Institut d'études politiques de Paris

ÉCOLE DOCTORALE DE SCIENCES PO

Programme doctoral de sociologie

médialab

Doctorat en sociologie

Retying the Knot

Four Articles on Science and Expertise in a Polarized World

Rodrigo REYES CORDOVA

Thesis supervised by

Achim EDELMANN, Associate Professor, Sciences Po Paris

Jean-Philippe COINTET, Professor, Sciences Po Paris

Defense on June 26th, 2026

Jury

Jean-Philippe COINTET, Professor, Sciences Po Paris

Achim EDELMANN, Associate Professor, Sciences Po Paris

Jacob HABINEK, Associate Professor, Linköping University (*reviewer*)

Tiffany MORISSEAU, Associate Researcher, Institut Catholique de Paris

Sally STARES, Associate Professor, LSE (*reviewer*)



Cover by Dakota Havard

People will try to understand me and to frame me in their words. They will seek truth. But the truth always carries the ambiguity of the words used to express it.

— *Frank Herbert, God Emperor of Dune*

Contents

Acknowledgments **10**

Introduction **12**

Science in modernity 13

Expertise 16

Education 17

Networked society 19

Political polarization 21

Research questions 22

Methodology 28

Chapter overview 29

1 Political Ideology-Driven Perceptions of Experts and Their Claims **32**

1.1 Summary 33

1.2 Introduction 33

1.3 Theoretical framework 36

1.3.1 From production to impact 36

1.3.2 Cultural cognition: A symmetrical account 37

1.3.3 Hypotheses 39

1.4 Methodology 40

1.4.1 Experimental design and procedure 40
1.4.2 Sample, measures, and analytical strategy 42
1.5 Results 44
1.5.1 Trust in experts 44
1.5.2 Perception of claims 46
1.5.3 Source specificity 48
1.6 Discussion and conclusion 50

2 Motivated Reasoning and Polarization Over Science in Europe 54
2.1 Summary 55
2.2 Introduction 56
2.3 Theoretical framework 58
2.3.1 Types of science 58
2.3.2 Motivated reasoning 59
2.3.3 Commitment and ability 61
2.3.4 Hypotheses 61
2.4 Methodology 63
2.4.1 Dependent variables 64
2.4.2 Independent variables 64
2.4.3 Analytical strategy 65
2.5 Results 66
2.5.1 Data and model evaluation 66
2.5.2 Political polarization 67
2.5.3 Motivated reasoning 68
2.6 Discussion and conclusion 71

3 Beliefs and Expectations in the Sharing of Expert Information 75

3.1 Summary . . . 76
3.2 Introduction . . . 77
3.3 Theoretical framework . . . 78
3.3.1 Relevance theory . . . 78
3.3.2 Information theory . . . 80
3.3.3 Beliefs and expectations . . . 81
3.4 Methodology . . . 83
3.4.1 Experimental design and procedure . . . 84
3.4.2 Sample, measures, and analytical strategy . . . 88
3.5 Results . . . 91
3.5.1 Overview . . . 91
3.5.2 Surprise and interestingness . . . 92
3.5.3 Sharing intention . . . 93
3.6 Discussion and conclusion . . . 96

4 Scientific Authority Cues Increase the Spread of Misinformation 100

4.1 Summary . . . 101
4.2 Introduction . . . 101
4.3 Results . . . 104
4.4 Discussion and conclusion . . . 114
4.5 Methods . . . 118
4.5.1 Observational study . . . 118
4.5.2 Experimental study . . . 121

Conclusion 124

References 131

List of Figures and Tables **154**

Appendices **156**

Résumé **262**

Resumen **273**

Acknowledgments

I want to thank Achim Edelmann and Jean-Philippe Cointet for their guidance through these three and a half years. Achim taught me to be more rigorous in my writing and coding, and encouraged me to design survey experiments. Jean-Philippe offered valuable guidance on STS literature and, most importantly, supported me at times when I could not see the finish line. I will be forever grateful to both for giving me the opportunity to pursue this PhD.

I thank Tiffany Morisseau and Stefaan Blancke for being part of my *comité de suivi*. Their guidance and suggestions were very appreciated and helped me steer this thesis in a more concrete direction. I also want to thank Sally Stares and Jacob Habinek for being part of my jury and for reading my work. I look forward to discussing it.

I am grateful to the many people I have worked with academically: Peter Achterberg, who was enthusiastic about writing together and who first introduced me to the public understanding of science field seven years ago; Ismail Harrando, with whom I worked closely over the last year and a half, provided constant moral support, and played a leading role in the computational aspect of this thesis; Noam Titelman for his feedback in the design of the "surprise" experiments; Roderik Rekker, with whom I discussed ideas at the very beginning of my PhD; and Kelly Christensen and Inès Girard, with whom I worked closely during the first months on data cleaning and preparation.

The médialab welcomed me with open arms, and I am especially grateful for the openness and friendliness of its members. I thank the methodological team—Benjamin Ooghe-Tabanou, Guillaume Plique, and Béatrice Mazoyer—for introducing me to new data collection and processing tools. I want to thank my fellow PhD candidates, including Maud Barret Bertelloni, Andrea Benedetti, Manon Berriche, Emma Bonutti D'Agostini, Duncan Cassells, Camille Chanial, Annina Claesson, Salomé Do, Héloïse Eloi-Hammer,

Tim Faverjon, Jean-Baptiste Garrocq, Valentin Goujon, Rubing Shen, Benjamin Tainturier, Jules Villa, Ömer Faruk Metin, Elise Noga, Pearl Pandya, and Armin Pournaki. The lunches and coffees we shared were welcome break in long working days.

I also want to thank Alex Kindel, for his generous feedback on my first article and support during the final year of this project. I am grateful to Pedro Ramaciotti Morales and Sylvain Parasie for their support and the pleasant exchanges we had. Even if brief, they were always appreciated. Finally, I thank Barbara Bender and Amélie Vairelles for helping me navigate the complexities of the French administration.

Without the support and conversations with friends from Paris, Tilburg, and Nürnberg, this thesis would not have been possible. Italo and Hiroki, thank you for making my years in Paris more enjoyable than they would have been otherwise. Dakota, I am especially grateful for our conversations, often sprinkled with a healthy dose of conspiracy ideation. They shaped how I approached this project and helped me reflect on facts and political correctness. Oleg and Franzi, you are role models and I feel fortunate to be close to you and the wonderful Nika.

I am extremely grateful to my parents, Monica and Carlos, who have always supported me without ever constraining me. From enrolling me in English lessons every Saturday at a very young age (something I found very annoying at the time), to your immediate support when I decided to migrate to Europe ten years ago. Your choices and encouragement laid the path that brought me here and this achievement is as much yours as it is mine. I am also thankful to my sister, Sofia, who is always there when I need her. Thank you for your emotional support, for listening when I have had a bad day, and for making me laugh, often with inside jokes nobody else understands.

Finally, thank you Marita. You are a caring and thoughtful partner, pushing me to be better, listening in moments of self-doubt, and quick to proofread my work. Your feedback was very helpful in the chapters that follow. Thank you for your patience and your love.

Introduction

At the beginning of the millennium, George W. Bush withdrew from the Kyoto Protocol to the dismay of the scientific community, in what was perhaps the most visible challenge of his administration to expert authority. Fast-forward twenty years and during the COVID-19 pandemic, governments that mandated wearing masks, social distancing, and vaccination were met with resistance and open contestation among the public. In the first case, political decisions trumped scientific opinion, and in the second, segments of the public rejected expert guidance. Both illustrate a common thread: scientific expertise is increasingly contested.

Whether heeded or not, our contemporary democratic societies rely heavily on scientific expertise to inform policy, economic regulation, and risk management. However, over the past decades, public disagreement over science has intensified, particularly when factual claims are used in governance. Even if such disagreement is often framed as disputes over facts, it instead reflects broader tensions about the exercise of power, the use of science in governance, and the authority of expertise. All of this unfolds against a backdrop of a networked society that has become increasingly polarized.

Drawing from existing literature, I outline four major developments over the past seventy years that help explain the current dispute over scientific expertise. First, the rise of technocracy has elevated science to a central role in governmental decision-making, generating resistance among groups who perceive expert authority as a challenge to democratic processes, as unelected experts increasingly shape policy. Second, the expansion of higher education has familiarized broad segments of the public with science, allowing citizens not only to rely on expert authority but also to challenge it. Third, the dawn of a new Information Age has given rise to the internet and social media, thereby facilitating the production and distribution of information that is catered to specific audiences. Finally,

this has all taken place amid increasing polarization in public opinion. Together, these developments have fostered a context where scientific expertise is in "crisis": opposition to technocratic governance and regulation can be articulated by citizens familiar with how science works, while alternative online sources of information provide the resources for contesting expert authority and advancing alternative truths, often in the service of political goals. To understand why scientific expertise has become a key point of contestation, it is first necessary to clarify what science is and the role it plays in modern societies.

Science in modernity

Science is highly regarded in modern societies. It is commonly understood as an organized cumulative method for knowing the world, through which knowledge is produced by direct experimentation and is ideally free from bias (Dooremalen et al., 2015). Science's privileged status (Cologna et al., 2025) rests on its capacity to generate truth claims that appear to speak for themselves (Latour, 1993), positioning it as our most powerful tool for producing knowledge (Shapin, 2008). By virtue of being perceived as relying on methodical processes, science is positioned as a rational, "objective arbiter of 'Truth'" (Lewis, 2010, p. 11) in our culture. This perceived capacity is a central source of scientific legitimacy and, in turn, confers authority to scientific experts.

The "matter-of-fact" character of science is central to modernity (Latour, 1993) and plays a key role in narratives of progress that modern societies use to distinguish themselves from those who came before. This boundary between the modern and the old hinges on the separation between the social and the natural, between facts and interests (Forman, 2007), framing science as apolitical and objective, and, in its most extreme form, as the only legitimate way of knowing (Sorell, 2013). Because science is framed as disinterested, the "scientific" label becomes a powerful source of legitimacy, granting authority over truth to those who can claim it (Shapin, 2008).

The idealized separation of science is what Latour describes as the cutting of the Gordian Knot in modernity. It refers to a division between knowledge and culture, where "on the left, they have put knowledge of things; on the right, power and human politics" (Latour, 1993, p. 3). By cutting the knot, truth is purified from values, allowing science to

appear neutral. Yet in practice, knowledge production is a social endeavor and therefore inseparable from its social context (Shapin, 1994). From this perspective, science is a large-scale *social* activity used to produce facts. It is embedded in "all the building blocks of what we term the *social*" (Jasanoff, 2004a, p. 3), inseparable from the social arrangements that shape it and susceptible to political and economic interests (Foucault, 2008; Jasanoff, 2004b). The Knot, then, is impossible to untie, and cutting it obscures the entanglement of facts and interests.

