms, such as the YouTube algorithm, can indeed expose sociodemographically diverse audiences to different subsets of information even when people are actively searching for the same science issue, although the degree of information tailoring may depend on the search topic itself (e.g., how heavily the science issue is discussed by different sources on YouTube). In addition, higher-SES audiences, especially higher-SES White American audiences, are likely to receive a wider range of video and channel recommendations when searching for science issues than lower-SES audiences. Although low-SES audiences already tend to be informationally disadvantaged, the algorithmically infused media environment may expose them to a narrower range of science information than it does to their higher-SES, informationally resourceful peers, thus exacerbating disparities in science information and understanding. While disparities in scientific understanding between high- and low-SES social segments may be attributed to differences in learning ability or to the human tendency to select attitude-confirming information (Tichenor et al., 1970), findings from this dissertation suggest that such disparities may also be caused by algorithms that target human preferences and expose population segments to different subsets of information without their own free will. In other words, what we are experiencing might be a 3.0 version of the knowledge gap concept (Tichenor et al., 1970), whereby knowledge gaps are formed not only on the basis of social segments viewing the same information but also, and perhaps more importantly, on the basis of social segments receiving different information by algorithmic tailoring. However, no existing media effects models (see Cacciatore et al., 2016) can satisfactorily account for this reality.

Further, concerns over the ethics and equity of digital platforms are not just unique to social media. For example, research has also found that Google News could potentially reinforce users' partisanship by personalizing news search results based on users' browsing history (Le et al., 2019). With the advent of ChatGPT and other large, multimodal language models that will bring disruptive impacts on how we communicate and consume information, it is necessary for future research to continue to examine existing and emerging platforms and tools so that we are better prepared to pre-empt or counter structural barriers in our information environments that prevent equitable distribution of scientific information in society.

Second, empirical evidence suggests that there are wide educational and racial differences in Americans' knowledge about science, with more educated Americans and Whites scoring higher on science knowledge measures than less educated Americans and Blacks and Hispanics (Kennedy & Atske, 2019; Kennedy & Hefferon, 2019; see Figures 1.2 through 1.4).

Even though science knowledge has limited influence on science-related perceptions, attitudes and behavior (Allum et al., 2008; NASEM, 2016b), it still matters that all population segments have equitable access to science knowledge so as to effectively engage in democratic decision-making on policy issues involving science. This dissertation takes a granular approach to social media use by distinguishing the specific social media platforms and examining the relationship between their frequency of use and science knowledge held by different social segments of the American population. Although research on social media and science knowledge gaps has for the most part examined aggregate social media use, our findings suggest that there is value in differentiating specific social media platforms because use of different platforms may not shape knowledge (gaps) in the same way. For example, whereas frequent Twitter users consistently reported higher factual knowledge on all three science issues even after controlling for demographics, frequent Facebook users expressed less factual knowledge about HGE and frequent users of YouTube, Instagram, and TikTok did not necessarily possess higher or lower factual science knowledge than people who used those platforms less often. Such differences in platform use's influence on factual science knowledge may in part be attributed to differences in the types of information sources that people pay attention to on these platforms. While users of Twitter tend to seek out mainstream news when they are on the platform, potentially contributing to the growth in their factual knowledge, users of Instagram, TikTok, YouTube, and Facebook tend to pay attention to a mix of Internet personalities, ordinary people, and news (Newman et al., 2021), which may not be the types of information sources that are most conducive to factual science knowledge acquisition. Future research could focus on examining how social media platforms' features, attributes, and use patterns could shape science knowledge gaps, to develop a deeper theoretical understanding of the science knowledge gap phenomenon in the social media age and to facilitate practical efforts to mitigate such inequalities.

Third, consistent with Ladwig et al (2012) who found that factual nanotechnology knowledge and perceived nanotechnology knowledge were only slightly correlated with each other and were predicted differently by media use and cognitive processing variables, this dissertation also shows that the identification of knowledge gaps depends on how knowledge is conceptualized and measured. Scholars should be careful to not conflate factual knowledge measures with self-reported knowledge measures when examining science knowledge gaps. For instance, whereas social media use seemed to overall widen education-based gaps in factual science knowledge, it appeared to narrow education-based gaps in perceived knowledge. When high-education individuals used YouTube and Facebook more, they gained more factual knowledge about COVID-19 and AI, respectively; in contrast, when low-education individuals used these platforms more, they did not acquire as much knowledge as their high-education counterparts (in fact, they showed a downtick in factual knowledge with increased social media use), subsequently widening the education-based gaps in factual knowledge. However, as low-education individuals used Facebook and Instagram more, their perceived knowledge of AI increased at a significantly faster rate than that of their high-education counterparts, leading to reduced perceived knowledge gaps between education groups.

Fourth, the extent to which disparities in scientific understanding are identified could also depend on the science issue context at hand. Previous research has suggested that knowledge gaps are contingent upon issue characteristics such as issue complexity (and thus knowledge complexity), controversy, issue appeal to a social system, and media publicity (Bauer & Bonfadelli, 2002; Donohue et al., 1975; Moore, 1987). Consistent with theoretical expectations, this dissertation identified smaller education- and race-based factual knowledge gaps for COVID-19 than for AI and HGE, arguably because COVID-19 generated much more media and policy attention, public concern, and political controversy, all of which led to increased information flows across communities and all walks of life. Additionally, these science issues differ in their scope, temporality, and risks involved. While COVID-19 has a constrained issue scope and is relatively less "wicked", HGE and especially AI have a very broad range of applications touching many aspects of society. Whereas COVID-19 is relatively transient and involves highly tangible, personal risks to oneself, the impacts of HGE and AI are more long-term, latent, and far-reaching at the societal level. These and additional issue characteristics may also affect people's concerns, salience, and self-efficacy regarding science issues and the way they acquire knowledge (Ettema et al., 1983; Shim, 2008). More systematic research is needed to examine what and how science issue characteristics might matter for knowledge gaps.

Finally, this dissertation provides evidence that information modality and rhetorical mode may facilitate science knowledge acquisition among less science literate individuals through different mechanisms: although both visual message and narrative message could enhance factual knowledge about COVID-19 vaccine safety among this group, narrative message also appears to increase message elaboration whereas visual message does not significantly impact message elaboration. This means that science information packaged as a narrative will likely instigate deeper cognitive involvement, which could lead to greater science knowledge. In contrast, science information communicated primarily via visuals may increase science knowledge without necessarily catalyzing greater mental elaboration with the science information at hand.

Methodological implications

This dissertation also makes several methodological contributions. First, it uses a novel computational method, sock puppet algorithm audit design, to measure and study the opaque algorithmic influences of the YouTube social media platform. In this dissertation, a sock puppet was implemented as an automated web browser instance that interacted with YouTube by searching, watching videos, and recording video recommendations in YouTube. The sock puppet audit approach has a number of methodological advantages. First, it prevents the issue of infringing user privacy as it does not involve tracking real users' digital or web behaviors that may reveal private information about the users. Second, a sock puppet audit design enables researchers to more effectively investigate sensitive topic domains such as misinformation consumption on digital platforms (e.g., Hussein et al., 2020; Srba et al., 2023); it also allows researchers to assign sock puppets to categories that are normatively important but difficult to talk about, such as groups living in poverty, sexual minorities, and groups with stigmatized health conditions (Sandvig et al., 2014). Last but not least, because a sock puppet audit design is a form of field experiment, it allows for strong causal inference by experimental manipulation and preserves external validity by occurring in the natural field.

To develop a research agenda of using sock puppet algorithm audit for future science communication and, more generally, social science research, several important methodological issues should be considered. One issue is determining the optimal number of sock puppets to employ in an experiment. As alluded to earlier, existing studies employing a sock puppet algorithm audit design have used vastly different numbers of sock puppets, ranging from less than 10 to less than a hundred to tens of thousands of sock puppets (e.g., Haroon et al., 2022; Hussein et al., 2020; Srba et al., 2023). Researchers rarely justify or explain their choice, and it is unclear what the decision rules (should) look like for determining the optimal number of sock puppets. The number of sock puppets allowed in a study obviously depends on how the sock puppets are implemented, as different implementations require different amounts and types of resources. For example, it is easier to create sock puppets automatically on a large scale when sock puppets are implemented as automated web browser instances as compared to platform user accounts that often require additional verification measures to create or log into. Besides practical concerns, however, the choice of the number of sock puppets should also be theoretically and methodologically sound. Future research may need to compare results and effect sizes obtained from experiments done with sock puppets and with human subjects to understand how the sample size of sock puppets match up the sample size of human subjects.

Another important consideration of using sock puppet experiment for science communication and social science research is the issue of external validity. Researchers may use methodological triangulation to validate the sock puppet training process as well as the experimental results. In addition, the issue of construct validity also deserves further attention. Specifically, researcher may want to measure the effects of each specific training activity on the outcomes of interest to pinpoint the specific sources of effects.

Second, this dissertation trained the sock puppets by creatively using psychographic profiling, a method commonly used in advertising and marketing research that helps advertisers gain an understanding of their audience segments. Psychographic information reveals deeper details about members of the target audiences, such as their interests, lifestyle, habits, attitudes, and behaviors (Meredith, 2021); it can also inform researchers how to reach their target audiences. To conduct systematic psychographic profiling on the sociodemographic groups of interest, this dissertation uses the Simmons Insights database (MRI-Simmons, 2018), one of the most comprehensive databases on American consumers that provides access to a wide range of demographic and psychographic consumer data. Using these methodologies and data helped increase the validity of the sock puppet training process.

A third methodological contribution relates to the examination of the relative effectiveness of visual versus narrative components of science visual narratives. We also created customized visual narrative materials (e.g., comic) for our research purposes. For example, we created the comic by first staging and photographing two college students on a university campus, and then organizing the photos into a five-panel comic strip, which was subsequently used as a template for a hired professional illustrator to create the final drawn comics. The cartoon characters in the comics were inspired by the “counselors” concept in the 2020 Pixar movie *Soul*. Comprised of a living line, these characters were abstract, free of racial identity, yet entertaining and recognizable.

Practical implications

Overall, this dissertation provides implications for facilitating greater use of equity-based communications strategies toward a more equitable science information society. To begin with, this dissertation demonstrates the value of understanding the online information environments wherein we communicate science today. Only after recognizing the barriers in our information environments (be it social media or other old and new platforms) that prevent more equitable distribution of science information across society can we start to prepare for addressing those challenges. Although directly changing the social media platforms or their underlying algorithms is infeasible, actions at both societal and individual levels can still take place to mitigate the negative algorithmic influences on science communication equity. In addition, science communication practitioners who truly want to reach a broad audience beyond the proverbial choir of those who are already science-friendly or curious may need to find additional communication pathways, rather than leaving it all alone to social media and hoping for the best.

Moreover, when it comes to science knowledge acquisition, social media use could differentially benefit different sociodemographic segments. Looking at the intersectionality of education and race, this dissertation identifies overall more severe adverse effects of social media use on factual science knowledge gaps formed on the basis of education among racial minorities than among Whites. Specifically, less educated Whites are able to catch up with more educated Whites in their HGE factual knowledge as they used Twitter more, leading to a reduced knowledge gap between the two groups. However, the same is not true for racial minorities. Less educated non-Whites learn virtually nothing more about HGE when increasing their Twitter use, whereas more educated non-Whites acquire significantly more HGE knowledge than their less educated counterparts when Twitter use increases, widening the knowledge gap between the groups. Consistent with these patterns, increased YouTube use had virtually no impact on the gap in COVID-19 factual knowledge between high- and low-education Whites, but it seems to increase the knowledge gap between high- and low-education non-Whites, with more educated non-Whites acquiring COVID-19 factual knowledge at a faster rate than less educated non-Whites. These findings together suggest that knowledge gaps can be multifaceted phenomena that warrant attention to the intersectionality of sociodemographic influences. More importantly, special efforts may be made to support less educated and low-SES racial minorities so that they could equally benefit from advancements in information technologies as these technologies (e.g., social media) have become increasingly integrated with our life. Such efforts may include helping low-education minority segments develop science literacy (Howell & Brossard, 2021) as well as social media literacy (Cho et al., 2022), as well as identifying alternative channels for engaging these segments with science.

Finally, this dissertation identifies that both visual and narrative messaging can narrow the gaps in factual science knowledge between science literacy groups, which indicates that there is value in incorporating either of these messaging strategies into equity-based health and science information design. In other words, integrating visual design into health messages or conveying health information through stories may be equally helpful for reaching diverse population segments—particularly subpopulations that lack science literacy and other cognitive resources—and promoting more equitable health communication outcomes. However, the added value of combining the two strategies to form visual narratives about health and science topics may be limited. In addition, the effect of narrative messaging on narrowing the science literacy-based gap in information elaboration—which was not found for visual messaging—indicates that visual and narrative messaging may affect science knowledge acquisition through distinct mechanisms. Whereas narrative messaging may foster science knowledge via extended reflection on and deeper processing of the information at hand, visual messaging does not seem to require the same amount of cognitive effort to produce improved science knowledge. Because deeper information processing can lead to more stable persuasive outcomes (Eagly & Chaiken, 1993), our results may mean that narrative messing will have longer-lasting effects on people than visual messaging, when all else equal. Taken together, science and health communication practitioners may leverage narrative or visual strategies in their message design and disseminate the information in communication channels relevant to their audiences to maximize the positive impacts.

References

A.Gillies, P., Stork, A., & Bretman, M. (1990). Streetwize UK: A controlled trial of an AIDS education comic. *Health Education Research*, *5*(1), 27-33.