For more than a century, science has been entangled with the production of wealth and the projection of power. As Shapin (2008) argues, "the authority of science is increasingly based not on what scientists know but on what they can help to happen" (p. 443). A key source of science's authority, then, lies in the technologies, capacities, and interventions that its specialized knowledge makes possible. This capacity to "make things happen" is reflected in portrayals of science as the driving force behind some of the most important transformations since the nineteenth century, including "globalization, the networked society, [and] the knowledge economy" (p. 433).

This entanglement of science and power is not unique to contemporary science, as early scientific practices were already linked to private and state interests. In the late sixteenth and early seventeenth centuries, Galileo understood that, in order to secure the economic means and reputation necessary to pursue his work on celestial bodies, he had to seek patronage in the Venetian and Medici courts (Biagioli, 1990a, 1990b). In seventeenth-century England, authority in "truth-telling" was attributed to *gentlemen*, a status granted to men on the basis of their economic wealth and social standing (Shapin, 1994). Eighteenth-century botany developed in close connection to colonial state interests (Schiebinger, 2004), and the organization of science into specialized disciplines in German universities reflected growing state control and the demand for bureaucrats (Habinek, 2010).

The US in the twentieth century provides a more recent example of this instrumental dynamic. The entanglement of science with power intensified after World War II, when it became closely aligned with economic growth and state military power. During this period, scientific research was increasingly framed as a means for delivering industrial innovation, medical advances, and a specialized workforce by enhancing human capital

through education (Gauchat, 2023). Celebrated scientific advances such as the harnessing of nuclear energy and the space program came to symbolize the capacity of science to secure military dominance. In this context, technology, understood as the application of science in the service of productivity and power, came to matter more than the production of knowledge (Forman, 2007; Schnaiberg, 1980).

Since scientific authority in modernity rests on its capacity to "make things happen", it is unevenly distributed across disciplines and reflected in the cultural position of the natural sciences. These fields receive greater funding than those in the humanities and social sciences, because they are typically seen as capable of delivering goods desired by the military and by industry (Shapin, 2008). Their contribution to society is thus easily palpable and more readily measured in economic terms than that of the social sciences.

The uneven distribution of authority helps explain why there is disagreement over what makes a research field scientific. Despite the broad recognition of science's legitimacy, there is far less consensus about its defining features. There is no single scientific method but rather different sciences for different *uses* (Eyal, 2019). Fields such as physics can develop abstract theories and allow engineers to apply them in ways that contribute to the economy. However, when it comes to social theory and its application to policy and societal issues, not everyone would take it for granted that we are dealing with "science." The object of study in these fields is embedded in the world of values, societal institutions, and power relations, as evidenced by the development of the social sciences alongside the expansion of the nation-state (Smith, 2022) and their significant funding by the US in its pursuit of global influence (Nugent, 2010). As a result, the boundary between knowledge production and its practical application is less clear in the social sciences. Research fields in which the separation between facts and values is not easily maintained may see their scientific legitimacy contested.

In practice, as Shapin (2008) observes, "knowledge is just what works and [...] the scientist is understood as an aid to the technologist" (p. 444). In a similar vein, Latour (1993) describes modern science as "a mere intermediary allowing access to Nature" (p. 37), enabling its exploitation. Thus, scientific authority is not only epistemic but also deeply instrumental in modernity: it derives from the combination of producing truth claims and the material, economic, and political interventions those claims enable. Science may be

presented as disinterested, yet its authority is inseparable from how it is used.

The idealized separation of facts from interests is paradoxical, given the role science has come to play in modern governance. Over the past seventy years, scientific knowledge has been increasingly used in policy decision-making, with experts tasked to produce and lend credibility to the knowledge that justifies regulation.

Expertise

Expertise[1] refers to the possession of substantive, specialized knowledge in a domain. Experts are distinguished from others not by credentials alone, but by the internalized skills and ways of speaking acquired through socialization within expert circles and through involvement in knowledge production (Eyal, 2019). This process of socialization is historically situated: what counts as expertise depends on the institutions, technologies, and problems to which it is applied.

After World War II, the administrative state rose to challenge corporate domination through regulatory authority, and experts were increasingly involved in decisions at the public policy level (Gauchat, 2023). This growing reliance on experts rested on their perceived autonomy from the state and economic spheres and their ability to manage uncertainty (Eyal, 2019). As expertise became integral to the state and to newly formed bureaucratic agencies tasked with regulating economic activity, it became instrumental for governance. By drawing from the "matter-of-factness" of science, political choices could be presented as technically necessary rather than normative.

Relying on experts gave rise to what is often called *regulatory science*—"the kind of science that regulatory agencies rely on in seeking to protect human health, safety, or the environment, as well as to serve social goals such as nondiscrimination or environmental justice" (Dear & Jasanoff, 2010, p. 773). The contemporary dispute over scientific expertise is, in part, a consequence of this development, as scientific knowledge became increasingly embedded in regulatory bodies and entangled with policy, law, and governance. For some members of the public, the entanglement appeared as a moment in which "the

[1]Expertise is an encompassing term, as both scientists and laypeople may be experts in specific domains. However, when using the term *expert*, I refer specifically to *scientific* experts, reflecting science's position in modern societies as the dominant mode of producing specialized knowledge.

democratically elected and accountable politicians were co-opted by a cadre of scientific and technical experts" (Shapin, 2008, p. 435).

Segments of the public may have become anti-intellectual because of the combination of the involvement of non-elected experts in policy-making (Merkley, 2020; Motta, 2018) and the high-profile instances of conflicting regulatory rulings and failed economic predictions, including the 2008 financial crisis (Eyal, 2019). Importantly, this anti-intellectual sentiment is not rooted in ignorance, but is articulated through populist rhetoric that characterizes experts as using scientific knowledge to control "the people" (van der Meer & Hameleers, 2025).

From this perspective, the "crisis of expertise" does not seem to be one of *science*, but of the science used in policy-making. Evidence shows that while general trust in science remains high, disagreement emerges around *specific* issues where factual scientific claims are involved in regulation (Blank & Shaw, 2015; Funk et al., 2019; Pfänder et al., 2025). What is contested, then, may not be scientific methods or knowledge themselves, but the way claims are linked to authority and the use of such claims in the exercise of power.

For many years, the proposed solution to this problem was shaped by the knowledge-deficit model (Bak, 2001; Simis et al., 2016; Sturgis & Allum, 2004), which attributed distrust of experts and science to insufficient scientific knowledge and education. The assumption is that better-informed citizens are more rational and trust experts more. Yet empirical evidence suggests the relationship between knowledge and trust in experts and their claims is more complicated and, at times, paradoxical.

Education

Over the second half of the twentieth century, access to and enrollment in higher education expanded rapidly (Breen, 2010; Morrisson & Murtin, 2009; Schofer & Meyer, 2005). As a result, many citizens have become familiarized with science through formal education, to the point that scientific thought, once esoteric, is "now broadly spread through the educated world" (Shapin, 2008, p. 434). This educational expansion has increased scientific literacy and cultivated positive attitudes toward science among the public (Allum et al., 2008; Brossard & Nisbet, 2007; Noy & O'Brien, 2019).

Beyond deference, such familiarity with how science works endows the educated public with *science capital*, allowing "individuals to interpret scientific information that they come across in their everyday lives, make better informed choices (e.g., around personal health) and enable them to access, understand, respond to, and even contribute to shaping scientific developments in society" (Archer et al., 2015, p. 926). Yet science capital does not automatically translate into unquestioning deference. It can also be used to contest science and expert claims (Drummond & Fischhoff, 2017; Gauchat, 2012) because such familiarity with science includes exposure to the practices of experts, transforming the relationship between experts and the educated public.

As Eyal (2019) suggests, learning how experts produce knowledge may have a paradoxical outcome, making individuals *more* inclined to challenge expert authority. Once the public is aware of the assumptions, methodological choices, and uncertainties that are part of knowledge production, people may become vigilant and even skeptical of scientific claims. In this sense, the expansion of higher education has broadened not only trust in science but also the public's capacity to engage with it meaningfully and critically. As a result, since the last half of the twentieth century, education has both elevated the cultural authority of science *and* undermined the capacity of experts to utter uncontested factual claims.

Empirical evidence reflects the paradoxical relationship between education and the perceptions of experts and their claims. On the one hand, higher levels of education have been associated with increased trust in experts and science (Alper et al., 2023), and a narrower gap between trust in scientific methods and trust in scientific institutions (Achterberg et al., 2017), suggesting that education predisposes citizens to appreciate the virtues of science *in general*. On the other hand, higher levels of education and scientific literacy have been associated with increased disagreement with experts on politicized issues (e.g., Drummond & Fischhoff, 2017; Joslyn & Sylvester, 2019; Kahan et al., 2012). In these settings, when a *specific* topic is discussed, political motivations can outweigh deference predispositions and lead individuals to contest expert authority.

This apparent paradox can be understood as a transformation of trust in expertise. When knowledge is limited, individuals may rely on experts out of necessity (Baier, 1986). As familiarity with science grows and the object of trust becomes specific, educated in-

dividuals are better equipped to evaluate experts and their claims rather than relying on them automatically. In this sense, such vigilance does not imply a rejection of science in general, but rather a rejection of traditional scientific institutions as uncontested arbiters of truth (Achterberg et al., 2017; van der Meer & Hameleers, 2025). Consequently, individuals are no longer passive consumers of knowledge: they can identify relevant experts, evaluate competing claims, and seek out information beyond a single authoritative source (Shapin, 2008). In an age when information is abundant and easily accessible, they can "do their own research", drawing from alternative sources to evaluate and contest expert claims.

Networked society

In the last century, the expansion of higher education took place alongside a technological revolution. While information has been abundant and central to societies since, at least, the invention of the printing press, the volume of information available, the speed at which it circulates, and the capacity to process and recombine it have reached unprecedented levels. As Castells (2000) argues, contemporary societies everywhere have entered a new Information Age, defined by technologies centered "around microelectronics-based information/communication technologies" (p. 5). This marks a departure from hierarchical and centralized organizations to flexible, decentralized, and global forms of networked production and governance.

In this transformed networked society, information is continuously produced and reinterpreted. This condition accelerates knowledge production by enabling the real-time recombination of information drawn from multiple sources by multiple actors at any time. As a result, epistemic authority becomes increasingly diffused. Knowledge production becomes more distributed and interactive, incorporating the public and alternative communities in processes that were originally confined to traditional institutions.

The rise of the internet and social media best illustrates the idea of the networked society. Today, more people rely on the online environment for information (Wilson et al., 2020), exposing users to a myriad of information outlets (Klein, 2001; Nikolov et al., 2015; Phillips, 2010) and alternative sources of knowledge. Rather than resulting in

a standardized environment, the enhanced capacity for production and communication enables the targeting of information toward specific groups (Castells, 2000). In turn, traditional news media have lost their monopoly over the distribution of information (Van Aelst et al., 2025), and the authority of scientific institutions over the production of facts is increasingly contested (Gauchat, 2023).

While these technological developments have brought benefits—including improved access to health support (Moorhead et al., 2013), enhanced interaction between governments and citizens (Khan et al., 2014), and expanded participation in local and national public spheres (Ellison & Hardey, 2014; Lin & Kant, 2021)—they have also increased the visibility of misinformation. Even though misinformation is not a new phenomenon, its salience is tied to the current technological paradigm. In a global context where information production and consumption are central to power and culture (Castells, 2000), and information is available anywhere at any time, misinformation becomes a central problem in the digital public sphere.

Online platforms tend to reward emotionally charged and sensational content (Brady et al., 2020; Leach et al., 2026; Puryear et al., 2024), which can distort how users perceive what others believe. Shared content may be mistaken as being part of the social or scientific consensus, leading to biased beliefs that are widespread and resistant to corrections (Friggeri et al., 2014; Walter & Tukachinsky, 2020). Notorious cases, such as *Infowars* and *4chan*, show how online platforms can turn marginal claims into widely diffused narratives.

In the context of expert knowledge, these conditions may promote the spread of claims that deviate from scientific consensus but align with an individual's preexisting beliefs (Kata, 2010; Lewandowsky et al., 2013). For instance, claims denying the anthropogenic nature of climate change or the benefits of nuclear power may gain undue visibility and be seen as credible alternatives rather than marginal opinions. More broadly, in an environment where competition for attention is high, simple and emotionally charged messages are particularly effective, incentivizing political actors to communicate in increasingly polarizing terms.

Political polarization

Polarization among the public does not necessarily entail opinion radicalization; rather, it is a process of alignment and sorting in which attitudes become increasingly correlated with exclusive partisan identities (Baldassarri & Gelman, 2008). In the literature, the distinction is often made between *ideological* polarization, which refers to the divergence in issue attitudes across partisan lines, and *affective* polarization, referring to the emotional dislike and distrust of those from other parties (Iyengar et al., 2019).

For democratic societies, polarization in general is problematic because it simplifies politics and social life into terms of "us" versus "them" (McCoy et al., 2018), making individuals less critical of their own group's positions as they are more concerned with their negativity toward the out-group (Wilson et al., 2020). This phenomenon has been associated with unequal labor opportunities (McConnell et al., 2018), economic conditions (Frye, 2002; Pelizzo & Babones, 2007), and the deterioration of interpersonal relationships (Berntzen et al., 2024).

In the US, there is evidence of an increase in political polarization (e.g., DellaPosta, 2020; Finkel et al., 2020; Young et al., 2026). This pattern reflects the role of partisanship as a strong identity marker in the US context: it is typically acquired early in life, remains stable over time, and is made salient by a constant political campaign cycle (Iyengar et al., 2019). Specifically, rather than a broader radicalization of the public's opinion, US Americans have gone through a process of party sorting (Fiorina & Abrams, 2008): an increased alignment between their ideology and issue attitudes, on the one hand, and their party affiliation, on the other (Gauchat, 2023; Levendusky, 2009).

Polarization in the US is often attributed to political elites. Since the 1970s, party leaders have increasingly adopted more extreme policy positions, placing themselves at opposite sides of the liberal-conservative spectrum (Baldassarri & Gelman, 2008). As party elites have grown more polarized, voters find it easier to sort themselves along partisan lines, adopting clearer positions and drawing generalized inferences about the stances of the opposing side (Iyengar et al., 2019). Consequently, citizens increasingly overestimate the ideological gap between their views and those of the "other" (Ahler, 2014; Levendusky & Malhotra, 2016; Westfall et al., 2015), focusing on what they perceive

as extreme, morally appalling opinions of the opposing side. This creates a state of *false* polarization (Druckman et al., 2022) where the ideological gap between them is exaggerated. Such exaggerated distance, in turn, contributes to affective polarization and, over time, may lead to ideological polarization as people become more loyal to their own party (Wilson et al., 2020).

The problem is not exceptional to the US. Polarization has also emerged in democratic multi-party systems in the rest of the world. Cases such as Turkey, Hungary, and Venezuela show how political leaders can create and intensify disagreements, using the resulting crises to consolidate power (McCoy et al., 2018). Even in Western Europe, polarization has been linked to the rise of the populist radical right, which embraces a simplified narrative of "us," the people, versus "the elite" (Rekker, 2021). The rise of populism creates a two-way dynamic of distrust where populist voters see themselves as defenders of democracy, while non-populist voters often feel the same way (Davis et al., 2025). As a result, both sides feel disdain for one another, contributing to strong partisan cleavages across Europe (Gidron et al., 2023; Reiljan, 2020; Westwood et al., 2018).

In Asia, evidence of political polarization has been found in Taiwan (Huang & Kuo, 2022), South Korea (Cheong & Haggard, 2023) and Japan (Lee et al., 2022). In Africa, polarization has been documented in Kenya (Kibet & Ward, 2018), Ivory Coast, and the sub-Saharan region (Davis & Wilfahrt, 2025). In Latin America, political polarization had traditionally been associated with revolutionary movements, but now also occurs within the democratic sphere (Sarsfield et al., 2024).

Research questions

Taken together, four developments in the last seventy years have transformed the public's engagement with expertise: the increasing role of science in governance, the expansion of higher education, the rise of the internet and social media, and the intensification of political polarization. These developments have contributed to a context in which expertise is more politically salient, widely understood, and alternative sources of information allow for a more critical engagement with expertise than in the past.

Even if the context is sometimes described as a crisis of science's legitimacy (Gauchat,

2023; Saltelli & Funtowicz, 2017), this characterization may not be accurate. Science itself remains highly regarded (Cologna et al., 2025; van der Meer & Hameleers, 2025), but evidence suggests that specific claims and the authority of those who advance them are contested (e.g., McCright et al., 2013; Pfänder et al., 2025; Rekker, 2025). The crisis, then, is not of science as a method of knowledge production, but of expert authority and its applications. This raises the question of what experts are trusted, what claims are found credible, and what shapes those judgments.

Research on public perceptions of science has begun to move away from knowledge-deficit models toward approaches that combine sociological and social psychological explanations. Gauchat (2010) paved the way for an analysis of science attitudes by taking into account the different cultural representations that people have of science, what they consider to be science, and their evaluations of expert credibility. In a similar vein, Kahan (2012, 2013) proposes the theory of cultural cognition, which explains differences in the evaluation of scientific information in terms of the alignment between an individual's factual beliefs and their cultural values. Complementing these perspectives, McCright and Dunlap (2010) distinguish between types of science and experts that appeal to individuals of different political ideologies, rather than analyzing science in general.

Building on their work, I assess the contemporary "crisis of expertise" by moving away from evaluations of science toward the perceptions of expertise by asking:

What shapes public perceptions of experts and their claims?

Given its growing importance in shaping public opinion (Baldassarri & Gelman, 2008; Gauchat, 2023; Wilson et al., 2020), political identification is likely to shape how individuals evaluate science. Rather than treating polarization over science as the divide between its general acceptance and rejection, this thesis challenges the common characterization of liberals as "pro-science" and conservatives as "anti-science" (e.g., Drummond & Fischhoff, 2017; Gauchat, 2012; Mann & Schleifer, 2020; Oreskes & Conway, 2022). Such accounts overlook findings that most citizens express generalized trust in science (Cologna et al., 2025; Pfänder et al., 2025). Instead, research increasingly highlights political conflicts over the experts and claims that are seen as legitimate, particularly when these are associated with specific political positions and mobilized in policy-making decisions (Altenmüller et al., 2024; Rekker, 2021).

Recent literature has shifted the focus away from wholesale *rejection* toward *preferences* (e.g., Gligorić et al., 2025; Rekker, 2025; Shi et al., 2017), reflecting the regulatory applications of expert claims and the symbolic consequences of accepting them (Kahan et al., 2015). To capture these preferences, it is necessary to distinguish between forms of expertise by drawing from the theoretical distinction between two forms of science: *impact* and *production* science (McCright & Dunlap, 2010; Schnaiberg, 1980). Impact science refers to research that addresses the social, environmental, and economic consequences of industry and public policy, and is often used in support of market regulation. By contrast, production science refers to research that is aligned with industrial and commercial interests, helping innovation, increasing productivity, and prioritizing economic growth.

This distinction echoes broader cultural differences between the natural sciences and the social sciences regarding their perceived neutrality, usefulness, and their roles in legitimizing state intervention. It captures how different forms of expertise are entangled with political and economic interests, and why some experts and their claims may be contested by some political groups and accepted by others. In particular, it highlights one of the most polarizing domains of contemporary politics: economic redistribution and the size of the state (Webster & Abramowitz, 2017). Accordingly, I ask:

To what extent does political identification shape preferences for expertise along the distinction between impact science and production science?

If individuals hold preferences for different types of experts, the attribution of information to those experts may shape how they evaluate it. In contemporary online environments, individuals may be reluctant to invest cognitive resources to evaluate claims (van der Meer & Hameleers, 2025). Because of the high volume of information and intense competition for attention, people rely on heuristics to process information (Chaiken, 1980; Metzger & Flanagin, 2013). In these environments, experts and scientific sources may function as credibility cues, shaping perceived accuracy and the willingness to share information.

Additionally, science contestation operates at different interrelated levels (Rekker, 2021): skepticism may be directed at scientific knowledge in general, specific research fields and experts, or individual claims. Experts, then, may not solely act as credibility

cues, but are contested themselves. If the "crisis of expertise" does not concern science itself but expert authority, then claim perceptions may depend not only on their content but on who they are attributed to.

This implies that identical information may be evaluated differently depending on whether—and which—experts and scientific entities are cited as its source. If individuals hold political preferences for different sources, they may respond differently to the same claim when it is attributed to an impact or production expert. Disagreement may emerge not because the public disagrees about factual beliefs, but because they assign different levels of authority and credibility to different sources of information. Thus, when expert authority is politically contested, scientific entity cues may have different effects across political groups.