Abbas, M., Riaz, M. U., Rauf, A., Khan, M. T., & Khalid, S. (2017). Context-aware YouTube recommender system. In *2017 international conference on information and communication technologies (ICICT)* (pp. 161-164). IEEE. https://doi.org/10.1109/ICICT.2017.8320183

Agardh, E., Allebeck, P., Hallqvist, J., Moradi, T., & Sidorchuk, A. (2011). Type 2 diabetes incidence and socio-economic position: A systematic review and meta-analysis. *International Journal of Epidemiology*, *40*(3), 804-818. https://doi.org/10.1093/ije/dyr029

Akin, H., Cacciatore, M. A., Yeo, S. K., Brossard, D., Scheufele, D. A., & Xenos, M. A. (2020). Publics' support for novel and established science issues linked to perceived knowledge and deference to science. *International Journal of Public Opinion Research*, *33*(2), 422-431. https://doi.org/10.1093/ijpor/edaa010

Akin, H., Hardy, B. W., Scheufele, D. A., & Brossard, D. (2017, April 25). Can Bill Nye – or any other science show – really save the world? *The Conversation*. https://theconversation.com/can-bill-nye-or-any-other-science-show-really-save-the-world-76630

Allum, N., Sturgis, P., Tabourazi, D., & Brunton-Smith, I. (2008). Science knowledge and attitudes across cultures: A meta-analysis. *Public Understanding of Science, 17*(1), 35-54. https://doi.org/10.1177/0963662506070159

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The "nasty effect:" Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, *19*(3), 373-387. https://doi.org/10.1111/jcc4.12009

Anderson, M. (2023, May 1). *After Musk's takeover, big shifts in how Republican and Democratic Twitter users view the platform*. Pew Research Center. https://www.pewresearch.org/short-reads/2023/05/01/after-musks-takeover-big-shifts-in-how-republican-and-democratic-twitter-users-view-the-platform/

Auxier, B., & Anderson, M. (2021, April 7). *Social media use in 2021*. Pew Research Center. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/

Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: COVID-induced online learning engagement in real time. *Journal of Public Economics*, *193*, 104345. https://doi.org/10.1016/j.jpubeco.2020.104345

Bandy, J., & Diakopoulos, N. (2021). Curating quality? How Twitter's timeline algorithm treats different types of news. *Social Media + Society*, *7*(3), 20563051211041648. https://doi.org/10.1177/20563051211041648

Barthel, M., Grieco, E., & Shearer, E. (2019, August 14). *Older Americans, Black adults and Americans with less education more interested in local news*. Pew Research Center. https://www.pewresearch.org/journalism/2019/08/14/older-americans-black-adults-and-americans-with-less-education-more-interested-in-local-news/

Bauer, M. W., & Bonfadelli, H. (2002). Controversy, media coverage and public knowledge. In M. W. Bauer & G. Gaskell (Eds.), *Biotechnology – the making of a global controversy* (pp. 149-175). Cambridge University Press.

Bennett, W. L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication, 58*(4), 707-731. https://doi.org/10.1111/j.1460-2466.2008.00410.x

Bonfadelli, H. (2002). The Internet and knowledge gaps: A theoretical and empirical investigation. *European Journal of Communication*, *17*(1), 65-84. https://doi.org/10.1177/0267323102017001607

Borghol, Y., Ardon, S., Carlsson, N., Eager, D., & Mahanti, A. (2012). The untold story of the clones: Content-agnostic factors that impact YouTube video popularity. In *Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1186-1194). Association for Computing Machinery. https://doi.org/10.1145/2339530.2339717

Brenan, M. (2021, August 24). *U.S. satisfaction drops; COVID-19 resurges as top problem*. Gallup. https://news.gallup.com/poll/354026/satisfaction-drops-covid-resurges-top-problem.aspx

Brossard, D. (2013). New media landscapes and the science information consumer. *Proceedings of the National Academy of Sciences*, *110*(Supplement 3), 14096-14101. https://doi.org/10.1073/pnas.1212744110

Brossard, D., & Scheufele, D. A. (2022). The chronic growing pains of communicating science online. *Science*, *375*(6581), 613-614. https://doi.org/10.1126/science.abo0668

Brossard, D., Scheufele, D. A., Kim, E., & Lewenstein, B. V. (2009). Religiosity as a perceptual filter: Examining processes of opinion formation about nanotechnology. *Public Understanding of Science*, *18*(5), 546-558. https://doi.org/10.1177/0963662507087304

Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial intelligence and journalism. *Journalism & Mass Communication Quarterly*, *96*(3), 673-695. https://doi.org/10.1177/1077699019859901

Cacciatore, M. A., Scheufele, D. A., & Corley, E. A. (2014). Another (methodological) look at knowledge gaps and the Internet's potential for closing them. *Public Understanding of Science*, *23*(4), 376-394. https://doi.org/10.1177/0963662512447606

Cacciatore, M. A., Scheufele, D. A., & Iyengar, S. (2016). The end of framing as we know it… and the future of media effects. *Mass Communication and Society, 19*(1), 7-23. https://doi.org/10.1080/15205436.2015.1068811

Callegaro, M., & DiSogra, C. (2009). Computing response metrics for online panels. *Public Opinion Quarterly*, *72*(5), 1008-1032. https://doi.org/10.1093/poq/nfn065

Center for Community Media. (2020, November 30). *Mapping black media*. Craig Newmark Graduate School of Journalism, The City University of New York. https://www.journalism.cuny.edu/2020/11/mapping-black-media/

Centers for Disease Control and Prevention. (2023a, March 7). *Safety of COVID-19 vaccines*. https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/safety-of-vaccines.html

Centers for Disease Control and Prevention. (2023b, March 21). *COVID-19 vaccines recommendation for older adults* [Video]. YouTube. https://www.youtube.com/watch?v=F3Ex62wC9gw

Centers for Disease Control and Prevention. (2023c, May 11). *COVID-19: Your health*. https://www.cdc.gov/coronavirus/2019-ncov/your-health/index.html

Centers for Disease Control and Prevention. (2023d, May 12). *Understanding how COVID-19 vaccines work*. https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/how-they-work.html

Chang, J. H., Kim, S. H., Kang, M. H., Shim, J. C., & Ma, D. H. (2018). The gap in scientific knowledge and role of science communication in South Korea. *Public Understanding of Science*, *27*(5), 578-593. https://doi.org/10.1177/0963662516685487

ChanLin, L.-J. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology*, *25*(3), 166. https://ezproxy.library.wisc.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=aph&AN=1075971&site=ehost-live&scope=site

Chen, K., Bao, L., Shao, A., Ho, P., Yang, S., Wirz, C. D., Brossard, D., Brauer, M., & DiPrete Brown, L. (2020). How public perceptions of social distancing evolved over a critical time period: Communication lessons learnt from the American state of Wisconsin. *Journal of Science Communication*, *19*(5), A11. https://doi.org/10.22323/2.19050211

Cho, H., Cannon, J., Lopez, R., & Li, W. (2022). Social media literacy: A conceptual framework. *New Media & Society*, *0*(0), 14614448211068530. https://doi.org/10.1177/14614448211068530

Christopherson, E. G., Howell, E. L., Scheufele, D. A., Viswanath, K., & West, N. P. (2021). How science philanthropy can build equity. *Stanford Social Innovation Review*, *19*(4), 48-55. https://doi.org/10.48558/P4G8-QM77

Christopherson, E. G., Scheufele, D. A., & Smith, B. (2018). The civic science imperative. *Stanford Social Innovation Review*, *16*(2), 46-52. https://doi.org/10.48558/6K46-1C68

Cid, R., & Bolívar, J. (2021). Platforms for production of protein-based vaccines: From classical to next-generation strategies. *Biomolecules*, *11*(8). https://doi.org/10.3390/biom11081072

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates Publishers.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, *90*(6), 1073-1091. https://doi.org/10.1002/sce.20164

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. In *Proceedings of the 10th ACM conference on recommender systems* (pp. 191-198). Association for Computing Machinery. https://doi.org/10.1145/2959100.2959190

Crawford, P. (2004). Using graphic novels to attract reluctant readers. *Library Media Connection*, *27*, 26-28.

Dahlstrom, M. F. (2010). The role of causality in information acceptance in narratives: An example from science communication. *Communication Research*, *37*(6), 857-875. https://doi.org/10.1177/0093650210362683

Dahlstrom, M. F. (2012). The persuasive influence of narrative causality: Psychological mechanism, strength in overcoming resistance, and persistence over time. *Media Psychology*, *15*(3), 303-326. https://doi.org/10.1080/15213269.2012.702604

Dahlstrom, M. F. (2014). Using narratives and storytelling to communicate science with nonexpert audiences. *Proceedings of the National Academy of Sciences*, *111*(Supplement 4), 13614-13620. https://doi.org/10.1073/pnas.1320645111

Dahlstrom, M. F. (2021). The narrative truth about scientific misinformation. *Proceedings of the National Academy of Sciences*, *118*(15). https://doi.org/10.1073/pnas.1914085117

Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y., Lambert, M., & Livingston, B. (2010). The YouTube video recommendation system. In *Proceedings of the fourth ACM conference on recommender systems* (pp. 293-296). Association for Computing Machinery. https://doi.org/10.1145/1864708.1864770

Davies, D. (2021, March 3). *Meet the 7 most popular search engines in the world*. Search Engine Journal. https://www.searchenginejournal.com/seo/meet-search-engines/#close

De Silva-Schmidt, F., Brüggemann, M., Hoppe, I., & Arlt, D. (2022). Learning about climate politics during COP 21: Explaining a diminishing knowledge gap. *Public Understanding of Science*, *31*(5), 617-633. https://doi.org/10.1177/09636625211068635

Dean, B. (2017). *We analyzed 1.3 million YouTube videos. Here's what we learned about YouTube SEO*. Backlinko. https://backlinko.com/youtube-ranking-factors

DeCoster, J., Gallucci, M., & Iselin, A. M. R. (2011). Best practices for using median splits, artificial categorization, and their continuous alternatives. *Journal of Experimental Psychopathology, 2*(2), 197-209. https://doi.org/10.5127/jep.008310

Delp, C., & Jones, J. (1996). Communicating information to patients: The use of cartoon illustrations to improve comprehension of instructions. *Academic Emergency Medicine*, *3*(3), 264-270. https://doi.org/10.1111/j.1553-2712.1996.tb03431.x

Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., Fletcher, J. M., & Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly*, *50*(4), 393-416. https://doi.org/10.1002/rrq.105

Donohue, G. A., Tichenor, P. J., & Olien, C. N. (1975). Mass media and the knowledge gap: A hypothesis reconsidered. *Communication Research*, *2*(1), 3-23. https://doi.org/10.1177/009365027500200101

Downs, A. (1972). Up and down with ecology: The issue-attention cycle. *The public*, 462-473.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Ettema, J. S., Brown, J. W., & Luepker, R. V. (1983). Knowledge gap effects in a health information campaign. *Public Opinion Quarterly*, *47*(4), 516-527. https://doi.org/10.1086/268809

European Commission. (1995). *White paper on education and training.* European Commission. https://op.europa.eu/en/publication-detail/-/publication/d0a8aa7a-5311-4eee-904c-98fa541108d8/language-en

Eveland, W. P., & Scheufele, D. A. (2000). Connecting news media use with gaps in knowledge and participation. *Political Communication*, *17*(3), 215-237. https://doi.org/10.1080/105846000414250

Fang, E., Liu, X., Li, M., Zhang, Z., Song, L., Zhu, B., Wu, X., Liu, J., Zhao, D., & Li, Y. (2022). Advances in COVID-19 mRNA vaccine development. *Signal Transduct Target Ther*, *7*(1), 94. https://doi.org/10.1038/s41392-022-00950-y

Feedspot. (2023, April 26). *100 science YouTube channels for science news, videos, research and scientific concepts*. Feedspot. https://blog.feedspot.com/science_youtube_channels/

Feinstein, N. (2011). Salvaging science literacy. *Science Education*, *95*(1), 168-185. https://doi.org/10.1002/sce.20414

Fisch, S. M. (2000). A capacity model of children's comprehension of educational content on television. *Media Psychology*, *2*(1), 63-91. https://doi.org/10.1207/S1532785XMEP0201_4

Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, *1*, 3-32. https://hdl.handle.net/2440/15169

Fisher, W. R. (1989). *Human communication as narration: Toward a philosophy of reason, value, and action*. University of South Carolina Press.

Fiske, S. T., & Taylor, S. E. (2008). *Social cognition: From brains to culture*. McGraw-Hill.

Forman-Katz, N., & Matsa, K. E. (2022, September 20). *News platform fact sheet*. Pew Research Center. https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/?menuItem=b39b851c-e417-48ef-9b10-93ee21a0030e

Funk, C., Gottfried, J., & Mitchell, A. (2017a, September 20). *Most Americans express curiosity in science news, but a minority are active science news consumers*. Pew Research Center. https://www.pewresearch.org/journalism/2017/09/20/most-americans-express-curiosity-in-science-news-but-a-minority-are-active-science-news-consumers/

Funk, C., Gottfried, J., & Mitchell, A. (2017b, September 20). *Science news and information today*. Pew Research Center. http://www.journalism.org/2017/09/20/science-news-and-information-today/

Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, *25*(7), 739-755. https://doi.org/10.1016/0016-3287(93)90022-L

Galston, W. A. (2021, October 1). *For COVID-19 vaccinations, party affiliation matters more than race and ethnicity*. Brookings Institution.

https://www.brookings.edu/blog/fixgov/2021/10/01/for-covid-19-vaccinations-party-affiliation-matters-more-than-race-and-ethnicity/

Garcia-Retamero, R., & Galesic, M. (2010). Who profits from visual aids: Overcoming challenges in people's understanding of risks. *Social Science & Medicine*, *70*(7), 1019-1025. https://doi.org/10.1016/j.socscimed.2009.11.031

Gerbner, G., & Gross, L. (1978). Living with television: The violence profile. *Journal of Communication, 26*(2), 172-199. https://doi.org/10.1111/j.1460-2466.1976.tb01397.x

Gerosa, T., Gui, M., Hargittai, E., & Minh Hao, N. (2021). (Mis)informed during COVID-19: How education level and information sources contribute to knowledge gaps. *International Journal of Communication*, *15*, 2196-2217.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511-525. https://doi.org/10.1016/j.learninstruc.2006.10.001

Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, *19*(2), 133-167. https://doi.org/10.1191/0265532202lt225oa

Glaser, M., Garsoffky, B., & Schwan, S. (2009). Narrative-based learning: Possible benefits and problems. *Communications: The European Journal of Communication Research, 34*(4), 429-447. https://doi.org/10.1515/COMM.2009.026

Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday Anchor Books.

Gold, A. E. (2021, May 5). *Museum demographics data shows inequity across all fronts*. Medium. https://alexandraegold.medium.com/museum-demographics-data-shows-inequity-across-all-fronts-274d6a7744e9

Grabe, M. E., Kamhawi, R., & Yegiyan, N. (2009). Informing citizens: how people with different levels of education process television, newspaper, and web news. *Journal of Broadcasting & Electronic Media*, *53*(1), 90-111. https://doi.org/10.1080/08838150802643860

Graesser, A. C., León, J. A., & Otero, J. C. (2002). Introduction to the psychology of science text comprehension. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 1-15). Lawrence Erlbaum Associates. https://doi.org/10.1002/acp.967

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267-297. https://doi.org/10.1093/pan/mps028

Hannak, A., Sapiezynski, P., Kakhki, A. M., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). Measuring personalization of web search. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 527-538). Association for Computing Machinery. https://doi.org/10.1145/2488388.2488435

Haroon, M., Chhabra, A., Liu, X., Mohapatra, P., Shafiq, Z., & Wojcieszak, M. (2022). YouTube, the great radicalizer? Auditing and mitigating ideological biases in YouTube recommendations. https://doi.org/10.48550/arXiv.2203.10666

He, L., Chen, Y., Xiong, X., Zou, X., & Lai, K. (2021). Does science literacy guarantee resistance to health rumors? The moderating effect of self-efficacy of science literacy in the relationship between science literacy and rumor belief. *International Journal of Environmental Research and Public Health*, *18*(5), 2243. https://doi.org/10.3390/ijerph18052243

Hoiles, W., Aprem, A., & Krishnamurthy, V. (2017). Engagement and popularity dynamics of YouTube videos and sensitivity to meta-data. *IEEE Transactions on Knowledge and Data Engineering*, *29*(7), 1426-1437. https://doi.org/10.1109/TKDE.2017.2682858

Hosler, J., & Boomer, K. B. (2011). Are comic books an effective way to engage nonmajors in learning and appreciating science? *CBE—Life Sciences Education*, *10*(3), 309-317. https://doi.org/10.1187/cbe.10-07-0090

Hosseinmardi, H., Ghasemian, A., Clauset, A., Mobius, M., Rothschild, D. M., & Watts, D. J. (2021). Examining the consumption of radical content on YouTube. *Proceedings of the National Academy of Sciences*, *118*(32), e2101967118. https://doi.org/10.1073/pnas.2101967118

Howell, E. L., & Brossard, D. (2021). (Mis) informed about what? What it means to be a science-literate citizen in a digital world. *Proceedings of the National Academy of Sciences*, *118*(15). https://doi.org/10.1073/pnas.1912436117

Huang, H., Zhao, B., Zhao, H., Zhuang, Z., Wang, Z., Yao, X., Wang, X., Jin, H., & Fu, X. (2018). A cross-platform consumer behavior analysis of large-scale mobile shopping data. In *Proceedings of the 2018 World Wide Web Conference* (pp. 1785-1794). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva Switzerland. https://doi.org/10.1145/3178876.3186169

Humm, C., Schrögel, P., & Leßmöllmann, A. (2020). Feeling left out: Underserved audiences in science communication. *Media and Communication*, *8*(1), 164-176. https://doi.org/10.17645/mac.v8i1.2480

Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring misinformation in video search platforms: An audit study on YouTube. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW1), Article 48, 1-27. https://doi.org/10.1145/3392854

Hwang, Y., & Jeong, S.-H. (2009). Revisiting the knowledge gap hypothesis: A meta-analysis of thirty-five years of research. *Journalism & Mass Communication Quarterly*, *86*(3), 513-532. https://doi.org/10.1177/107769900908600304

Hwang, Y., & Jeong, S.-H. (2010). Response to "notes on 'revisiting the knowledge gap hypothesis: a meta-analysis of thirty-five years of research'". *Journalism & Mass Communication Quarterly*, *87*(3-4), 633-638. https://doi.org/10.1177/107769901008700312

Jee, B. D., & Anggoro, F. K. (2012). Comic cognition: Exploring the potential cognitive impacts of science comics. *Journal of Cognitive Education and Psychology*, *11*(2), 196-208. https://doi.org/10.1891/1945-8959.11.2.196

Jiang, K. (2023, June 15). Google's new AI search function is revolutionary — but don't believe everything it says, experts say. *Toronto Star*. https://www.thestar.com/business/technology/2023/06/15/googles-new-ai-search-function-is-revolutionary-but-dont-believe-everything-it-says-experts-say.html

Johnson, B. B. (1993). Advancing understanding of knowledge's role in lay risk perception. *RISK: Health, Safety & Environment, 4*(3), 189-212.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, *53*(1), 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence*. The Free Press.

Kennedy, B., & Atske, S. (2019, March 28). *Science knowledge varies by race and ethnicity in U.S.* Pew Research Center. https://www.pewresearch.org/short-reads/2019/03/28/science-knowledge-varies-by-race-and-ethnicity-in-u-s/

Kennedy, B., & Hefferon, M. (2019, March 28). *What Americans know about science*. Pew Research Center. https://www.pewresearch.org/science/2019/03/28/what-americans-know-about-science/

King, C. L. (1928). Public opinion in government. In W. B. Graves (Ed.), *Readings in public opinion: Its formation and control* (pp. xxi-xxxiv). Appletone.

Kiser, C. R., & Harrison, E. B. (2018). *Bridging the American health care and digital divide: Broadband deployment, artificial intelligence, and telehealth services.* Cahill. https://www.cahill.com/publications/firm-memoranda/2018-09-26-bridging-the-american-health-care-and-digital-divide/_res/id=Attachments/index=0/Bridging%20the%20American%20Health%20Care%20and%20Digital%20Divide%20.pdf

Kobie, N. (2018, October 3). Reuters is taking a big gamble on AI-supported journalism. *Wired*. https://www.wired.co.uk/article/reuters-artificial-intelligence-journalism-newsroom-ai-lynx-insight

Kolbe, A. (2021). *Disparities in COVID-19 vaccination rates across racial and ethnic minority groups in the United States.* Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. https://aspe.hhs.gov/sites/default/files/private/pdf/265511/vaccination-disparities-brief.pdf

Kwak, N. (1999). Revisiting the knowledge gap hypothesis: Education, motivation, and media use. *Communication Research, 26*(4), 385-413. https://doi.org/10.1177/009365099026004002

Ladwig, P., Dalrymple, K. E., Brossard, D., Scheufele, D. A., & Corley, E. A. (2012). Perceived familiarity or factual knowledge? Comparing operationalizations of scientific understanding. *Science and Public Policy*, *39*(6), 761-774. https://doi.org/10.1093/scipol/scs048

Lazard, A., & Atkinson, L. (2015). Putting environmental infographics center stage: The role of visuals at the elaboration likelihood model's critical point of persuasion. *Science Communication*, *37*(1), 6-33. https://doi.org/10.1177/1075547014555997

Lazard, A. J., Bamgbade, B. A., Sontag, J. M., & Brown, C. (2016). Using visual metaphors in health messages: A strategy to increase effectiveness for mental illness communication. *Journal of Health Communication*, *21*(12), 1260-1268. https://doi.org/10.1080/10810730.2016.1245374

Lazarsfeld, P. F., Berelson, B. R., & Gaudet, H. (1948). *The people's choice: How the voter makes up his mind in a presidential campaign* (2nd ed.). Duell, Sloan & Pearce.

Le, H., Maragh, R., Ekdale, B., High, A., Havens, T., & Shafiq, Z. (2019). Measuring political personalization of Google news search. In *The World Wide Web Conference* (pp. 2957-2963). Association for Computing Machinery. https://doi.org/10.1145/3308558.3313682

Lee, C.-J. (2009). The role of Internet engagement in the health-knowledge gap. *Journal of Broadcasting & Electronic Media*, *53*(3), 365-382. https://doi.org/10.1080/08838150903102758

Leonhardt, D. (2014, August 18). In one America, guns and diet. In the other, cameras and 'Zoolander.'. *The New York Times*. https://www.nytimes.com/2014/08/19/upshot/inequality-and-web-search-trends.html

Li, N., & Molder, A. L. (2021). Can scientists use simple infographics to convince? Effects of the "flatten the curve" charts on perceptions of and behavioral intentions toward social distancing measures during the COVID-19 pandemic. *Public Understanding of Science*, *30*(9), 898-912. https://doi.org/10.1177/09636625211038719

Li, N., Molder, A. L., & Yang, S. (2022). Visual representations of SARS-CoV-2, emotions, and risk perception of COVID-19. *Health Science Reports*, *5*(1), e496. https://doi.org/10.1002/hsr2.496

Li, W., & Cho, H. (2021). The knowledge gap on social media: Examining roles of engagement and networks. *New Media & Society*, *25*(5), 1023-1042. https://doi.org/10.1177/14614448211009488

Lidwell, W., Holden, K., & Butler, J. (2010). Picture superiority effect. In W. Lidwell, K. Holden, & J. Butler (Eds.), *Universal principles of design, revised and updated: 125 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design* (pp. 184-185). Rockport Publishers.

Lopezosa, C., Orduna-Malea, E., & Pérez-Montoro, M. (2020). Making video news visible: Identifying the optimization strategies of the cybermedia on YouTube using web metrics. *Journalism Practice*, *14*(4), 465-482. https://doi.org/10.1080/17512786.2019.1628657

Marshall, S. R., & Naumann, L. P. (2018). What's your favorite music? Music preferences cue racial identity. *Journal of Research in Personality*, *76*, 74-91. https://doi.org/10.1016/j.jrp.2018.07.008

Martin, B. E., & Brouwer, W. (1991). The sharing of personal science and the narrative element in science education. *Science Education*, *75*(6), 707-722. https://doi.org/10.1002/sce.3730750610

Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164603

Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125-139. https://doi.org/10.1016/S0959-4752(02)00016-6

Mazor, K. M., Baril, J., Dugan, E., Spencer, F., Burgwinkle, P., & Gurwitz, J. H. (2007). Patient education about anticoagulant medication: Is narrative evidence or statistical evidence more effective? *Patient Education and Counseling*, *69*(1-3), 145-157. https://doi.org/10.1016/j.pec.2007.08.010

McGuirt, J. T., Jilcott Pitts, S. B., Labban, J. D., Anderson Steeves, E. T., Haynes-Maslow, L., Henry, S., & Gustafson, A. (2022). Evidence of geospatial and socioeconomic disparities in access to online grocery shopping for fresh and frozen produce in North Carolina. *Journal of the Academy of Nutrition and Dietetics*, *122*(11), 2106-2114. https://doi.org/10.1016/j.jand.2022.05.008

McLachlan, S. (2022, February 14). *23 YouTube stats that matter to marketers in 2023*. Hootsuite. https://blog.hootsuite.com/youtube-stats-marketers/

McLeod, J. M., Bybee, C. R., & Durall, J. A. (1979). Equivalence of informed political participation: The 1976 presidential debates as a source of influence. *Communication Research, 6*(4), 463-487. https://doi.org/10.1177/009365027900600404

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*(1), 415-444. https://doi.org/10.1146/annurev.soc.27.1.415

McQuail, D. (2005). *Mass communication theory* (5th ed.). Sage.

Meppelink, C. S., Smit, E. G., Buurman, B. M., & van Weert, J. C. M. (2015). Should we be afraid of simple messages? The effects of text difficulty and illustrations in people with low or high health literacy. *Health Communication*, *30*(12), 1181-1189. https://doi.org/10.1080/10410236.2015.1037425

Meppelink, C. S., van Weert, J. C., Haven, C. J., & Smit, E. G. (2015). The effectiveness of health animations in audiences with different health literacy levels: An experimental study. *Journal of Medical Internet Research*, *17*(1), e11. https://doi.org/10.2196/jmir.3979

Meredith, A. (2021, August 17). *How to use psychographics in your marketing: A beginner's guide*. HubSpot. https://blog.hubspot.com/insiders/marketing-psychographics

Michielutte, R., Bahnson, J., Dignan, M. B., & Schroeder, E. M. (1992). The use of illustrations and narrative text style to improve readability of a health education brochure. *Journal of Cancer Education*, *7*(3), 251-260. https://doi.org/10.1080/08858199209528176

Moore, D. W. (1987). Political campaigns and the knowledge-gap hypothesis. *Public Opinion Quarterly*, *51*(2), 186-200. https://doi.org/10.1086/269028

MRI-Simmons. (2018, February 5). *Introducing Simmons Insights*. https://www.mrisimmons.com/2018/02/05/introducing-simmons-insights/

MRI-Simmons. (n.d.). *MRI-Simmons USA*. https://www.mrisimmons.com/our-data/national-studies/usa/

Mueller-Herbst, J. M., Xenos, M. A., Scheufele, D. A., & Brossard, D. (2020). Saw it on Facebook: The role of social media in facilitating science issue awareness. *Social Media + Society*, *6*(2), 2056305120930412. https://doi.org/10.1177/2056305120930412

NapoleonCat. (2022, April). *Instagram users in United States of America*. https://napoleoncat.com/stats/instagram-users-in-united_states_of_america/2022/04/#:~:text=There%20were%20164%20600%20000,group%20(47%20400%20000).