The role of experts as sources has rarely been studied in isolation. Existing research often evaluates experts at the same time as claims associated with their fields (e.g., McCright et al., 2013; Washburn & Skitka, 2018) or assesses claims without reference to an expert or field (e.g., Rekker, 2025). This leaves open the question of how experts independently shape the perception and sharing of expert information, and has consequences for the veracity of information that circulates online. Prior research shows that people often share misinformation without being aware of its accuracy, relying on heuristics (Ceylan et al., 2023; Pennycook & Rand, 2019; Vellani et al., 2023). As a result, in today's information context, if individuals rely on source cues, scientific authority may be used to lend credibility to *misinformation*.

Consequently, the next question shifts the focus from *which* experts and claims are favored to *how* expert information is evaluated, treating experts as sources of information and examining how their attribution determines the perception and sharing of their claims. I ask:

How does political identification shape the way expert sources influence the perception and sharing of expert claims?

Prior research has documented partisan differences in expertise perceptions and attributed them to motivated reasoning (e.g., Campbell & Kay, 2014; Hornsey, 2020; Pasek, 2018). However, the presence of these differences alone is not enough evidence for the

mechanism (Druckman & McGrath, 2019). Motivated reasoning requires an individual to have both the *motivation* to defend their group's beliefs and the *ability* to do so (Kahan et al., 2012; Kunda, 1990; Lewandowsky & Oberauer, 2016). As a result, partisan differences in expert perceptions may not only reflect motivated reasoning, but also differences in information environments or exposure to elite cues (Williams, 2023).

Empirically, studies typically operationalize motivated reasoning by examining interactions between political ideology and measures of cognitive capacity, such as numeracy, treating these as proxies for the ability to process information in directional ways (e.g., Chung et al., 2023; Persson et al., 2021; Strömbäck et al., 2024). Yet results from this approach remain inconsistent, perhaps because conceptualizing motivated reasoning as a function of cognitive capacity is too narrow.

Building on the idea that engaging with expert claims may be less about *knowledge* and more about familiarity with knowledge-production processes (Eyal, 2019), I focus on motivated reasoning by taking into account the interactions between political identification and both (*i*) partisan commitment and (*ii*) education. Here, education is not treated as a measure of cognitive capacity, but as an indicator of familiarity with science that may enable critical engagement with expertise. This approach indirectly captures both core components of motivated reasoning: motivation to defend group beliefs and the capacity to process information in ways that align with those beliefs.

Additionally, because the entanglement between science and politics is more salient in some research fields than in others, motivated reasoning may drive preferences regarding *specific* types of science. Therefore, I assess whether it operates differently across science in general, impact science, and production science. Accordingly, I ask:

Does motivated reasoning drive political preferences in perceptions of experts and their claims, and does this differ across general, impact, and production science?

The political preferences I examine, and the motivated reasoning that has been theorized to guide them, rely on the premise that the interaction between preexisting beliefs and information drives how expert claims are perceived and shared. This premise is consistent with confirmation bias accounts (Jonas et al., 2001; Meppelink et al., 2019) and

with relevance theory, according to which belief-confirming information is relevant and therefore more likely to be processed by an observer (Wilson & Sperber, 1995, 1998).

Yet focusing only on beliefs, understood as what individuals take to be true, may be too limited. In online environments, individuals often share information without assessing its accuracy, suggesting that factors beyond perceived truth may drive information sharing (Pennycook et al., 2021). In such contexts, emotional reactions play a key role in determining what information is shared (Berger & Milkman, 2010; Chen et al., 2022; Kim, 2015; Stieglitz & Dang-Xuan, 2013). *Surprise*, triggered by expectation violations, may be one of those emotional reactions. Surprising messages have been theorized to carry high informational value (Baldi, 2002; Shannon, 1948; Tellis et al., 2019), making these messages valuable to observers.

This may seem paradoxical: confirmation- and relevance-based accounts suggest that information is more likely to be shared when it aligns with preexisting beliefs, while information-theoretic accounts imply that surprising information, in the form of expectation violations, is particularly likely to be shared. These perspectives seemingly offer competing explanations regarding information sharing. The tension arises in part because existing frameworks rarely distinguish between different types of individual assumptions. Relevance theory, for instance, treats all successful interactions between information and assumptions—including reinforcements and contradictions—as relevant (Wilson & Sperber, 1995, 1998). Similarly, Goebel et al. (2024), using an information-theoretic approach, operationalize surprise as the result of belief contradictions rather than of expectation violations. The lack of distinction leaves unclear whether and how information interacts with individual assumptions to shape the perception and sharing of information.

To address this gap, I shift away from political explanations and adopt a more cognitive perspective, examining how beliefs and expectations, as two different types of individual assumptions, interact with expert claims. *Beliefs* concern propositions taken to be true (Egan, 1986), and *expectations* refer to predictions about what is likely to occur in the future (Barto et al., 2013). I distinguish between these two types and the ways they interact with information, rather than treating assumptions as a single set of mental representations.

Specifically, interactions between information and assumptions are conceptualized as

occurring through different channels: belief contradictions and reinforcements, and expectation violations and validations. This distinction allows for assessing whether these channels operate differently in shaping the perception and sharing intentions of expert claims. Accordingly, I ask:

Does the interaction between information and individuals' beliefs and expectations differentially shape the perception and sharing intentions of expert claims?

Overall, this thesis examines how perceptions of experts and their claims are shaped by political identification, motivated reasoning, and individual assumptions. It combines cultural, political, and cognitive explanations to study how expertise is perceived today and how expert information flows online. To assess these different explanations, I use several methods to study not only perceptions and information-sharing intentions, but also sharing behavior.

Methodology

I adopt a multi-method, quantitative approach to empirically address the research questions. These seek to (*i*) examine the *public's* perceptions of expertise, (*ii*) identify *causal* factors influencing how expert information is evaluated, and (*iii*) observe how such information is *shared* in online contexts.

First, by focusing on how individuals evaluate experts and their claims, this thesis addresses public perceptions. Particularly, the questions concerning the alignment between political identification and preferences for impact and production science require establishing whether these preferences exist among the public. Addressing this calls for data that are generalizable to a broader population. Therefore, the thesis uses national survey data to examine how perceptions of expertise are associated with political and individual characteristics, drawing on survey questionnaires as long-established tools for measuring public opinion in representative samples (Berinsky, 2017).

Second, the research questions addressing the roles of expert sources and individual assumptions concern causal relationships. These are difficult to identify through obser-

vational data alone and require experimental designs. I use online survey experiments to randomly present participants with messages that differ only in a single aspect, allowing me to attribute differences in perceptions and sharing intentions to specific manipulations (Gaines et al., 2007), thereby isolating the effects of expert sources and individual assumptions.

Third, one research question concerns how expert information is shared in online environments. To address it, the thesis incorporates social media data to observe how expert information is shared, which allows for the analysis of large-scale behavior and how information spreads in online social networks (Edelmann et al., 2020).

Together, these three data sources address the research questions by combining the generalizability of survey data, the ability of experiments to isolate causal relationships, and the capacity of social media data to capture online behavior. Nevertheless, these methods have important limitations. The survey and experimental data rely on self-reports and assisted questionnaires, and may be affected by social desirability bias. They also operationalize constructs such as trust, accuracy perceptions, beliefs, and expectations through numerical indicators, which cannot fully capture the cultural meanings and subjective experiences of engaging with expertise—something that qualitative approaches are better equipped to address (e.g., Harambam & Aupers, 2015; Wynne, 1991; Yearley, 2000). Despite these limitations, the methods used here are well-suited to the population-level and causal questions addressed across the chapters that follow.

Chapter overview

Chapter 1 addresses how political identification shapes preferences for expertise and the effect of expert sources on the perception of claims. To identify preferences and claim perceptions, it uses a pre-registered online experiment with US adults in which participants reported their trust in sociologists, environmental scientists, economists, agricultural scientists, and scientists in general. Additionally, they evaluated the accuracy of four claims randomly attributed to experts from impact fields (focused on the consequences of industry and policy), production fields (industry-focused), scientists in general, or no source (control condition).

Trust in experts was modeled as three latent factors: trust in (*i*) impact science (sociologists and environmental scientists), (*ii*) production science (economists and agricultural scientists), and (*iii*) scientists in general. These factors were analyzed using structural equation modeling (SEM), allowing for the simultaneous estimation of a latent structure and its relationship with independent variables. Perceived accuracy was modeled using robust linear models with clustered standard errors at the participant level to account for repeated measures. These analyses aim to uncover whether participants display political preferences for experts in impact and production science, and whether the same claim attributed to different experts, or no source, is evaluated differently.

Chapter 2 builds on this by extending the analysis of expertise evaluations to a European context and examining whether motivated reasoning drives preferences. Using survey data from 11 European countries, it focuses on motivated reasoning by operationalizing it with two complementary interaction terms: (*i*) between left-right identification and closeness to a political party, and (*ii*) between left-right identification and education. These interactions were tested as predictors of three latent trust factors: trust in science as a whole, impact science, and production science using SEM.

Chapter 3 shifts the focus from political identification to cognitive mechanisms driving expert information-sharing intentions. It uses a pre-registered online experiment with US participants in which beliefs and expectations were measured by asking them to evaluate the truth of statements and to assess the likelihood of future outcomes, respectively. Participants were then randomly exposed to expert claims that either contradicted or confirmed their beliefs, violated or validated their expectations, or combined these manipulations. They then rated their intention to share these claims, as well as how surprising and interesting they found them. These measures were analyzed using robust linear models, testing whether belief reinforcements and expectation violations differentially affect surprise, perceived interestingness, and willingness to share expert information.

Finally, Chapter 4 examines how Scientific Authority Cues (SACs) influence the spread of information online. It combines an observational analysis of millions of tweets, shared between 2010 and 2023, that link to fact-checked content with a pre-registered online experiment. In the observational study, SACs were identified using a large language model (LLM), and users were classified as Republicans or Democrats. Sharing behavior

was modeled as a function of the presence of SACs, fact-check ratings, and the original poster's political identification.

In the experiment, participants rated their willingness to share headlines randomly attributed to either scientific entities or to no source. Afterward, they indicated their perceived accuracy, as well as their trust in and familiarity with the entities mentioned in the attributed claims. Perceived accuracy ratings were specified as mediators using SEM, and sharing intentions were modeled using robust linear models.

The empirical work presented in these four chapters does not answer all questions regarding the "crisis of expertise." It addresses the factors that shape perceptions of experts and their claims, and how those claims spread in polarized online environments. Above all, this thesis pushes against knowledge-deficit models by highlighting the biases and preferences that individuals hold but may not always recognize when evaluating expertise.

Chapter 1

Political Ideology-Driven Perceptions of Experts and Their Claims

This chapter was published as a single-authored article[1] in *Public Understanding of Science*.