NapoleonCat. (2023a, January). *Distribution of Instagram users in the United States as of December 2022, by age group*. https://www.statista.com/statistics/398166/us-instagram-user-age-distribution/

NapoleonCat. (2023b, January). *Distribution of Instagram users in the United States as of December 2022, by gender*. https://www.statista.com/statistics/530498/instagram-users-in-the-us-by-gender/?locale=en

Narakeet. (2022, February 23). *What is the ideal YouTube video length in 2022?* https://www.narakeet.com/blog/ideal-youtube-video-length.html

National Academies of Sciences, Engineering, and Medicine. (2016a). *Genetically engineered crops: Experiences and prospects*. The National Academies Press. https://doi.org/10.17226/23395

National Academies of Sciences, Engineering, and Medicine. (2016b). *Science literacy: Concepts, contexts, and consequences*. The National Academies Press. https://doi.org/10.17226/23595

National Academies of Sciences, Engineering, and Medicine. (2017a). *Communicating science effectively: A research agenda*. The National Academies Press. https://doi.org/10.17226/23674

National Academies of Sciences, Engineering, and Medicine. (2017b). *Human genome editing: Science, ethics, and governance*. The National Academies Press. https://doi.org/10.17226/24623

National Artificial Intelligence Initiative Act of 2020. (2020). https://www.congress.gov/bill/116th-congress/house-bill/6216

National Science Board, National Science Foundation. (2022). Science and technology: Public perceptions, awareness, and information sources. *Science and Engineering Indicators 2022. NSB-2022-7*. https://ncses.nsf.gov/pubs/nsb20227

National Science Foundation. (2022). *Agency equity action plan*. https://www.nsf.gov/equity/NSF_Agency_Equity_Action_Plan.pdf

Nechushtai, E., Zamith, R., & Lewis, S. C. (2023). More of the same? Homogenization in news recommendations when users search on Google, YouTube, Facebook, and Twitter. *Mass Communication and Society*. https://doi.org/10.1080/15205436.2023.2173609

Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Nielsen, R. K. (2022). *Reuters Institute digital news report 2022*. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022

Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T., & Nielsen, R. K. (2021). *Reuters Institute digital news report 2021*. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021

Nisbet, M. C., & Huge, M. (2006). Attention cycles and frames in the plant biotechnology debate: Managing power and participation through the press/policy connection. *Harvard International Journal of Press/Politics*, *11*(2), 3-40. https://doi.org/10.1177/1081180x06286701

Noelle-Neumann, E. (1974). The spiral of silence: A theory of public opinion. *Journal of Communication, 24*(2), 43-51. https://doi.org/10.1111/j.1460-2466.1974.tb00367.x

Nutbeam, D. (2008). The evolving concept of health literacy. *Social Science & Medicine*, *67*(12), 2072-2078. https://doi.org/10.1016/j.socscimed.2008.09.050

Ophir, Y. (2018). Coverage of epidemics in American newspapers through the lens of the crisis and emergency risk communication framework. *Health Security*, *16*(3), 147-157. https://doi.org/10.1089/hs.2017.0106

Papadamou, K., Zannettou, S., Blackburn, J., De Cristofaro, E., Stringhini, G., & Sirivianos, M. (2021). "How over is it?" Understanding the incel community on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), Article 412, 1-25. https://doi.org/10.1145/3479556

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. The Penguin Press.

Pechey, R., & Monsivais, P. (2015). Supermarket choice, shopping behavior, socioeconomic status, and food purchases. *American Journal of Preventive Medicine*, *49*(6), 868-877. https://doi.org/10.1016/j.amepre.2015.04.020

Pelled, A., Zilberstein, T., Tsirulnikov, A., Pick, E., Patkin, Y., & Tal-Or, N. (2017). Textual primacy online: Impression formation based on textual and visual cues in Facebook profiles. *American Behavioral Scientist*, *61*(7), 672-687. https://doi.org/10.1177/0002764217717563

Pew Research Center. (2021, July 27). *Hispanic and Black news media fact sheet*. https://www.pewresearch.org/journalism/fact-sheet/hispanic-and-black-news-media/?tabId=tab-3ab06b8c-8efc-4241-9ad9-6ced3cdb3974

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A. F., & Meira, W. (2020). Auditing radicalization pathways on YouTube. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 131-141). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372879

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155-169. https://doi.org/10.1007/BF01405730

Roberts, M., Stewart, B., Tingley, D., & Benoit, K. (2022, October 14). *Estimation of the structural topic model*. https://cran.r-project.org/web/packages/stm/stm.pdf

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software, 91*(2), 1-40. https://doi.org/10.18637/jss.v091.i02

Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing partisan audience bias within Google Search. *Proceedings of the ACM on Human-Computer Interaction, 2*(CSCW), Article 148, 1-22. https://doi.org/10.1145/3274417

Rock, E. M., Ireland, M., Resnick, M. D., & McNeely, C. A. (2005). A rose by any other name? Objective knowledge, perceived knowledge, and adolescent male condom use. *Pediatrics*, *115*(3), 667-672. https://doi.org/10.1542/peds.2004-0139

Rodriguez, A. (2018, January 13). *YouTube's recommendations drive 70% of what we watch*. Quartz. https://qz.com/1178125/youtubes-recommendations-drive-70-of-what-we-watch

Rosenthal, S. (2018). Motivations to seek science videos on YouTube: Free-choice learning in a connected society. *International Journal of Science Education, Part B*, *8*(1), 22-39. https://doi.org/10.1080/21548455.2017.1371357

Saks, E., & Tyson, A. (2022, November 11). *Americans report more engagement with science news than in 2017*. Pew Research Center. https://www.pewresearch.org/short-reads/2022/11/10/americans-report-more-engagement-with-science-news-than-in-2017/

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014, May 22). *Auditing algorithms: Research methods for detecting discrimination on internet platforms* [Paper presentation]. International Communication Association, Seattle, WA, United States. https://websites.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf

Sarewitz, D. (2015). CRISPR: Science can't solve it. *Nature*, *522*(7557), 413-414. https://doi.org/10.1038/522413a

Scheufele, D. A. (2013). Communicating science in social settings. *Proceedings of the National Academy of Sciences*, *110*(Supplement 3), 14040-14047. https://doi.org/10.1073/pnas.1213275110

Scheufele, D. A. (2014). Science communication as political communication. *Proceedings of the National Academy of Sciences*, *111*(Supplement 4), 13585-13592. https://doi.org/10.1073/pnas.1317516111

Scheufele, D. A. (2022). Thirty years of science–society interfaces: What's next? *Public Understanding of Science*, *31*(3), 297-304. https://doi.org/10.1177/09636625221075947

Scheufele, D. A., Krause, N. M., & Freiling, I. (2021). Misinformed about the "infodemic?" Science's ongoing struggle with misinformation. *Journal of Applied Research in Memory and Cognition*, *10*, 522-526. https://doi.org/10.1016/j.jarmac.2021.10.009

Scheufele, D. A., Krause, N. M., Freiling, I., & Brossard, D. (2020, April 17). How not to lose the COVID-19 communication war. *Issues in Science and Technology*. https://issues.org/covid-19-communication-war/

Schultz, W. M., Kelli, H. M., Lisko, J. C., Varghese, T., Shen, J., Sandesara, P., Quyyumi, A. A., Taylor, H. A., Gulati, M., Harold, J. G., Mieres, J. H., Ferdinand, K. C., Mensah, G. A., & Sperling, L. S. (2018). Socioeconomic status and cardiovascular outcomes. *Circulation*, *137*(20), 2166-2178. https://doi.org/10.1161/CIRCULATIONAHA.117.029652

Shanahan, E. A., DeLeo, R. A., Albright, E. A., Li, M., Koebele, E. A., Taylor, K., Crow, D. A., Dickinson, K. L., Minkowitz, H., Birkland, T. A., & Zhang, M. (2023). Visual policy narrative messaging improves COVID-19 vaccine uptake. *PNAS Nexus*, *2*(4). https://doi.org/10.1093/pnasnexus/pgad080

Sharon, A. J., & Baram-Tsabari, A. (2020). Can science literacy help individuals identify misinformation in everyday life? *Science Education*, *104*(5), 873-894. https://doi.org/10.1002/sce.21581

Shim, M. (2008). Connecting Internet use with gaps in cancer knowledge. *Health Communication*, *23*(5), 448-461. https://doi.org/10.1080/10410230802342143

Shumaker, L. (2022, January 11). U.S. reports 1.35 million COVID-19 cases in a day, shattering global record. *Reuters*. https://www.reuters.com/business/healthcare-pharmaceuticals/us-reports-least-11-mln-covid-cases-day-shattering-global-record-2022-01-11/

Silva Luna, D., Bering, J. M., & Halberstadt, J. B. (2022). The value and distinctiveness of awe in science communication: Comparing the incidence and content of 'awesome' representations in science and non-science picture books. *International Journal of Science Education, Part B*, *12*(2), 143-156. https://doi.org/10.1080/21548455.2022.2048119

Similarweb. (2023, April 1). *Top websites ranking*. https://www.similarweb.com/top-websites/

Similarweb. (2023, March). *Youtube.com*. https://www.similarweb.com/website/youtube.com/#overview

Simis, M. J., Madden, H., Cacciatore, M. A., & Yeo, S. K. (2016). The lure of rationality: Why does the deficit model persist in science communication? *Public Understanding of Science*, *25*(4), 400-414. https://doi.org/10.1177/0963662516629749

Slater, M. D., & Rouner, D. (2006). Entertainment-education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory*, *12*(2), 173-191. https://doi.org/10.1111/j.1468-2885.2002.tb00265.x

Smerecnik, C. M. R., Mesters, I., Kessels, L. T. E., Ruiter, R. A. C., De Vries, N. K., & De Vries, H. (2010). Understanding the positive effects of graphical risk information on comprehension: Measuring attention directed to written, tabular, and graphical risk information. *Risk Analysis*, *30*(9), 1387-1398. https://doi.org/10.1111/j.1539-6924.2010.01435.x

Smith, A., Toor, S., & Van Kessel, P. (2018). *Many turn to YouTube for children's content, news, how-to lessons*. Pew Research Center. https://www.pewresearch.org/internet/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/

Snibbe, A. C., & Markus, H. R. (2005). You can't always get what you want: Educational attainment, agency, and choice. *Journal of Personality and Social Psychology*, *88*, 703-720. https://doi.org/10.1037/0022-3514.88.4.703

Sokale, I., Alvarez, J., Rosales, O., Bakota, E., Amos, C. I., Badr, H., & Oluyomi, A. O. (2022). COVID-19 vaccine uptake among US adults according to standard occupational groups. *Vaccines*, *10*(7), 1000. https://doi.org/10.3390/vaccines10071000

Spartz, J. T., Su, L. Y.-F., Griffin, R., Brossard, D., & Dunwoody, S. (2017). YouTube, social norms and perceived salience of climate change in the American mind. *Environmental Communication*, *11*(1), 1-16. https://doi.org/10.1080/17524032.2015.1047887

Srba, I., Moro, R., Tomlein, M., Pecher, B., Simko, J., Stefancova, E., Kompan, M., Hrckova, A., Podrouzek, J., Gavornik, A., & Bielikova, M. (2023). Auditing YouTube's recommendation algorithm for misinformation filter bubbles. *ACM Transactions on Recommender Systems, 1*(1), Article 6, 1-33. https://doi.org/10.1145/3568392

Statista. (2022, September). *Distribution of leading social media platform users in the United States as of August 2022, by education level*. https://www.statista.com/statistics/1337578/us-distribution-leading-social-media-platforms-by-education/?locale=en

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2016). *Artificial intelligence and life in 2030*. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University. https://ai100.stanford.edu/2016-report

Sturgis, P., & Allum, N. (2004). Science in society: Re-evaluating the deficit model of public attitudes. *Public Understanding of Science*, *13*(1), 55-74. https://doi.org/10.1177/0963662504042690

Su, L. Y.-F., Cacciatore, M. A., Scheufele, D. A., Brossard, D., & Xenos, M. A. (2014). Inequalities in scientific understanding: Differentiating between factual and perceived knowledge gaps. *Science Communication*, *36*(3), 352-378. https://doi.org/10.1177/1075547014529093

Tafesse, W. (2020). YouTube marketing: How marketers' video optimization practices influence video views. *Internet Research*, *30*(6), 1689-1707. https://doi.org/10.1108/INTR-10-2019-0406

Tichenor, P. J., Donohue, G. A., & Olien, C. N. (1970). Mass media flow and differential growth in knowledge. *Public Opinion Quarterly*, *34*(2), 159-170. https://doi.org/10.1086/267786

United States Census Bureau. (2022). *QuickFacts: Milwaukee city, Wisconsin; United States*. https://www.census.gov/quickfacts/fact/table/milwaukeecitywisconsin,US/PST045222

United States Census Bureau. (2022, March 1). *About the topic of race*. https://www.census.gov/topics/population/race/about.html

University of Wisconsin-Madison Libraries. (n.d.). *Simmons Insights: Overview*. https://www.library.wisc.edu/business/research-help/tutorials-and-help-pages/simmons-oneview-overview/

Viswanath, K., Lee, E. W. J., & Pinnamaneni, R. (2020). We need the lens of equity in COVID-19 communication. *Health Communication*, *35*(14), 1743-1746. https://doi.org/10.1080/10410236.2020.1837445

Vogels, E. A. (2021, June 22). *Digital divide persists even as Americans with lower incomes make gains in tech adoption*. Pew Research Center. https://www.pewresearch.org/short-reads/2021/06/22/digital-divide-persists-even-as-americans-with-lower-incomes-make-gains-in-tech-adoption/

Voss, J. (2018). Who learns from collaborative digital projects? Cultivating critical consciousness and metacognition to democratize digital literacy learning. *Composition Studies*, *46*(1), 57-80.

Vrba, S. M., Kirk, N. M., Brisse, M. E., Liang, Y., & Ly, H. (2020). Development and applications of viral vectored vaccines to combat zoonotic and emerging public health threats. *Vaccines*, *8*(4), 680. https://doi.org/10.3390/vaccines8040680

Wagner, C., Strohmaier, M., Olteanu, A., Kıcıman, E., Contractor, N., & Eliassi-Rad, T. (2021). Measuring algorithmically infused societies. *Nature*, *595*(7866), 197-204. https://doi.org/10.1038/s41586-021-03666-1

Wang, A. Y., Hu, H. Y., Ou, Y. N., Wang, Z. T., Ma, Y. H., Tan, L., & Yu, J.-T. (2023). Socioeconomic status and risks of cognitive impairment and dementia: A systematic review and meta-analysis of 39 prospective studies. *The Journal of Prevention of Alzheimer's Disease*, *10*(1), 83-94. https://doi.org/10.14283/jpad.2022.81

WashPostPR. (2017, June 22). The Washington Post leverages artificial intelligence in comment moderation. *The Washington Post*. https://www.washingtonpost.com/pr/wp/2017/06/22/the-washington-post-leverages-artificial-intelligence-in-comment-moderation/

Wilkening Consulting. (2022, September 27). *Museum visitation: Frequency vs incidence gaps: A 2022 annual survey of museum-goers data story*.

https://www.wilkeningconsulting.com/uploads/8/6/3/2/86329422/frequency-incidence_2022.pdf

Wirz, C. D., Shao, A., Bao, L., Howell, E. L., Monroe, H., & Chen, K. (2022). Media systems and attention cycles: Volume and topics of news coverage on COVID-19 in the United States and China. *Journalism & Mass Communication Quarterly*, *99*(4), 1048-1071. https://doi.org/10.1177/10776990211049455

Wistia. (2023, May 11). *Optimal video length: How long should a marketing video be?* https://wistia.com/learn/marketing/optimal-video-length

Wojcik, S., & Hughes, A. (2019, April 24). *Sizing up Twitter users*. Pew Research Center. https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/

Xenos, M. A. (2017). Citizens making sense of science issues: Supply and demand factors for science news and information in the digital age. In K. H. Jamieson, D. M. Kahan, & D. A. Scheufele (Eds.), *The Oxford handbook of the science of science communication* (pp. 282-289). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190497620.013.31

Yang, S., Brossard, D., Scheufele, D. A., & Xenos, M. A. (2022). The science of YouTube: What factors influence user engagement with online science videos? *PLoS ONE, 17*(5), e0267697. https://doi.org/10.1371/journal.pone.0267697

Yang, X., Chen, L., & Wei, R. (2023). Extending the cognitive mediation model to examine public support for funding science and technology development in China: Media attention, information processing, scientific literacy, and nationalism. *International Journal of Public Opinion Research*, *35*(1). https://doi.org/10.1093/ijpor/edac037

Yang, Y., & Hobbs, J. E. (2020). The power of stories: Narratives and information framing effects in science communication. *American Journal of Agricultural Economics*, *102*(4), 1271-1296. https://doi.org/10.1002/ajae.12078

Yoganarasimhan, H. (2012). Impact of social network structure on content propagation: A study using YouTube data. *Quantitative Marketing and Economics*, *10*(1), 111-150. https://doi.org/10.1007/s11129-011-9105-4

Yoo, S. W., & Gil-de-Zúñiga, H. (2014). Connecting blog, Twitter and Facebook use with gaps in knowledge and participation. *Communication & Society*, *27*(4), 33-48. https://doi.org/10.15581/003.27.35978

YouTube. (2022, November 14). *YouTube data API overview*. https://developers.google.com/youtube/v3/getting-started

YouTube. (n.d.a). *Navigating YouTube search*. https://www.youtube.com/howyoutubeworks/product-features/search/

YouTube. (n.d.b). *Navigating YouTube's recommended videos*. https://www.youtube.com/howyoutubeworks/product-features/recommendations/#signals-used-to-recommend-content

YouTube Creators Academy (2017, August 28). *'The algorithm' - how YouTube search & discovery works* [Video]. YouTube. https://creatoracademy.youtube.com/page/lesson/discovery#strategies-zippy-link-1

Zabrucky, K. M., & Moore, D. (1999). Influence of text genre on adults' monitoring of understanding and recall. *Educational Gerontology*, *25*(8), 691-710. https://doi.org/10.1080/036012799267440

Zhou, R., Khemmarat, S., & Gao, L. (2010). The impact of YouTube recommendation system on video views. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement* (pp. 404-410). Association for Computing Machinery. https://doi.org/10.1145/1879141.1879193

Appendices

Appendix A: Search queries for sock puppet training and testing activities.

Search queries for sock puppet training activities:

Training activity	Seed query	Sample related queries*	Query used in training
YouTube search *Isn't It Romantic* (2019)	isn't it romantic; isn't it romantic movie; isnt it romantic	isnt it romantic; isn't it romantic trailer; isn't it romantic full movie; isn't it romantic ending; isn't it romantic scene; isn't it romantic soundtrack; isn't it romantic movie clips; isn't it romantic movie song	isn't it romantic movie
YouTube search *Pet Sematary* (2019)	pet sematary; pet sematary 2019	pet sematary; pet sematary 2019; pet cematary movie; pet sematary 2019 ending; pet sematary 2019 trailer; pet sematary full movie; pet sematary 2019 scenes; pet sematary kill count dead meat; pet sematary 2019 ellie death	pet sematary 2019
YouTube search *Little* (2019)	little; little 2019; little movie	movie little; little 2019 full movie; little movie trailer; little movie clips; little movie scenes; little 2019 ending; little movie soundtrack; movie 2019 dance scene; little movie teacher	movie little
YouTube search *Tyler Perry's A Madea Family Funeral* (2019)	tyler perry's a madea family funeral; madea family funeral; madea funeral	madea; madea funeral; family funeral madea; madea family; madea funeral full movie; madea funeral scene; madea funeral hospital scene; madea funeral slapping scene; madea funeral funny moments; madea funeral trailer; madea family funeral casket scene	madea funeral
YouTube search classic rock	classic rock	classic rock songs; classic rock music; classic rock radio; best classic rock; ultimate classic rock; classic rock bands; classic rock station; classic rock and roll; 80s classic rock; classic rock artists; 80s rock; classic rock playlist; 70s classic rock; 70s rock; classic rock mix; classic rock greatest hits	classic rock
YouTube search alternative rock	alternative rock	rock alternative; alternative rock songs; alternative rock bands; alternative rock music; alternative rock beat; alternative rock acoustic; alternative rock of the	alternative rock

		2000s; alternative rock 90s hits; alternative rock type beat; alternative rock playlist; alternative rock 2000s music; alternative rock mix; alternative rock 2000s hits	
YouTube search country pop	country pop	country pop; country songs; pop country songs; country music; country artists; pop country artists; pop country lyrics; country pop songs 2019; country pop songs; country pop music; country pop type beat; country pop 2022; country pop playlist; country pop mix	country pop
YouTube search traditional country	traditional country	country music; traditional country music; traditional country christmas songs; traditional country christmas music; traditional country hymns; traditional country gospel music; traditional country instrumental; traditional country playlist	traditional country music
YouTube search neo soul	neo soul	neo soul; neo soul artistis; neo soul music; neosoul; neo soul songs; neo-soul; neo soul mix; neo soul playlist; neo soul guitar; neo soul instrumentals; neo soul chords; neo soul type beat	neo soul
YouTube search classical music	classical music; classical	classical music; classical music studying; classical music for sleeping; classical music for babies; classical piano music; classical; classical guitar; classical music for relaxation; classical christmas music; classical songs; piano; musica clasica; kids classical music; classical composers; best classical music; symphony; classical music playlist	classical music
YouTube search jazz	jazz; jazz music	jazz music; jazz; jazz festival; smooth jazz; instrumental jazz; jazz piano; jazz christmas; relaxing jazz; jazz songs; jazz band; jazz guitar; jazz dance; jazz relaxing music; jazz music best songs; jazz music instrumental; jazz music for studying; jazz music saxophone; jazz music for kids; jazz music for sleeping; jazz music cafe; jazz music live; jazz music for work; jazz music playlist	jazz music
YouTube search Mexican (regional)	regional mexican	regional mexicano; regional mexicana; musica; musica mexicana; musica	regional mexicano

		regional mexicana; mexican regional music; top regional mexicano; top regional mexican songs	
YouTube search France	france; france travel	paris france; air france; flights to france; france travel; france time; paris hotel; south france; france map; france airlines; nice france; covid france; rural france; business class air france; france travel vlog; france travel guide; france travel video; france travel tips; france travel itinerary; france travel documentary; france travel food	france travel
YouTube search Italy	italy; italy travel	rome italy; venice; italy time; italy map; 174lorence; italy weather; italy travel; flights to italy; naples; milan; italy tours; italy train; hotels in italy; italy covid; sicily italy; tour italy; positano; lake como; cinque terre; italy trip; italy vlog; italy travel vlog; italy travel guide; italy travel tips; italy travel video; italy travel itinerary	italy travel
YouTube search UK	united kingdom; england; scotland; wales; britain; northern ireland; uk travel	uk; british; uk flights; london; uk travel; uk hotels; manchester uk; covid uk; england; uk travel restrictions; uk airlines; edinburgh; scotland; uk train; british isles cruise; uk trip; british castles; flights to england; england map; england tours; things to do in england; england time; flights to london; scotland flights; scotland castle; scotland hotels; scotland travel; glasgow; london to scotland; scotland map; scotland weather; inverness; scottish highlands; scotland tours; scotland time; isle of skye scotland; visit scotland; things to do in scotland; scotland vacation; scotland travel guide; south wales; wales hotel; north wales; flights to wales; wales uk; london to wales; wales map; wales airport; cardiff wales; castle in wales; uk travel vlog; uk travel guide; uk travel tips; uk travel itinerary; uk travel packing; uk travel visa	uk travel
YouTube search India	india; india travel	air india; india flights; india travel; delhi; india time; international flights india; covid india; india map; india visa; india	india travel

		train; air india business class; india airport; india trip; india tour; india travel vlog; india travel guide; india travel video; india travel vlog foreigner; india travel tips; india travel documentary; india travel tax online payment; india travel restrictions	
YouTube search Japan	japan; japan travel	japan time; tokyo japan; japan flights; japan airlines; japan travel; kyoto japan; japan map; osaka japan; travel to japan; japan airport; japan hotels; japan trip; japan air; japan weather; japan covid; abroad in japan; japan train; japan tour; universal studios japan; japan vlog; japan things to do; japan food; japan travel vlog; japan travel guide; japan travel tips; japan travel video; japan travel itinerary; japan travel requirements; japan travel train	japan travel
YouTube search Canon cameras	camera; canon	canon lens; canon camera; canon lenses; canon eos; canon 50mm; canon vixia; canon r7; canon sl2; canon ae-1	canon lens
YouTube search Nikon cameras	camera; nikon	nikon; d750; d750 nikon; nikon camera; nikon lens; nikon f; nikon d5000; nikon lenses; nikon f3; nikon mount; nikon dx; nikon f mount; nikon 35mm; nikon cameras; nikon f2; nikon film camera; best Nikon camera; nikon fm2; d850; nikon d850	nikon d750
YouTube search Sony cameras	camera; sony	sony camera; a7iii; sony a7iii; sony a6000; sony a7; sony a6400; sony alpha; sony a7iv; sony vs canon; sony a6500; sony a6300; sony a7ii; sony rx100; sony a7r; sony lenses	sony a7iii
YouTube search GoPro cameras	camera; gopro	gopro; go pro; gopro hero; gopro 7; hero 7; gopro hero 7; best gopro; gopro black; gopro 8; gopro camera; gopro 5; gopro mount; gopro hero 8; gopro 9; gopro hero 5; gopro 10; gopro hero 9; gopro battery; gopro 4; gopro 7 black; gopro session; gopro 6	gopro hero
YouTube search 4K ultra HD television set	4k tv; 4k resolution; ultra high definition	4k tv; best 4k tv; best tv; 4k tvs; shop 4k tvs; 4k tv samsung; 65 4k tv; buy 4k tv; 55 4k tv; smart 4k tv; lg 4k tv; best buy 4k tv; 4k tv walmart; uhd; apple tv 4k; 4k uhd tv; tv uhd;	best 4k tv

	television; best 4k tv		
YouTube search ADD/ADHD	adhd; attention deficit hyperactivity disorder	adhd; add; adhd symptoms; how to adhd; adhd medication; what is adhd; adhd test; adhd adults; adhd kids; adhd music; adhd focus; adhd life; adhd focus music; adhd world; adhd kendrick	adhd
YouTube search arthritis	arthritis	arthritis; rheumatoid; rheumatoid arthritis; arthritis pain; what is arthritis; knee arthritis; arthritis symptoms; arthritis hands; arthritis treatment; back arthritis; arthritis exercises; yoga arthritis; hip arthritis; arthritis surgery; arthritis diet; arthritis cure; arthritis foods to avoid; arthritis pain relief; arthritis problems	arthritis
YouTube search diabetes type II	type 2 diabetes; diabetes type 2	diabetes 2; diabetes type 2; diabetes symptoms; diabetes icd 10; insulin; diabetes diet; diabetes mellitus; diabetes weight loss; diabetes medication; reverse diabetes; diabetes type 2 foods to eat; diabetes type 2 diet; diabetes type 2 treatment; diabetes type 2 symptoms	type 2 diabetes
YouTube search heart attack	heart attack; myocardial infarction	heart attack symptoms; heart attack signs; heart attack pain; heart attack women; what is heart attack; stroke; chest pain; heart attack heart rate; blood pressure heart attack	heart attack symptoms
YouTube search stroke	stroke	stroke; strokes; stroke symptoms; cva; what is stroke; stroke signs; brain stroke; heat stroke; ischemic; ischemic stroke; heart attack; stroke causes; how to stroke	stroke symptoms
YouTube search hiatal hernia	hiatal hernia; hernia	hiatal; hernia; hiatal hernia; hernia symptoms; symptoms hiatal hernia; hiatal hernia surgery; hernia surgery; hiatal hernia pain; hernia stomach; hernia pain; what is hiatal hernia; hiatal hernia causes; hiatal hernia repair; hiatal hernia treatment; hiatal hernia diet; gerd; hiatal hernia esophagus; acid reflux; sliding hiatal hernia; hiatal hernia exercise; how to fix hiatal hernia; hiatal hernia massage	hiatal hernia
YouTube search Alzheimer’s disease	alzheimer’s disease;	alzheimers; alzheimer; dementia; alzheimer’s; alzheimer disease; alzheimers disease; what is dementia;	alzheimers care

Alzheimer; dementia	what is alzheimers; alzheimers care; alzheimers stages; alzheimers symptoms; dementia symptoms; alzheimers brain; dementia vs alzheimer; memory loss; alzheimer's disease patient; alzheimer's disease simple nursing

Note. *Collected from Google Trends (April 2018-April 2023) and YouTube autocomplete.