Abstract

US conservatives are often seen as distrusting scientists, and liberals as more trusting. This article examines how alignment between an expert's field and individual political ideology affects claim perceptions. US adults (N = 1,054) participated in a pre-registered online experiment, indicating their trust in five experts and evaluating the accuracy of four claims. Claims were attributed to experts from impact fields (focused on the consequences of industry and policy), production fields (industry-focused), scientists in general, or no source. Results show that liberals trust all experts more than conservatives and generally perceive claims as more accurate. However, the trust gap between liberals and conservatives is smaller for production experts. While no difference was found between the perceived accuracy of claims attributed to production versus impact experts, expert attribution increased some claims' perceived accuracy. These findings reveal some political-ideology preferences and that attributing a claim to an expert can improve its perception.

[1] https://doi.org/10.1177/09636625251372081

1.1 Summary

US conservatives are often seen as anti-science, and liberals as pro-science (Kennedy et al., 2022; Mann & Schleifer, 2020; McCright & Dunlap, 2011). This view overlooks recent scholarship showing that these political groups do not reject or support *all* science (Gligorić et al., 2025; Rekker, 2025; Shi et al., 2017). Rather, these positions depend on the research field or scientific claim under discussion. As such, people may have preferences for certain claims and experts, representing research fields, over others. This chapter examines how the alignment between an expert's field and an individual's political identity affects claim perceptions, testing whether the same claim may be perceived differently depending on the expert uttering it.

Results show that liberals trust experts more than conservatives and generally perceive claims as more accurate. However, the trust gap between liberals and conservatives is smaller for production experts. No difference was found between the perceived accuracy of claims attributed to impact versus production experts, although expert attribution increased the perceived accuracy of some claims.

These findings provide some answers to the research questions concerning preferences for expertise and experts as sources of claims. They suggest *some* political-ideology preferences, but these are not clearly aligned with the impact-production distinction as originally theorized. Experts seem to "stand on their own" when they are evaluated by US participants. In addition, there is evidence of political bias in the evaluation of expert information, but simply including an expert as the source of a claim can improve its perceived accuracy, regardless of the reader's political ideology. Together, the results indicate that political preferences exist but are weaker than expected, and that expert sources can improve the credibility of claims.

1.2 Introduction

In the United States, conservatives and liberals are divided in their views on science, with conservatives being less likely than liberals to believe that scientists act in the public's best interests (Kennedy et al., 2022). Research has shown that conservatives are more likely to

reject consensual scientific claims (McCright & Dunlap, 2011) and be more distrustful of scientists than liberals (Mann & Schleifer, 2020). Gauchat (2012) shows that this has not always been the case: for the 1974–2010 period, conservatives started with the highest levels of trust in scientific institutions and steadily declined to the point of being the group with the least trust. Over a similar period, Gauchat also found that liberals' trust in scientific institutions increased sharply, a pattern observed by Lee (2021) as well.

The growing distrust by conservatives is particularly concerning when it extends to expert claims intended to inform the public about scientific findings (Dahlstrom, 2021; Del Vicario et al., 2016). Distrust in such claims may lead the public to dismiss important information meant to mitigate risks, as seen during the COVID-19 pandemic. This distrust may be fueled by the spread of alternative-expert information (Oreskes & Conway, 2011) and facilitate the spread of misinformation that is anti-expert consensus (Havey, 2020). Such distrust negatively affects individuals, possibly leading them to engage in hazardous health-related decisions that go against public health guidelines[2], including ignoring the need for mask use or vaccine immunization (Barua et al., 2020).

While research evidence shows conservatives as anti-science, I move away from anti- and pro-science labels. Instead, I focus on how political ideology shapes perceptions of different experts and the fields they represent. People hold stereotypical beliefs about the political orientation of experts and their research fields (Altenmüller et al., 2024), and these can shape how they perceive claims. According to the cultural cognition model (Kahan, 2012), individuals are more likely to reject claims made by experts whom they perceive as politically misaligned.

This study distinguishes between two types of research fields: *impact* fields, assessing the social and environmental consequences of industrial activity and policy, and *production* fields, focused on promoting industrial efficiency and economic growth. Experts are treated as representatives of their fields to examine how this distinction affects the perceptions of claims. I hypothesize that an individual's political ideology shapes their trust in experts from these fields, as well as their evaluation of the claims those experts make. On the one hand, conservatives are more likely to distrust experts from impact fields and to view their claims with skepticism, as these fields are often associated with regulatory

[2]It is important to note that public health guidelines during the pandemic did not always align with scientific consensus, as seen with the closure of parks (Tufekci, 2020)

change and challenges to the economic status quo. In contrast, they favor experts from production fields and are more accepting of their claims, which align with beliefs about individual success, economic growth, and system stability (Jost et al., 2003, 2017). Liberals, on the other hand, are more likely to distrust experts from production fields and their claims, viewing them as prioritizing economic gain over social welfare. However, they trust experts from impact fields and favor their claims, as these align with their beliefs about the value of social justice, environmental protection, and progressive change (Sterling et al., 2019).

The hypotheses were tested in a randomized online experiment, where US participants rated their trust in experts from different fields and the accuracy of four claims. Although much of the literature focuses on trust, this study uses perceived accuracy as a proxy of trust in a claim. While trust is a multidimensional concept (Bauer & Freitag, 2017), accuracy plays an important role in shaping it in the contexts of risk assessment (Chryssochoidis et al., 2009) and belief formation (Wilson & Sperber, 1995). Moreover, this measure has been used effectively in experimental studies to assess how individuals perceive information (e.g., Altay et al., 2022; Epstein et al., 2023; Pennycook et al., 2021).

Results indicate that, although the ideological gap is smaller for production experts, liberals trust *all* experts more than conservatives. The perceived accuracy of claims is not affected by the expert to whom the claim is attributed, though in some cases, mentioning an expert as the source increases a claim's perceived accuracy. Moreover, while conservatives generally perceive claims as less accurate than liberals, they rate the taxation claim (which states that taxing the rich and corporations reduces government income) as more accurate than liberals do. The absence of clear ideological alignment of expert types—both in terms of trust and the perceived accuracy of claims attributed to them—suggests that participants are not consistent in their evaluation of experts from impact and production fields. However, conservatives' higher perceived accuracy of the taxation claim aligns with previous findings that show political preferences over specific claims (e.g., Nisbet et al., 2015; Rekker, 2025; Washburn & Skitka, 2018). Taken together, these results and prior research suggest that the political polarization over science is more pronounced at the level of claim evaluation than at the level of experts and their research fields. In addition, attributing a claim to an expert can improve its perceived accuracy.

I will proceed in four steps. First, I contextualize the current political polarization over science in the United States and introduce cultural cognition to describe the mechanisms behind it. Second, I briefly outline the study design, focusing on the experimental setup, as well as the study sample, measures, and analytical strategy. Third, I describe the results moving from the general picture of the responses to the effects of expert sources on perceived accuracy. I conclude by discussing the main findings, limitations of this study, and its broader implications.

1.3 Theoretical framework

The role of science has changed, and cultural cognition offers a framework for understanding how political ideology shapes perceptions of risk and expert claims. Drawing from the cultural cognition model, I discuss science polarization at the research field and claim levels. This discussion leads to the development of hypotheses, focusing on the importance of expert sources in claim perception.

1.3.1 From production to impact

The political polarization over science may find its roots in a societal transition. As societies have become increasingly aware of the environmental, societal, and health consequences of industrial activity, the role of science has changed (Beck, 1996; Beck et al., 1994; McCright & Dunlap, 2010). According to Beck, science has shifted from generating knowledge for industrial gains to understanding and dealing with the consequences of industrial activity. In this new stage, governmental institutions rely on experts to legitimize policies that mitigate the risks of industrial activity (Gauchat, 2010).

The change in the role of science is reflected in how research fields are perceived. Schnaiberg (1980) explains that some fields are typically associated with identifying and addressing the negative environmental impact of industry and technology, whereas others help industrial production by developing technologies that improve efficiency. Rekker (2025) extends this distinction from the natural sciences to social sciences to differentiate between impact fields as those typically seen as challenging individualism and evaluating policy effects on well-being, and production fields, as those typically seen as advancing

economic growth. Similarly, claims made by experts can be grouped based on their implications, either as questioning the impact of industrial activity and policies or as serving the industry.

In the United States, the impact-production distinction of research fields and claims aligns with political ideology. As science has increasingly taken on the role of identifying and mitigating the risks of industrialization, it has become associated with liberal ideology (Altenmüller et al., 2024). This perceived alignment between science and liberalism has led to the impression that liberals are pro-science and conservatives are anti-science (Gauchat, 2010; Lee, 2021). However, conservative skepticism of science is directed at institutions rather than science itself and may reflect discontent with the cultural authority of scientific institutions (Achterberg et al., 2017). In other words, while conservatives tend to value the rationality and truth-seeking nature of the scientific method, they distrust specific institutions (Aupers, 2012; Millstone & van Zwanenberg, 2000). Such distrust may trickle down to specific claims and research fields like environmental science, medical science, and sociology (Kennedy et al., 2022; Rekker, 2025).

Therefore, to understand the political polarization over science, it may be more useful to focus on the effect of political ideology on trust in research fields and perceptions of claims rather than anti- or pro-science labels. This approach shows how individuals may not be opposed to or in favor of science as a whole but be skeptical of fields and claims. The distinction of fields and claims, based on their relation to industry and their alignment with political beliefs, has received empirical support (e.g., McCright et al., 2013; Rekker, 2025; Shi et al., 2017), reinforcing the idea that political ideology shapes the public's perception of experts and their claims. The above has contextualized the political polarization over science and used the impact-production distinction to map conservatives and liberals in their preferences. However, why these preferences form is yet to be explained.

1.3.2 Cultural cognition: A symmetrical account

The cultural cognition model (Kahan, 2012) details why the public disagrees about societal risks along political lines. The model suggests that an individual's political beliefs shape risk perceptions by (i) attaching credibility to experts whom they perceive as sharing

their beliefs, and (*ii*) dismissing information that conflicts with their beliefs[3]. Through these mechanisms, people interpret risk information with regard to its social, symbolic, and policy consequences (Kahan et al., 2015).

While this model has primarily been used to explain risk perceptions, it can also help explain how people respond to experts and their claims more broadly (e.g., Kahan et al., 2011; Newman et al., 2018). Individuals associate experts and their fields with particular political orientations (Altenmüller et al., 2024). As a result, trusting an expert and accepting their claims may lead to an alignment or misalignment with one's political beliefs. Similarly, Rekker (2021) proposes that political ideology can motivate distrust in science at various levels. Among those, two are relevant for the present study: (*i*) research fields, and (*ii*) specific claims. Distrust at either of these two levels is motivated by a perceived misalignment between an individual's political ideology and the ideas and technologies typically associated with the expert and the claim. In other words, individuals evaluate claims based on who makes them and what they entail.