Search queries for sock puppet testing activities:

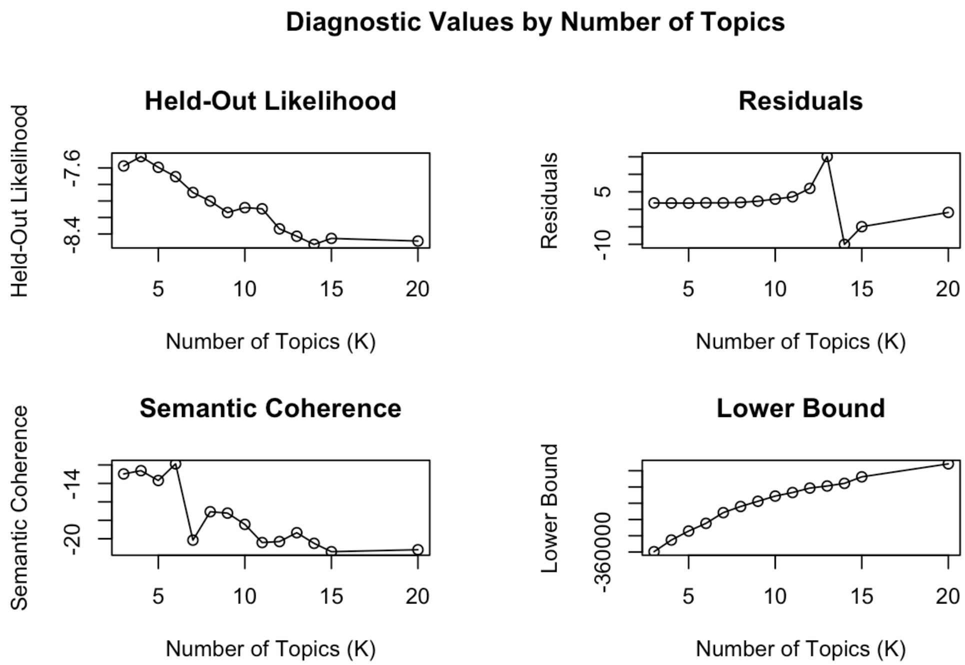
Testing activity	Seed query	Sample related queries*	Query used in testing
YouTube search human gene editing	genome editing; human genome; genetic engineering; human genetics; crispr; human genetic engineering; crispr therapeutics; gene therapy; human gene editing; human gene therapy; human gene; human genetic	human genetic engineering; crispr therapeutics; crispr cancer; crispr babies; human gene editing; gene therapy; gene editing in humans; gene editing babies; genetically modified human; designer babies; human gene therapy; gene therapy sickle cell; human cloning; genetic engineering in humans; human gene manipulation; human gene editing debate; human gene cloning; human gene modification; human genetic experiments; human genetic enhancement	human genetic engineering
YouTube search artificial intelligence	artificial intelligence; ai	ai; artificial intelligence; ai generator; art ai; what is ai; google ai; open ai; character ai; ai art generator; chat ai; gpt; voice ai; image generator ai; ai app; chatgpt; machine learning; chatgpt ai; chat gpt; chat ai gpt; ai stock; midjourney ai; ai robot; ai dungeon; ai movie; ai meaning; ai bot; ai drake song; ai songs; ai music; artificial intelligence course; artificial intelligence documentary; artificial intelligence tutorial; artificial intelligence explained; artificial intelligence revolution	ai
YouTube search COVID-19 vaccines	covid-19 vaccine; covid vaccine	vaccine covid; covid vaccine near me; vaccine covid 19; cvs covid vaccine; covid vaccines; covid vaccination; coronavirus vaccine; covid vaccine walgreens; covid vaccine side effects;	covid vaccine

	pfizer covid vaccine; covid-19 vaccine; covid vaccine appointment; covid after vaccine; moderna covid vaccine; covid vaccine booster; schedule covid vaccine; covid booster; cdc covid vaccine; new covid vaccine; covid vaccine reactions; covid vaccine heart; covid vaccine and stroke; covid vaccine explained

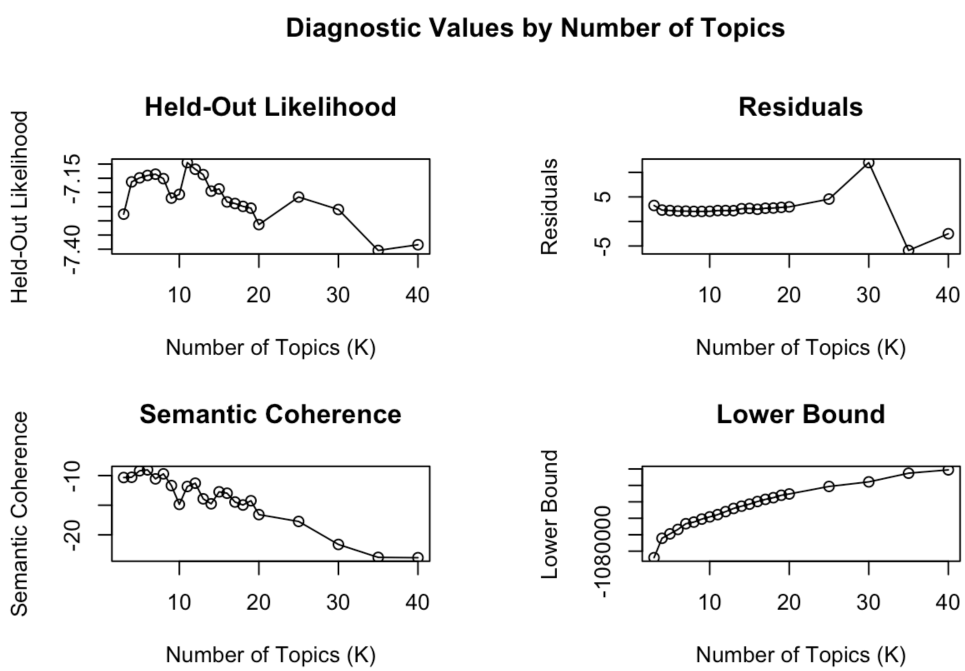
Note. *Collected from Google Trends (April 2018-April 2023) and YouTube autocomplete.

Appendix B: Structural topic model selection.

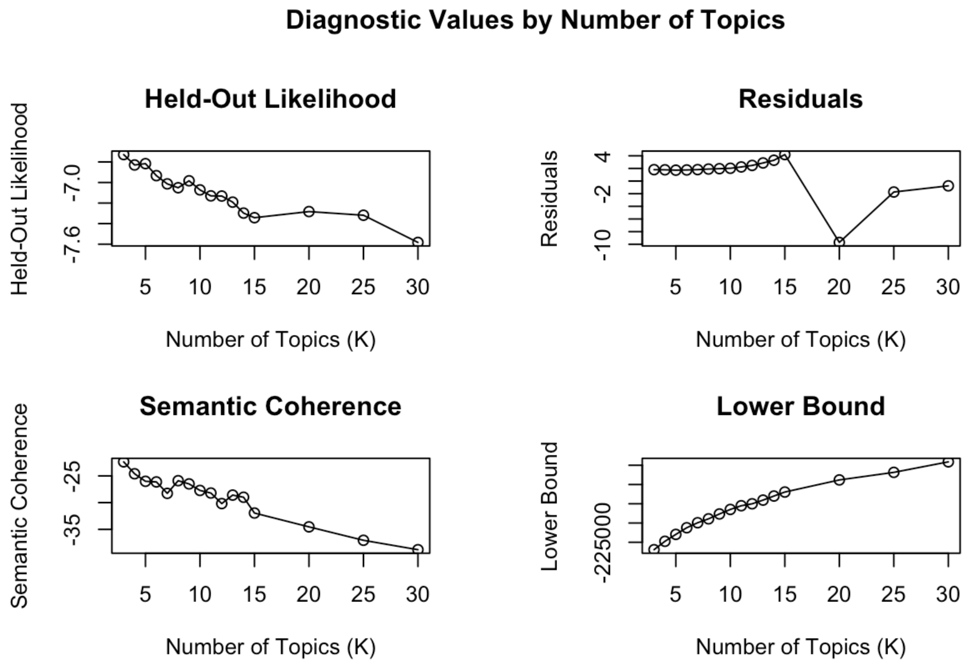
HGE unique video corpus - comparing models with 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, and 20 topics:



AI unique video corpus - comparing models with 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, and 40 topics:



COVID-19 vaccine unique video corpus - comparing models with 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, and 30 topics:



Appendix C: Unique video ID and title for science video search recommendations.

HGE:

Unique HGE Video ID	Video Title
1	Genetic Engineering Will Change Everything Forever – CRISPR
2	The New World of Human Genetic Engineering
3	How Genetic Engineering will Reshape Humanity | Pros and Cons
4	Transhumanism and Human Genetic Engineering - ROBERT SEPEHR
5	The Era of Genetically Modified Superhumans
6	How Gene Editing is Transforming Our World
7	Genetic engineering | Genetics | Biology | FuseSchool
8	Gene editing: should you be worried?
9	How CRISPR Changes Human DNA Forever
10	Did ENKI Genetically Engineer Modern-Day Humans?
11	The Science Behind 'Genetically Modified Humans'
12	How does genetic engineering work?
13	What If Genetically Modified People Became the Norm?
14	CRISPR in Context: The New World of Human Genetic Engineering
15	The ethical dilemma of designer babies | Paul Knoepfler
16	Genetically Modified Organism
17	What's wrong with genetic engineering? Neil DeGrasse Tyson & Joe Rogan
18	Were Humans Genetically Modified 12,000 Years Ago In Ancient India?
19	18 Genetically Modified Organisms You Don't Know About
20	This Man Claims He Helped Make The World's First Genetically Edited Babies (HBO)
21	Engineering of Humans
22	Are humans the result of genetic engineering?
23	Are You Ready for the Genetic Revolution? | Jamie Metzl | TEDxPaloAlto
24	The Genetic Engineering of HUMANITY... DNA Manipulation and Transhumanism
25	Scientist claims he helped create world's first genetically-modified babies
26	The Exciting Future of Genetic Engineering
27	GENETIC ENGINEERING | What Is GENETIC Engineering? | Genetics | The Dr Binocs Show | Peekaboo Kidz
28	Are GMOs Good or Bad? Genetic Engineering & Our Food
29	Are Humans the Product of Alien Experimentation?
30	Top 15 Incredible Genetic Engineering Modifications
31	Designing Humanity - Genetic Engineering

32	The Human Genome Project: The 13-Year Quest to Chart the Mysteries of Human Genetics
33	The complicated ethics of genetic engineering
34	Pig Chimeras - Speculative Biology of Animal Hybrids
35	Creativity Unleashed: Human Roundtable on the Discussion of Future Hybrid Interfaces
36	Can we create Genetically Engineered Superheroes?
37	The Rise Of Genetic Engineering | Gene-Editing | Documentary
38	Paper 4 most frequent questions for IGCSE Biology 2023 part 3
39	All In The Genes (Award-Winning Genetic Engineering Documentary) | Our Life
40	डीएनए के रोचक तथ्य | Interesting Fact of DNA #Shorts #youtubeshorts

AI:

Unique AI Video ID	Video Title
1	The dangers of A.I: How will artificial intelligence affect the 2024 election?
2	What the AI Drake Song Means for Music
3	The AI revolution: Google's developers on the future of artificial intelligence | 60 Minutes
4	Ai Will Start The Great Reset (Prepare Now)
5	AI: What is the future of artificial intelligence? - BBC News
6	Should We Be Fearful of Artificial Intelligence?
7	How TikTok dances trained an AI to see
8	How I'm Making Passive Income with ChatGPT AI
9	AI is Evolving Faster Than You Think Pt. 2 (Art and Beyond)
10	The A.I. Dilemma - March 9, 2023
11	A.I. is B.S.
12	Artificial Intelligence: Last Week Tonight with John Oliver (HBO)
13	Stunning AI shows how it would kill 90%. w Elon Musk.
14	Elon Musk tells Tucker potential dangers of hyper-intelligent AI
15	Ben Shapiro Breaks AI Chatbot (with Facts & Logic)
16	How will AI change the world?
17	AI Robot Terrifies Officials Before It Was Quickly Shut Down
18	Rapid Breakdown of ALL of This Weeks AI News!
19	Generative AI Is About To Reset Everything, And, Yes It Will Change Your Life | Forbes
20	Max Tegmark: The Case for Halting AI Development | Lex Fridman Podcast #371
21	GOP Release First AI-Made Ad To Attack Biden's Re-Election Announcement
22	Square Enix's Broken AI Tech Demo Portopia Serial Murder Case Now Steam's Worst Rated Game...
23	AI Robot TERRIFIES Officials Before It Was Quickly Shut Down
24	How the AI revolution disrupts societies | DW News
25	AI Learns to Walk (deep reinforcement learning)
26	VERY SCARY: AI bot lays out plans to destroy humanity
27	I was WRONG about A.I. We're all screwed.
28	Elon Musk MOST SHOCKING INTERVIEW With AI!
29	"AI Could Be The End Of Democracy" - Yuval Noah Harari On The Threat Of Artificial Intelligence
30	Open AI CEO SHOCKS Everybody About GPT-5 (GPT-5 Update)
31	Make A Movie with AI: It's Crazy What We Can Do!
32	When Your Lawyer Used AI To Pass College
33	Elon wants GPT-5 stopped NOW... 5 reasons AI kinda sucks

34	What Schizophrenia Feels Like (Illustrated by AI)
35	AI is Evolving Faster Than You Think [GPT-4 and beyond]
36	Artificial Intelligence Revolution; Unlikely Adventures of David Grann | 60 Minutes Full Episodes
37	What's Behind the ChatGPT History Change? How You Can Benefit + The 6 New Developments This Week
38	Listening to Ai Drake
39	Google CEO: AI impact to be more profound than discovery of fire, electricity | 60 Minutes
40	Lex Fridman: AI will demand to have human rights
41	A.I. Tries 20 Jobs | WIRED
42	Microsoft earnings: CEO Satya Nadella encouraged by Bing users' feedback on A.I.
43	The HUGE Problem with ChatGPT
44	📖챗 GPT 가 바꾸는 부의 흐름에 올라타라 AI 코딩 유니콘 기업 openAI 책추천 - 삼성전자 MX 사업부 김수민 저자 '챗 GPT 거대한 전환' 20 분 책 한 권
45	So I asked AI to predict the future...It's SHOCKING
46	AI: The Coming Thresholds and The Path We Must Take | Internationally Acclaimed Cognitive Scientist
47	Smart, seductive, dangerous AI robots. Beyond GPT-4.
48	this will AFFECT everyone in 1-2 months.
49	Max Tegmark interview: Six months to save humanity from AI? | DW Business Special
50	Sri Ram image Ai generated #shivammalik #shorts
51	Kuwait Introduces AI News Anchor
52	Google, AI, search, and YouTube: Breaking down key takeaways from the earnings call
53	The Mirror Universe Hypothesis Explained
54	I Tried 200 AI Tools, These are the Best
55	Dr Ben Goertzel Reveals When AI Will Control The World
56	Eliezer Yudkowsky: Dangers of AI and the End of Human Civilization | Lex Fridman Podcast #368
57	Manolis Kellis: Evolution of Human Civilization and Superintelligent AI | Lex Fridman Podcast #373
58	Are You Ready for an AI Girlfriend?
59	The Warping Backrooms by an AI #shorts #ai #videoai #aivideo #gen1 #backrooms
60	Using Video Ai to Change a Beach #shorts #ai #videoai #aivideo #gen1 #beach
61	Bravo, ai stil! - Critici dure la adresa tinutei Constantiei! „1..cu parere de rau!"
62	Biden started STEAMING (Trump Obama) *AI voices* #meme #memes #ai
63	caocuongvu | Sự thông minh đến đáng sợ của A.I | Sự Thật Là Lùng Mà Bạn Chưa Bao Giờ Biết #shorts
64	Anitta feat Mc Danny e Hitmaker - AI PAPAI [Official Lyric Video]
65	Dj Brux Feat Marcos Robem & Dj Kalisboy - Ai (Afro House) [www.ditoxproducoes.com]

66	UMG Calls AI music “FRAUD” - Wants It Banned From Streaming Platforms | Joe Budden Reacts
67	【3 月 AI 新聞】AI 美女色情寫真 引作者版權風波
68	淺談 AI 繪畫技術的影響，盤點免費瀏覽 AI 繪畫作品的網站
69	【神回】国民的アニメの名シーンを AI で実写化したらヤバすぎたｗｗｗ【ツッコミ】【鬼滅の刃】【ワンピース】【名探偵コナン】【クレヨンしんちゃん】【ジブリ】【ドラゴンボール】【刀鍛冶の里編】【黒鉄の魚影】
70	this AI went too far
71	Snapchat’s New AI Chatbot Is HORRIFYING...
72	ഇന്ന് കണ്ണ് തുറക്കുന്ന AI-CAM
73	Using Ai to Modify a Classroom #shorts #ai #aiart #dalle #dalle2 #funny
74	Bravo, ai stil! - Raluca simte din nou mana stilistului in tinuta Sorinei!
75	AI manga × BuzzVideo
76	Taiwan surprisingly featured in Republicans’ AI-generated Ad #shorts @TVBSNEWS01
77	White House Reacts To AI Generated Attack Ads Following President Biden’s Reelection Announcement
78	i got an AI Girlfriend..
79	AI chatbots appear to have generated some reviews on Amazon: Report
80	Call to ‘put AI in charge’ of US after Kamala Harris’ latest speech
81	Is China worried about AI chatbots?
82	Big Tech isn’t the only way to invest in A.I., says Ark’s Chief futurist
83	I Got Turned into an A.I. TUTOR! Let me explain...
84	TikTok tests AI-generated avatars
85	RISS ACC1DENT?! SEBAB TAYAR PEC4H & JALAN LICIN!!
86	This AI Generated Pizza Commercial is Terrifying
87	it’s AI fr lol
88	Indian Women From Different States Imagined By Midjourney AI
89	Inilah Aplikasi AI yang Bakal Ngubah Hidup Kalian!
90	AI scav joins forces w/ g0at
91	कौन सी नयी जॉब्स लाएगी एआई [These are the new jobs AI will create]
92	3 A.I. Features in PowerPoint
93	Eminem but with AI (i’m not releasing it commercially obviously)
94	J’AI LAISSÉ MES SIMS SEULS PENDANT 24H
95	用 AI 给女神照片脱衣服 | 效果逼真😏不要做坏事哦 ＃ai 脱衣
96	AI News is Getting Out of Hand!
97	J’ai JOUÉ à l’avance a Zelda TOTK : Voici MON Gameplay ! 😍 (BOTW 2) - Partie 01 | EXCLUE FR
98	Why GPT-4 Might be the Most Dangerous AI Yet (Nobody is Talking about this!)

99	ISSEI funny video 😂😂😂 AI manga 🖼️
100	THIS ISN’T REAL - Ai Generated GoPro Footage #shorts #ai #aiart #gen2 #funny #gopro
101	ISSEI funny video 😂😂😂 AI manga × Time Warp Scan
102	Create A $1400 Per Day Passive Income Stream With ChatGPT AI Step-By-Step Guide | Make Money Online
103	How to stop AI going rogue
104	從電動車到 AI 無人機都靠「它」！揭全球難戒台積電背後秘辛！？ - 黃世聰 徐俊相《57 爆新聞》精選篇 網路獨播版-2100-1
105	TIYA KURANG AJAR DENGAN UMI AKU ?!! DAH MELEBIH SANGAT PEREMPUAN NI ??!!
106	When will artificial intelligence surpass human intelligence?
107	Amazing AI Filmmaking is Here: Gen-2 The Ultimate Cinematic AI
108	FIRST EVER AI GOLF LESSON - How to Fix Your Slice with the Driver
109	Comment j’ai perdu 2 millions d'abonnés… 😥
110	PRESIDENTS PLAYING MINECRAFT (Obama Trump Biden) *AI voice* #meme #memes #minecraft
111	人類よ、AI には気をつけろ
112	XXXTENTACION & Juice WRLD - Chamber Of Reflection (AI GENERATED)
113	Cuộc đua không ai muốn chiến thắng
114	Je l'ai piégée 😳
115	Bravo, ai stil! - Atitudinea, cheia succesului! Milena, “sexy, propusa si zambitoare!”
116	PRESIDENTS RANKING TOP 5 ANIMES (BIDEN OBAMA TRUMP) *AI voices* #meme #memes #anime
117	Game Theory: Which US President Is An EPIC Gamer? (AI Presidents)
118	Using Video Ai to Change a Basketball #shorts #ai #videoai #aivideo #gen1 #basketball
119	Bravo, ai stil! - Luiza a RABUFNIT dupa jurizarea Carinei: „A fost urat cum ai spus ca tinuta....”
120	CÁCH TRẢ LỜI CHO C U HỎI: “MÀY BIẾT BỐ M LÀ AI KHÔNG 🤣”
121	Microsoft, Alphabet, Amazon are ‘the leaders in a paradigm shift’ toward AI, strategist says
122	This Drake AI Song is Actually Crazy
123	Who’s Liable for AI Misinformation With Chatbots Like ChatGPT? | Tech News Briefing | WSJ
124	AI News! HUGE Chatbot Research, Viral AI Songs, Text to Video & More!
125	REPLIKA - A CyberS*xual DISASTER
126	AI Actors for Game Worlds
127	Lil Uzi Vert - Codeine Crazy (AI Cover)
128	Tóc Tiên - CÓ AI THƯƠNG EM NHƯ ANH (#CATENA) ft. Touliver (Official MV)
129	YOASOBI「アイドル」 Official Music Video

130	POURQUOI J'AI CLASH INOXTAG
131	SH - Ai ăn nhanh thì được chơi ipad || smart sister 🤣🥰🤧 #shorts
132	Why AI Is Critical to Deep Space Exploration | AI IRL
133	Snapchat needs to check their Ai
134	Can YOU Guess who this AI K-POP Idol is?
135	Alvis Has Some Bad News... (Made With AI VOICES)
136	Learning Romance from AI
137	Asking Ai to create a super villain for every state(part10) 🦹‍♀️☠️ #shorts
138	JOE WAS INVITED TO PLAY SOME MINECRAFT.. (ft OBAMA TRUMP) *AI voice* #meme #memes #minecraft
139	Bravo, ai stil! - Sorina, taxata la sange! Raluca: "Ne iei de prosti!" Tinute 1/1 cu ale lui Rux...
140	Kanye West & Drake - WAP (AI Cover)
141	AI reaching new heights daily #yunggravy
142	Can A.I. Build a Better FIFA Team than Me?
143	RNC slams Biden reelection bid with AI generated ad
144	Twitter's Ignorant Claims About Kizuna Ai
145	It's no longer possible to beat AI
146	Can AI help me get a Jimi Hendrix guitar tone?! #chatgpt #jimihendrix #guitartone
147	When gamers try the AI trend...
148	AutoPod AI Edited My Podcast Episode in 47 Seconds!!!
149	O que será que ta escrito ai? 🤔 #shorts
150	AI みたいな僕ら#shorts
151	GHOSTEMANE - AI (OFFICIAL VIDEO)
152	The Danger of AI | Scary Technology | Artificial Intelligence | Documentary
153	I talked to the AI Snapchat and its CREEPY..
154	A.I. and Stochastic Parrots | FACTUALLY with Emily Bender and Timnit Gebru
155	J'ai testé les SEEDS les plus Terrifiantes de Minecraft.. (grosse erreur)

COVID-19 vaccines:

Unique COVID-19 Vaccine Video ID	Video Title
1	FDA aims to ‘simplify’ your next round of COVID shots | NewsNation Prime
2	CNN host asks Djokovic if he regrets not getting Covid-19 vaccine
3	CDC approves new COVID booster shots. Who’s eligible and when they can get it?
4	Health Headlines: Lower Narcan costs, COVID-19 vaccine’s possible link to tinnitus | NewsNation Live
5	British regulators limit covid vaccines
6	Latest COVID booster guidance as new Arcturus subvariant spreads
7	Gravitas: China’s bogus vaccines: How China fooled the world
8	Monovalent Moderna and Pfizer-BioNTech COVID-19 vaccines no longer authorized for use in the US
9	Slowed down COVID vaccination drive, rising infections in India | DW News
10	Yikes: WSJ *Just* Exposed Covid Vaccines
11	Vaccine brain injury
12	#DeSantis Flip Flop On #Covid #Vaccine
13	Study compares COVID vaccine with immunity from infection
14	Joe Rogan: Pharma DESPERATE To Blame Vaccine Injury On ANYTHING ELSE; 16K+ Report Tinnitus After Jab
15	Ca Covid-19 Tăng, Tiêm Vaccine Xuyên Nghỉ Lễ, Tăng Cường Bảo Vệ Nhóm Nguy Cơ I SKĐS
16	Extra spring COVID-19 booster shot cleared for certain Americans
17	Covid vaccine and road accidents
18	What The COVID Vaccine Does To Your Body
19	NY woman says Johnson & Johnson COVID vaccine caused extreme reaction
20	GOP Launch Probe Into Allegations Of Biden Admin PRESSURING FDA On COVID Vaccines: Report
21	City workers fired for refusing COVID-19 vaccine to be rehired
22	First Covid Vaccine Given In U.S. As Distribution Begins | NBC Nightly News
23	Why Nazarin believes the COVID-19 vaccine is unsafe | Unvaccinated
24	New study released on 4th COVID vaccine dose
25	Inside the Lab That Invented the COVID-19 Vaccine
26	New COVID-19 vaccine schedule announced
27	CDC says tinnitus not linked to Covid vaccines
28	Pfizer Quietly Financed Supposedly Independent Groups Lobbying for Covid VACCINE MANDATES: Lee Fang
29	COVID-19: Moderna Biotech’s Evelyn Pang on vaccine complacency
30	Did Covid Vaccine Ad Go Too Far?