For example, when evaluating environmental scientists—a field associated with promoting the restriction of commerce and industry—conservatives are motivated to distrust them. Likewise, when confronted with a claim highlighting the anthropogenic nature of climate change, conservatives are inclined to dismiss it. Trusting the expert and accepting the claim implies acknowledging the risks of industrialization and criticizing economic elites, which could lead to regulatory measures and systemic change. These consequences conflict with beliefs about the value of economic individualism, free-market principles, deference to established hierarchies, and preserving the *status quo* (Koch et al., 2016), which correspond with hierarchical-individualist values (Kahan et al., 2015) and, in turn, with conservatism (Jost et al., 2003, 2017). From this perspective, conservatives tend to favor production fields and experts because they are perceived as aligned with their belief in the legitimacy of commerce, economic growth, and preserving the current system.

Liberals, in turn, hold beliefs about the importance of social justice, the need to combat inequality, the value of environmental sustainability (Sterling et al., 2019), and altering the status quo (Koch et al., 2016). These beliefs make them more receptive to experts from fields that critically evaluate the consequences of industrial activity and advocate

[3]Kahan (2012) describes three more prior mechanisms: identity-protective cognition, culturally biased assimilation of information, and cultural availability.

for regulation. Trusting experts from such fields—and accepting their claims that challenge industrial practices—reinforces their political beliefs. These beliefs correspond to egalitarian-communitarian values (Kahan et al., 2015), which frame the industry as a source of harm that justifies intervention. From this perspective, liberals tend to favor impact fields and experts because they are seen as aligned with beliefs in the need for systemic change and the regulation of industry in pursuit of societal and environmental goals.

However, just like conservatives, liberals are expected to reject expert claims when they conflict with their beliefs. Experimental and correlational evidence supports this in the context of nuclear power (Washburn & Skitka, 2018), the hydraulic fracking of natural gas (Nisbet et al., 2015), and claims about taxation policies and the safety of genetically modified organisms (GMOs) (Rekker, 2025). These findings suggest that both conservatives and liberals may favor research fields and expert claims they perceive as aligned with their beliefs.

By testing this framework, the present study has two aims. First, it seeks to confirm previous findings of ideological preferences for research fields by evaluating the trust individuals place in experts representing impact and production fields. Second, it examines how political ideology shapes the perceived accuracy of expert claims, depending on who the claim is attributed to. Within the US context, this leads to the formulation of four hypotheses.

1.3.3 Hypotheses

As previous research shows, conservatives and liberals should favor different research fields. In particular, liberals trust impact fields, while conservatives trust production fields. This trust is extended to the experts who represent those fields. Accordingly, I hypothesize:

H1 - Liberals trust experts from impact research fields more than conservatives.

H2 - Conservatives trust experts from production research fields more than liberals.

An expert's research field may influence how an individual perceives claims. The effect of expert sources on a claim's perception has rarely been researched in isolation. Experts

are often assessed alongside claims typical for these fields (e.g., McCright et al., 2013), or claims are assessed without mentioning an expert or research field (e.g., Rekker, 2025). To address this gap, this study focuses on how attributing the same claim to experts from different research fields influences an individual's perception of those claims. This approach acknowledges that the source of a claim plays an important role in its evaluation. Accordingly, I hypothesize:

H3 - Liberals view claims attributed to experts from impact research fields more favorably than claims attributed to experts from production research fields.

H4 - Conservatives view claims attributed to experts from production research fields more favorably than claims attributed to experts from impact research fields.

1.4 Methodology

The hypotheses were tested using linear regression models with data from a pre-registered[4] randomized online experiment with US participants.

1.4.1 Experimental design and procedure

The experiment was implemented using a Qualtrics questionnaire. First, participants completed the experimental task, where they were asked to give accuracy ratings to four expert claims in a random order. Each participant was randomly assigned to one of four conditions, manipulating the expert uttering the claim: Impact ("sociologists" or "environmental scientists"), production ("economists" or "agricultural scientists"), general (generic "scientists"), and control condition (no expert mentioned). These experts have been used in previous research as representatives of impact and production research fields (McCright et al., 2013; Rekker, 2025).

Participants read the four expert claims shown in Table 1.1. Three of these[5] were adapted from those used by Rekker (2025). For each claim, the expert source was ma-

[4]https://osf.io/9wnm2

[5]The taxation, GMOs and climate change claims were the ones adapted from Rekker (2025). The immigration claim was added to allow it to be attributed to both "sociologists" and "economists", as opposed to Rekker's gender differences claim.

nipulated based on the condition to which the participant was assigned, as detailed in Table 1. After reading each statement, participants answered the question: "To the best of your knowledge, how accurate is the claim made above?" They selected one option out of [Not at all accurate] - [Not very accurate] - [Somewhat accurate] - [Very accurate] - [Completely accurate]. Following Altay et al. (2022), the main dependent variable categories were inflated to avoid ceiling effects. As mentioned in the introduction, perceived accuracy was used as a proxy of trust in claims.

In the second part, participants were asked to rate their trust in the five experts mentioned earlier using four questions (McCright et al., 2013) presented in random order. For each expert, they indicated how much they trust them to (*i*) "create knowledge that is unbiased and accurate", (*ii*) "create knowledge that is useful", (*iii*) "advise government officials", and (iv) "inform the public on important issues". They selected one option out of [Completely distrust] - [Partially distrust] - [Neither distrust nor trust] - [Partially trust] - [Completely trust]. Participants were then asked about their political ideology: "When it comes to politics, do you usually think of yourself as ..." and given the option to select one out of [Extremely liberal] - [Liberal] - [Slightly liberal] - [Moderate, middle of the road] - [Slightly conservative] - [Conservative] - [Extremely Conservative]. As outlined in the theoretical framework, cultural cognition's focus is on political ideology, rather than partisanship, and, while they may have converged in the US in recent decades (Mason, 2015), they are different concepts (Goren, 2005). Consequently, participants also reported their partisan identification, along with their age, gender, and educational level. The full questionnaire, along with the informed consent, can be seen in the OSF materials.

Table 1.1: **Claims and their variations across experimental conditions.** Each participant evaluated all four claims and was randomly assigned to a condition for each claim.

	Social Science		**Natural Science**	
Condition	**Immigration**	**Taxation**	**GMOs**	**Climate Change**
Impact	"**Sociologists** claim that immigration does not increase crime."	"**Sociologists** claim that taxing the rich and corporations reduces government income."	"**Environmental scientists** claim that genetically modified organisms don't harm people's health."	"**Environmental scientists** claim that the Earth's climate is changing because of human activity."
Production	"**Economists** claim that immigration does not increase crime."	"**Economists** claim that taxing the rich and corporations reduces government income."	"**Agricultural scientists** claim that genetically modified organisms don't harm people's health."	"**Agricultural scientists** claim that the Earth's climate is changing because of human activity."
General	"**Scientists** claim that immigration does not increase crime."	"**Scientists** claim that taxing the rich and corporations reduces government income."	"**Scientists** claim that genetically modified organisms don't harm people's health."	"**Scientists** claim that the Earth's climate is changing because of human activity."
Control	"Immigration does not increase crime."	"Taxing the rich and corporations reduces government income."	"Genetically modified organisms don't harm people's health."	"The Earth's climate is changing because of human activity."

1.4.2 Sample, measures, and analytical strategy

Participants were recruited through Prolific between the 13th of May and the 21st of June 2024. Prolific's prescreening tools were used to ensure that all participants were at least 18 years old and resided in the US at the time of the study. Recruitment aimed for a balanced

sample in terms of gender, education, political ideology, and partisan identification. A total of 1,340 responses were recorded.

After removing participants who failed the attention check (15%), submitted the same response to all the claims (3%), and did not self-identify as liberals or conservatives (3%), an analytic sample of 1,054 participants and 4,216 responses to the accuracy questions remained. This sample size was determined in advance with a power analysis and is sufficient to detect small effects (the motivation is shown in Appendix 1.1). Among the participants, 49% identified themselves as male, while the rest (51%) identified themselves as female and other identities. Half (50%) were college-educated, and the rest (50%) were not; 18% were between 18 and 29, 39% were between 30 and 44, 28% were between 45 and 69, and 15% were 69 years or older.

To measure political ideology, I differentiate between participants who identified as [Extremely conservative], [Conservative], or [Slightly conservative] and those who identified as [Extremely liberal], [Liberal], and [Slightly liberal]. In the analytical sample, half identified themselves as conservative, and the rest as liberal.

The analysis proceeds in three steps: First, I examine whether liberals and conservatives differ in their trust in impact and production experts. To do this, I conduct a confirmatory factor analysis by grouping answers to the trust-in-experts questions into three latent factors. The first, *trust in impact experts*, comprises eight items related to environmental scientists and sociologists. The second, *trust in production experts*, includes eight items related to agricultural scientists and economists. The third, *trust in scientists*, includes four items focused on scientists in general. I then regress these factors on participants' political ideology. The factor analysis and regression were estimated simultaneously using one structural equation model (SEM), treating the trust responses as ordered categorical variables to account for their non-normal distribution.

In the second step, I examine differences in how participants perceive the accuracy of a claim attributed to experts from either impact or production fields. The dependent variable, perceived accuracy, is treated as an interval scale from 0 ("Not at all accurate") to 4 ("Completely accurate"). I regress this variable on political ideology (conservative versus liberal), the experimental conditions (production, general, and control versus impact), and their interactions. For all analyses that aggregate all claims, I use robust regression models

to account for repeated measures from the same participants; for claim-specific analyses, I fit simple linear regression models.

Finally, in the third step, I explore whether the specificity of the source affects a claim's perceived accuracy. To do this, I conduct exploratory regression analyses comparing the impact and production conditions (treated as a combined, *specific* reference group) to the general and control conditions.

For each step, I include models that control for the typical demographic variables—age, gender, and education (Bernerth & Aguinis, 2016). For these models, age was recoded with its minimum value set to 0. The reference categories were defined as females and non-binary for gender, and no college for education. Partisan identification was also included as a control, with Democrats as the reference category. The analysis script, available in the OSF materials, includes all control models and robustness checks.

1.5 Results

I begin by evaluating the differences between conservatives and liberals in their trust in impact and production experts (H1 and H2), using results from the SEM. Next, I examine how liberals and conservatives differ in their perceived accuracy of claims made by impact and production experts (H3 and H4). Finally, I explore whether source specificity affects the perceived accuracy of claims.