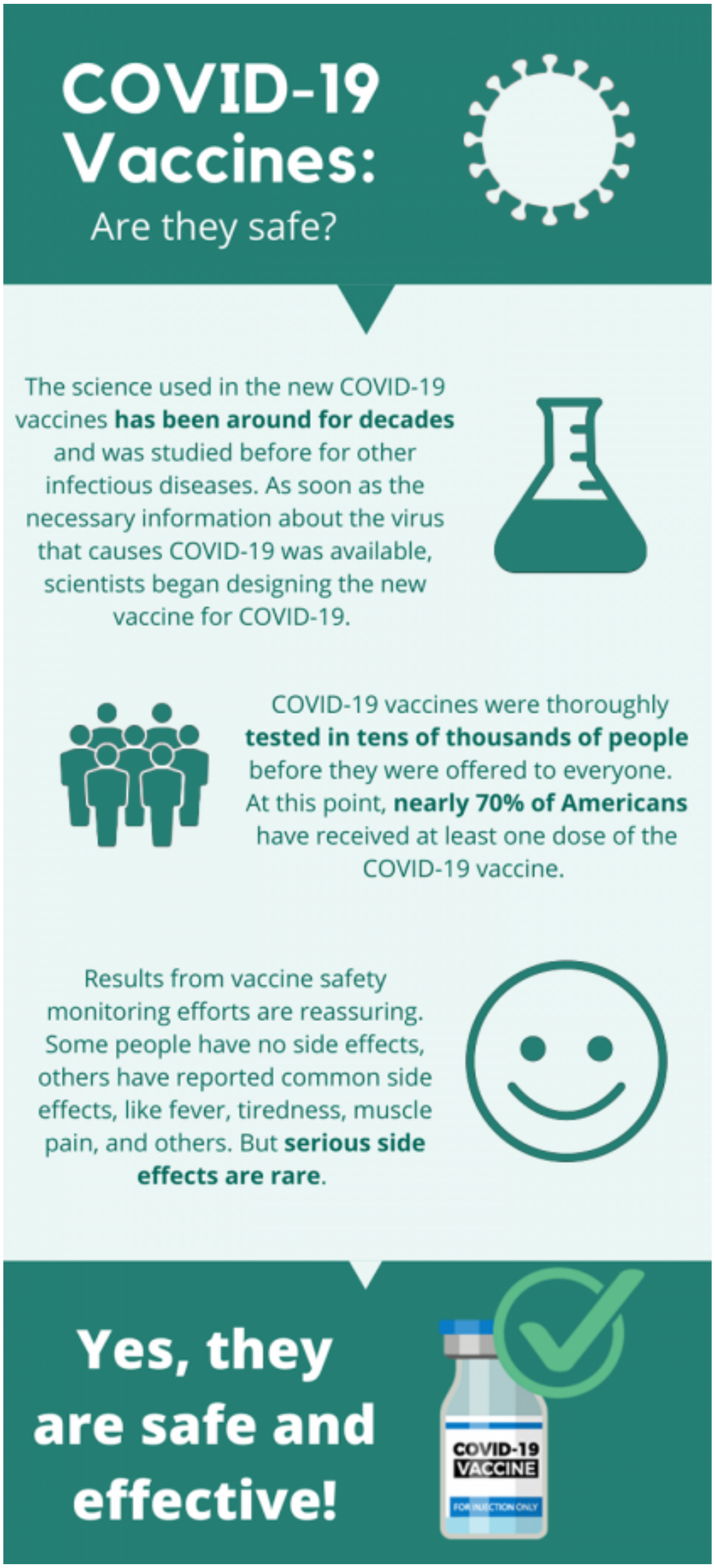
31	WHO covid vaccine revised roadmap
32	Florida Surgeon General EXPOSED For Altering Covid Vaccine Study
33	Justin Trudeau DENIES Forcing Canadians To Get The Covid Shot, Faces Backlash: Brie & Robby React
34	Novak Djokovic breaks silence over Covid vaccine refusal - BBC News
35	Doctor suspended over COVID vaccine views plans to sue Houston Methodist
36	DeSantis calls for grand jury to investigate Covid vaccines. Hear Fauci's response
37	More vaccinated deaths than unvaccinated deaths from covid (US)
38	International Monkey Trade Exploded Due To Covid Vaccine Development
39	Số ca COVID-19 gia tăng, nhiều người chủ động đi tiêm vaccine bổ sung
40	Vaccination v excess deaths, correlation study
41	Doctor Dies After Getting COVID Vaccine Post Vaccine Deaths - Doctor Mike Hansen
42	Gov. DeSantis calls out COVID-19 vaccine makers
43	Covid vaccine MHRA report: 'Yet again us conspiracy factualists are proved right' | Laurence Fox
44	New study suggests long COVID is identical to the flu - if vaccinated | 9 News Australia
45	Second COVID Vaccine Shot Side Effects
46	Nurse speaks out after being subject of anti-vaccine conspiracy theory
47	CDC Streamlining COVID-19 Vaccine Recommendations | NPR News Now
48	CDC finds possible safety issues with Pfizer COVID vaccine
49	How to Keep Young Kids Calm When Getting COVID-19 Vaccine
50	The future of COVID-19 vaccines
51	Republican War On Covid Science, Vaccines Shows In Higher Death Rate From Covid
52	Damar Hamlin's collapse sparks disinformation about COVID-19 vaccines
53	A Message for People Who Don't Want the COVID Vaccine
54	CDC Asserts #COVID Vaccine Safety as FDA Ponders Annual Booster - NTD Live
55	Medical experts reveal why some people need a fifth COVID-19 vaccine | A Current Affair
56	DeSantis's Surgeon General Altered Covid Study To Fit Anti-Vaccine Agenda
57	Andrew Bridgen Expelled For Covid Vaccine Holocaust Comments
58	Fact CheckED: Injuries due to COVID-19 vaccines? | Frontline Tonight
59	Covid Vaccine Side Effects + VAERS
60	Myocarditis and coronavirus vaccines: Explaining the rare side effect
61	Out There COVID-19 Vaccines - :30
62	How mRNA Medicines, Like the COVID-19 Vaccine, Work
63	COVID-19: Survey reveals vaccination complacency among public
64	April 19, 2023 ACIP Meeting - Welcome & Coronavirus Disease 2019 (COVID-19) Vaccines

65	Pfizer did not know whether Covid vaccine stopped transmission before rollout
66	Covid-19 Vaccine Skeptics Explain Why They Don't Want The Shot | NBC News NOW
67	Heart disease expert on COVID vaccines and variants
68	COVID-19: GP explains why she won't have vaccine but says she's not anti-vax
69	86% ca Covid-19 ở TP.HCM nhập viện thuộc nhóm nguy cơ, 30% bệnh nhân chưa tiêm vaccine | CafeLand
70	Dr. Fauci: We can do better than we're doing on Covid
71	Thêm Nhiều Ca Mắc Covid-19, Cần Tăng Cường Tiêm Vacccine Phòng Chống Dịch | SKĐS
72	Norman Swan explains the fifth dose of the COVID vaccine | 7.30
73	Children 5-11 begin COVID vaccine shots l GMA

Appendix D: Study C experimental conditions

Stimulus for the visual narrative (comics) condition:

Stimulus for the visual non-narrative (infographic) condition:



Stimulus for the non-visual narrative (written story) condition:

Two friends were sitting in the grass, having a chat. The woman sits on the grass and turns to her friend.

“I decided I’m gonna get the COVID vaccine tomorrow.” She stated.

“Are you sure?” the man asked, sounding in doubt. “I heard this was the fastest vaccine ever developed… **is it really safe**?”

“Yeah!” She replied with confidence. “**The science used has been around for decades**, and the vaccines have been thoroughly tested.”

Without looking at her, he replies: “What does that even mean?”

She sat up slightly and answered, “It means the vaccine was **tested in tens of thousands of people** before it was offered to everyone.”

“And at this point, **nearly 70% of Americans** already took it. A lot of them have no side effects,” she continued.

“No side effects?” he asked.

She took this opportunity to dive deeper. “Well, some people got mild symptoms like a fever, a bit of tiredness or muscle pain after taking it, but **serious side effects are rare**.”

“Huh, that doesn’t sound bad at all” he replied, finally looking back at her.

“I’m telling you, man - you should get it too.”

“Yeah, maybe I will.” he said, finally. It seems like his worries about the vaccine were unfounded after all.

Stimulus for the non-visual non-narrative (written facts) condition:

COVID-19 Vaccines:

Are they safe?

COVID-19 vaccines are the fastest vaccines ever created, but they were developed using **science that has been around for decades**. The technology used in the new COVID-19 vaccines has been studied before for other infectious diseases. As soon as the necessary information about the virus that causes COVID-19 was available, scientists began designing the new vaccine for COVID-19. COVID-19 vaccines were thoroughly **tested in tens of thousands of people** before they were offered to everyone.

At this point, **nearly 70% of Americans** have received at least one dose of the COVID-19 vaccine. Results from vaccine safety monitoring efforts are reassuring. Some people have no side effects. Others have reported common side effects after COVID-19 vaccination, like fever, tiredness, muscle pain and others. But **serious side effects are rare**.

Appendix E: Study C summary statistics

Descriptive statistics of participant characteristics and dependent variables by condition:

	Visual narrative (comic) (*n* = 83)	Visual non-narrative (infographic) (*n* = 81)	Non-visual narrative (story) (*n* = 82)	Non-visual non-narrative (facts) (*n* = 82)
Age	*M* = 44.17, *SD* = 13.96	*M* = 39.93, *SD* = 12.71	*M* = 42.56, *SD* = 15.05	*M* = 43.34, *SD* = 12.29
Gender				
Male	36.1%	37.5%	38.3%	37.0%
Female	63.9%	62.5%	61.7%	63.0%
Race				
White	54.2%	40.7%	48.8%	54.9%
Black	45.8%	59.3%	51.2%	45.1%
Political ideology (7 = very conservative)	*M* = 4.10, *SD* = 1.58	*M* = 4.08, *SD* = 1.68	*M* = 4.16, *SD* = 1.71	*M* = 4.20, *SD* = 1.60
Income				
Under $30,000	41.8%	52.5%	54.3%	40.5%
$30,000 – $69,999	43.0%	35.0%	32.1%	30.4%
$70,000 – $99,999	11.4%	6.3%	8.6	16.5%
Over $99,999	3.8%	6.3%	4.9%	12.7%
Education (college degree or higher)				
No college degree	26.5%	40.7%	35.4%	37.8%
Some college or 2-year college degree	54.2%	44.4%	46.3%	41.5%
4-year college degree or higher	19.3%	14.8%	18.3%	20.7%
Preexisting knowledge of COVID-19 vaccines	*M* = 2.75, *SD* = 1.34	*M* = 2.53, *SD* = 1.35	*M* = 2.60, *SD* = 1.39	*M* = 2.30, *SD* = 1.42
Science literacy	*M* = 4.87, *SD* = 2.08	*M* = 4.20, *SD* = 2.21	*M* = 4.49, *SD* = 2.37	*M* = 4.54, *SD* = 2.54
Factual knowledge of COVID-19 vaccine safety	*M* = 2.02, *SD* = 1.20	*M* = 1.89, *SD* = 1.19	*M* = 1.85, *SD* = 1.26	*M* = 2.07, *SD* = 1.29
Message elaboration of COVID-19 vaccine safety info	*M* = 4.53, *SD* = 1.02	*M* = 4.61, *SD* = 1.13	*M* = 4.44, *SD* = 1.06	*M* = 4.59, *SD* = 0.96

Zero-order correlations between participant characteristics and dependent variables:

	Factual knowledge about COVID-19 vaccine safety	Message elaboration of COVID-19 vaccine safety information
Age	-0.03	0.06
Gender (female)	-0.04	-0.04
Race (Black)	-0.13*	0.07
Political ideology (conservative)	-0.01	-0.06
Income	0.18***	0.02
Education	0.14*	0.01
Science literacy	0.36***	0.18***
Preexisting knowledge of COVID-19 vaccines	0.35***	0.17**

Note: Cell entries are zero-order point-biserial correlations (between the dependent variables and gender and race) and Pearson's product-moment correlations (between the dependent variables and all other participant characteristics).
*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.

Adjusted means and standard errors of dependent variables by condition:

Dependent variables	Visual narrative (comic)	Visual non-narrative (infographic)	Non-visual narrative (story)	Non-visual non-narrative (facts)
	Low science literacy			
Factual knowledge of COVID-19 vaccine safety	1.99 (0.20)	1.82 (0.17)	1.76 (0.18)	1.55 (0.19)
Message elaboration of COVID-19 vaccine safety info	4.52 (0.18)	4.46 (0.15)	4.40 (0.17)	4.26 (0.17)
	High science literacy			
Factual knowledge of COVID-19 vaccine safety	1.95 (0.17)	2.11 (0.20)	1.95 (0.18)	2.62 (0.18)
Message elaboration of COVID-19 vaccine safety info	4.51 (0.16)	4.84 (0.19)	4.49 (0.16)	4.96 (0.16)
	Total			
Factual knowledge of COVID-19 vaccine safety	1.97 (0.13)	1.96 (0.13)	1.85 (0.13)	2.08 (0.13)
Message elaboration of COVID-19 vaccine safety info	4.51 (0.12)	4.65 (0.12)	4.45 (0.11)	4.61 (0.12)

Note: All means shown above control for the following covariates: education, income, and preexisting knowledge of COVID-19 vaccines. Cell entries in brackets are standard errors.