1.5.1 Trust in experts

The confirmatory SEM (shown in Figure1.1) showed a reasonable model fit: the scaled comparative fit index (CFI) and Tucker–Lewis index (TLI) were acceptable, but the scaled root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) were not. An exploratory SEM revealed a better-fitting five-factor model (see Appendix 1.2), suggesting that participants assess their trust in different experts independently, rather than along the impact-production distinction. The hypothesized model is interpreted here for three reasons. First, the CFI and TLI compare the hypothesized model to one where there are no relations between variables, while the RMSEA and SRMR

compare the hypothesized model to one where all variables are correlated Xia & Yang, 2019. The CFI and TLI are prioritized because the aim is to test hypotheses and not to describe the data perfectly. Second, the hypothesized model, containing three factors, is more parsimonious than the five-factor exploratory model. Finally, political differences remain consistent across both the theorized and five-factor models: liberals report higher levels of trust in all experts compared to conservatives.

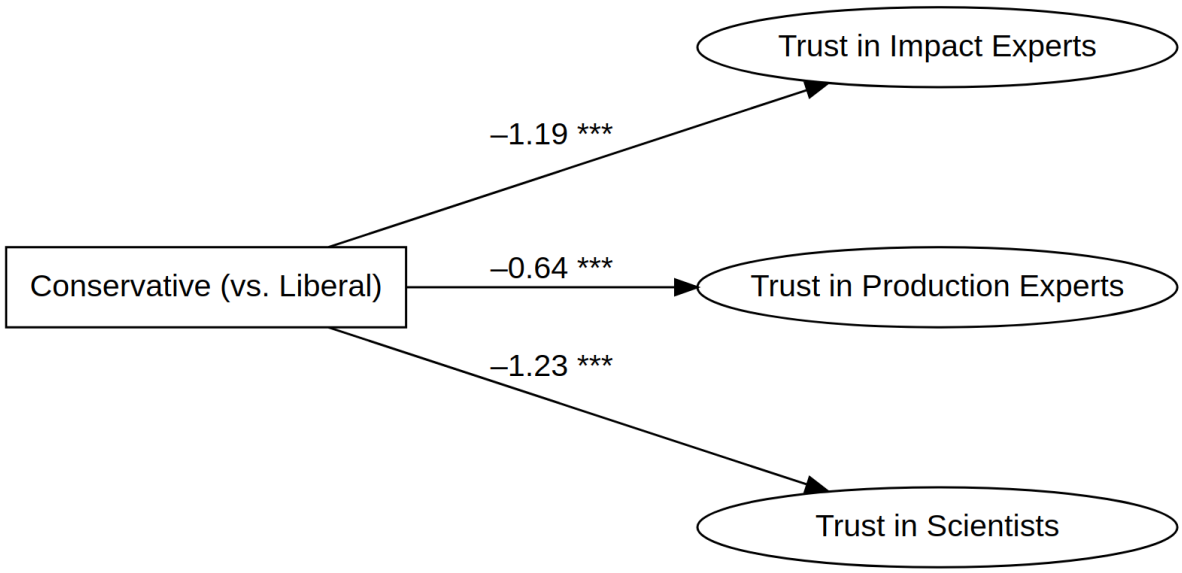


Figure 1.1: **Differences in trust in experts between conservatives and liberals.** N = 1,054, $\chi^2(184) = 5,058.45$, $p < 0.001$; CFI = 0.90, TLI = 0.90, RMSEA = 0.16, SRMR = 0.12. Factor loadings (shown in Appendix 1.3) and factor covariances are omitted for clarity. N = 1,054; *** $p < 0.001$.

The results of the hypothesized model are shown in Figure 1, and the standardized factor loadings and fit indices are shown in Supplement 3. Each trust factor shows strong internal consistency, with Cronbach's α values ranging between 0.91 and 0.95. All trust scores are positively correlated (between 0.69 and 0.79), indicating that higher trust in experts from impact fields is associated with higher trust in experts from production fields and scientists as well.

As expected in H1, liberals express substantially more trust in experts from impact research fields than conservatives ($b = -1.19$, $z = -18.61$): they report higher trust in sociologists and environmental scientists. Contrary to H2, however, conservatives express less trust in experts from production research fields than liberals ($b = -0.64$, $z =$

-10.38): they report lower trust in economists and agricultural scientists. In addition, conservatives show substantially lower trust in scientists than liberals ($b = -1.23$, $z = -16.77$). These differences remain statistically significant, though smaller, when adding all controls. Overall, the results provide evidence of ideology-based preferences for different types of experts: while conservatives are generally less trusting, the gap is smaller for production experts than for impact experts. This suggests that, although conservatives do not trust production experts more than liberals do, they are relatively more favorable toward them compared to impact experts. Still, support for grouping experts into impact and production categories is mixed. Participants seem to evaluate experts in isolation, and liberals consistently report higher levels of trust across *all* experts.

1.5.2 Perception of claims

Even if conservatives and liberals do not show the expected preferences, the source of a claim may still shape its perceived accuracy. According to Figure 1.2, claims are generally perceived as "not very" or "somewhat" accurate. Using the control condition as a baseline, on average, respondents perceived the climate change claim as the most accurate ($\bar{x} = 2.77$, SD = 1.25), followed by the GMOs ($\bar{x} = 1.85$, SD = 1.19), immigration ($\bar{x} = 1.58$, SD = 1.27), and taxation ($\bar{x} = 0.80$, SD = 1.00) claims. More variance is observed for the immigration and climate change claims, suggesting a lower level of consensus among participants.

Figure 2 and the regression models (Appendix 1.4) reveal differences in the way conservatives and liberals assess the accuracy of expert claims. Echoing the earlier analysis of trust in experts, conservatives generally perceive claims as less accurate than liberals do ($b = -0.64$, $t = -17.08$). This pattern holds for most claims: conservatives perceive the immigration ($b = -1.24$, $t = -18.62$), GMOs ($b = -0.53$, $t = -7.55$), and climate change ($b = -1.11$, $t = -18.25$) claims as significantly less accurate than liberals. The exception is the taxation claim, which conservatives perceive as marginally more accurate than liberals do ($b = 0.32$, $t = 5.24$). Politicization also varies per claim: the perceived accuracy of the immigration (adjusted $R^2 = 0.25$) and climate change (adjusted $R^2 = 0.24$) claims were more strongly shaped by ideology than the taxation (adjusted $R^2 = 0.03$) and GMOs claims (adjusted $R^2 = 0.05$). These political differences persist after adding control

variables to the models, except for the GMOs claim, where the difference disappears.

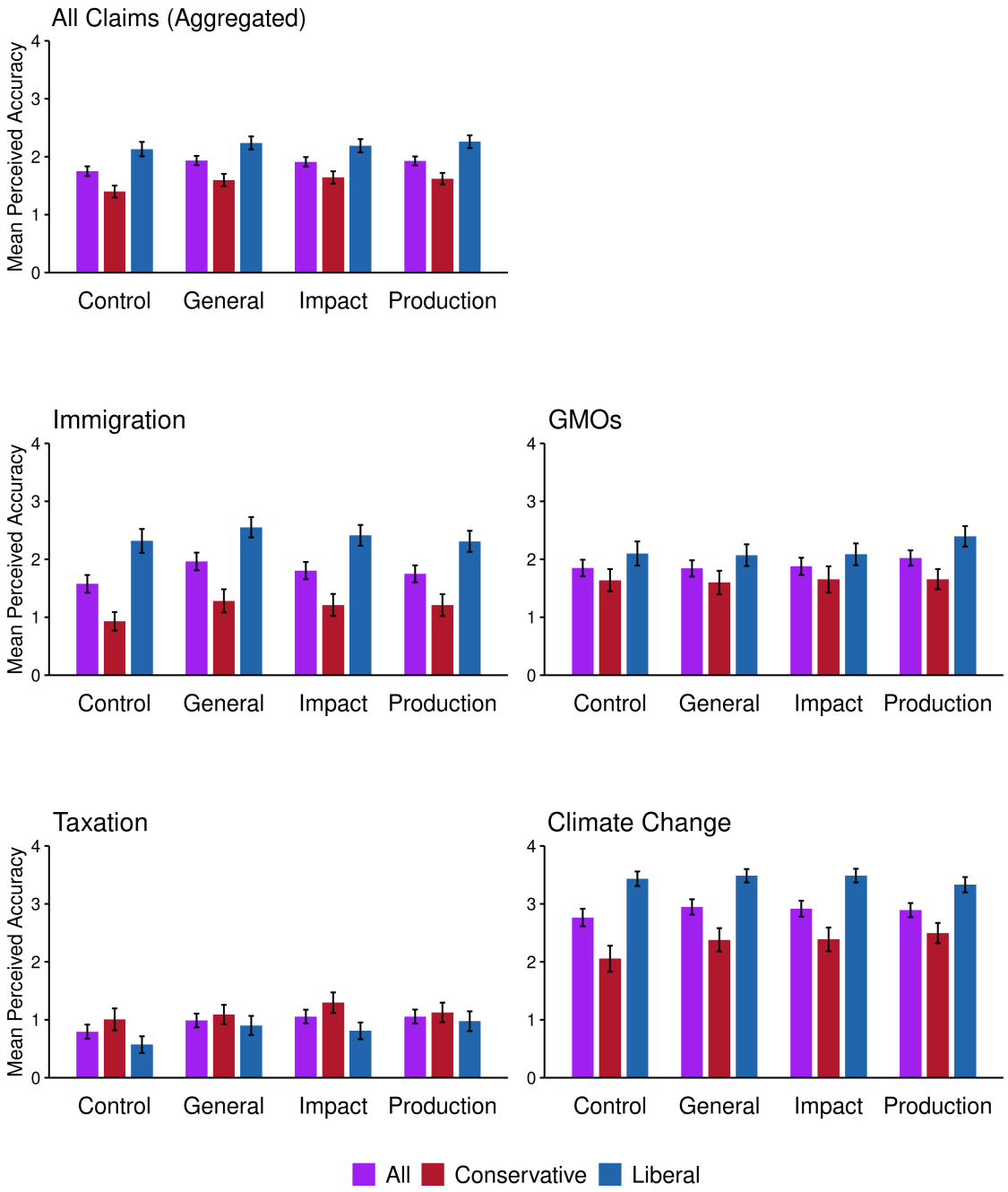


Figure 1.2: **Perceived accuracy of claims by experimental condition and political ideology, shown for all and specific claims.** Ratings are 0 ("Not at all accurate"), 1 ("Not very accurate"), 2 ("Somewhat accurate"), 3 ("Very accurate"), and 4 ("Completely accurate"). Error bars show 95% confidence intervals. N = 1,054.

Moving on to the analysis of the experimental manipulation, the results do not support H3 or H4. There is no significant difference in how participants perceive the accuracy of claims attributed to experts from impact versus production research fields (b = 0.02, t

= 0.44). The absence of this effect holds across all specific claims and does not change by political orientation[6]. In sum, although liberals and conservatives differ in how they perceive claims, both are similarly unaffected by whether the expert source is from an impact or production research field.

1.5.3 Source specificity

Regarding the effect of source specificity, the analyses show that attributing a claim to a specific expert improves its perceived accuracy compared to not mentioning a source. According to aggregated analyses seen in Table 1.2 (column M1), participants in the control condition, where no source is mentioned, perceive claims as marginally less accurate compared to mentioning a specific source ($b = -0.16$, $t = -3.31$). In contrast, claims attributed to "scientists" in general are rated similarly to those attributed to specific experts ($b = -0.01$, $t = -0.21$). This effect of source specificity on the perceived accuracy of a claim is consistent across ideological groups and persists after adding control variables.

The effect of source specificity and political ideology was explored separately for each claim as well. According to columns M1 in Table 1.2, overall, mentioning no source in a claim marginally deteriorates the perceived accuracy of the immigration ($b = -0.17$, $t = -2.07$), taxation ($b = -0.26$, $t = -3.49$), and climate change ($b = -0.18$, $t = -2.43$) claims compared to providing a specific source. When adding controls, the effect of source specificity persists for the taxation and climate change claims, but not for the immigration claim. Consistent with the aggregated analyses, there are no significant differences in perceived accuracy between claims attributed to scientists in general and those attributed to experts from either impact or production research fields.

[6]There is a significant interaction between conservatism and the production condition in the taxation claim, indicating that conservatives perceived the claim as more accurate than liberals in the impact condition. However, this difference disappears when control variables are added to the model.

Table 1.2: **Perceived accuracy of claims depending on the source: specific (impact or production), general, or control (no source)**. Columns M1 show overall differences, while M2 break them down by group. N = 1,054. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

	Aggregated		**Immigration**		**Taxation**		**GMO**		**Climate Change**	
	M1	**M2**	**M1**	**M2**	**M1**	**M2**	**M1**	**M2**	**M1**	**M2**
Intercept	2.25***	2.22***	2.41***	2.36***	0.89***	0.89***	2.21***	2.24***	3.48***	3.41***
Conservative (vs. Liberal)	−0.64***	−0.59***	−1.24***	−1.15***	0.32***	0.32***	−0.53***	−0.59***	−1.11***	−0.97***
Control (vs. Specific)	−0.16***	−0.09	−0.17*	−0.04	−0.26***	−0.32***	−0.08	−0.14	−0.18*	0.02
General (vs. Specific)	−0.01	0.01	0.13	0.19	−0.05	0.01	−0.12	−0.17	0.00	0.07
Conservative x Control	–	−0.14	–	−0.24	–	0.12	–	0.13	–	−0.41**
Conservative x General	–	−0.05	–	−0.12	–	−0.13	–	0.12	–	−0.14
Adjusted R^2	0.06	0.06	0.26	0.26	0.03	0.03	0.05	0.05	0.24	0.25
F-difference	–	–	–	1.08	–	1.05	–	0.38	–	3.85*

To clarify how source specificity interacts with political ideology, columns M2 for the taxation and climate change claims in Table 1.2 provide a detailed breakdown of the effects. As shown in Figure 1.3, when a specific source is mentioned, both liberals and conservatives show a similar increase in the perceived accuracy of the taxation claim compared to when no source is mentioned ($b = 0.12$, $t = 0.78$). However, the pattern changes for the climate change claim: conservatives perceive the claim as less accurate when no source is mentioned ($b = -0.41$, $t = -2.65$), compared to when it is attributed to an environmental or an agricultural scientist. In contrast, liberals are unaffected by source specificity ($b = 0.02$, $t = 0.32$), as attributing the claim to a specific expert does not change its perceived accuracy. This effect of source specificity among conservatives persists even after adding controls.

In sum, attributing claims to experts from impact (sociologists or economists) or production (environmental scientists or agricultural scientists) research fields increases their perceived accuracy. This effect is driven by the taxation and climate change claims. While

both liberals and conservatives react similarly to the taxation claim, their responses diverge for the climate change claim. Conservatives perceive it as more accurate when attributed to environmental or agricultural scientists, whereas liberals do not change their views based on the source. While these results stem from exploratory analyses, they highlight the importance of attributing claims to a specific expert source. Doing so enhanced perceived accuracy in two ways: uniformly across political ideology for the taxation claim, and differentially—only among conservatives—for the climate change claim.

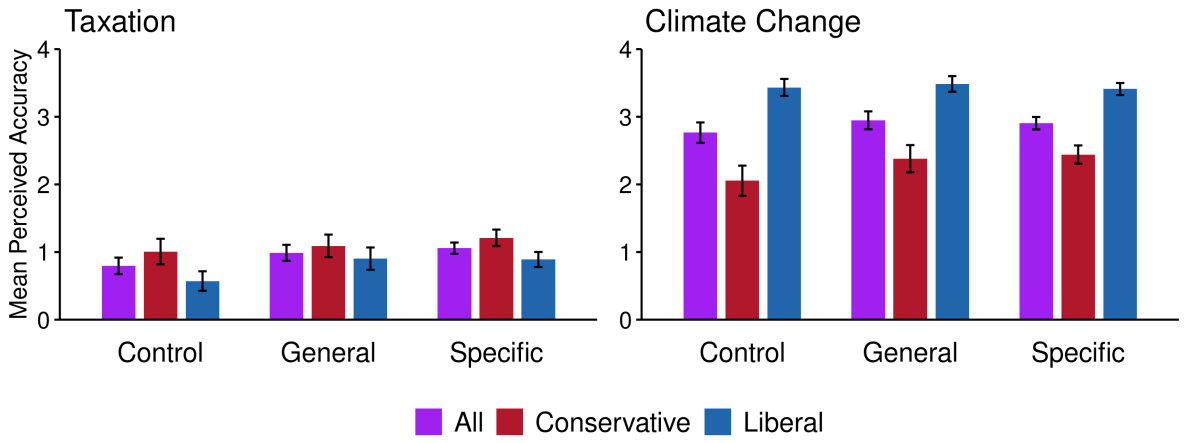


Figure 1.3: **Perceived accuracy of the taxation and climate change claims by experimental condition and political ideology.** Ratings are 0 ("Not at all accurate"), 1 ("Not very accurate"), 2 ("Somewhat accurate"), 3 ("Very accurate"), and 4 ("Completely accurate"). Source conditions are grouped into: control (no source), general (scientists), and specific (experts from an impact or production field). Error bars show 95% confidence intervals. N = 1,054.

1.6 Discussion and conclusion

In the political polarization over science literature, conservatives are often depicted as anti-science. However, their distrust is not directed at science but at the people conducting or using it. The cultural cognition model (Kahan, 2012) and Rekker (2021, 2025) highlight the role of expert sources in the perception of claims: if the expert's research field is perceived as misaligned with an observer's beliefs, their claims will likely be met with more skepticism. This led to the expectation that liberals trust experts from impact fields—those that focus on the negative impact of industrial activity and advocate for public welfare—more than conservatives. Conversely, conservatives trust experts from

production fields—those that support industrial activity and economic growth—more than liberals. This expert preference was also expected to influence the perceived accuracy of expert claims: liberals judge claims as more accurate when they come from impact fields, while conservatives judge claims as more accurate when they come from production fields. In this study, experts were treated as representatives of their research fields, allowing for an examination of how trust in fields influences the perceived accuracy of claims. These expectations were tested in a pre-registered randomized online experiment administered to US participants.

The analysis of trust revealed that there are political ideology-driven preferences for types of experts, but not the clear alignment that was theorized. Liberals trust experts from impact research fields (sociologists and environmental scientists) more than conservatives, as expected. Contrary to what was hypothesized, however, conservatives trust production experts (economists and agricultural scientists) less than liberals, though the gap is smaller than for impact experts. This indicates a relative preference for production experts among conservatives, but still no preference over liberals. Conservatives also report substantially less trust in scientists in general. In other words, while liberals trust *all* experts more than conservatives, political ideology may still play a role in differentiating trust across expert types, even if the pattern is overshadowed by conservatives' general aversion to experts. This general distrust may reflect the perception of scientists as being more liberal than conservative (Altenmüller et al., 2024). These results align with the findings of Gauchat (2012) but not with the findings of both McCright et al. (2013) and Rekker (2025). In addition, the findings provide mixed evidence for the theorized grouping of experts along the impact and production distinction. Participants evaluate experts individually, rather than grouping them, suggesting that experts "stand alone" rather than being seen as representatives of broader science categories.

The difference in trust between liberals and conservatives is reflected in the perceived accuracy of expert claims. Overall, conservatives judge the claims as less accurate than liberals, except for the taxation claim. The hypothesized interaction between expert source and political ideology was not supported. There is no difference in perceived accuracy between claims attributed to experts from impact versus production research fields, and this pattern holds across all claims and regardless of political ideology. In other words, attributing the same claim to one expert over the other does not affect its per-

ceived accuracy, suggesting that the impact-production distinction does not meaningfully influence how expert claims are perceived. Combined with the absence of clear clustering in trust around impact and production, this indicates that the grouping may not be salient enough for individuals to differentiate the content of a claim from its source. Alternatively, it may be that the distinction between impact and production fields does not map onto the liberal-conservative divide as clearly as previous research has suggested. Future research should examine which political values are stereotypically associated with different scientific fields and how these align, or not, with political ideologies.

According to the exploratory analyses, attributing a claim to an expert significantly improves its perceived accuracy compared to providing no source. This effect appears in the taxation and climate change claims and operates in two ways. First, for the taxation claim, the effect is uniform across political ideologies: both liberals and conservatives perceive the claim as more accurate when it is attributed to an economist or sociologist (even if liberals do not favor the claim). Second, for the climate change claim, the effect is ideology-dependent: conservatives perceive the claim as more accurate when it is attributed to agricultural or environmental scientists, compared to when no source is mentioned. Liberals, by contrast, are unaffected by the presence or absence of a source. These findings suggest that while participants may not respond differently to impact or production experts, they do respond positively to the presence of a specific expert source. This effect is substantial for conservatives in the case of the climate change claim—an important finding given the conservative elite's opposition to it (McCright & Dunlap, 2010). Overall, mentioning expert sources may improve the credibility of claims even if that claim conflicts with the observer's political ideology

This study has four limitations and an important clarification. First, the sample is not representative of the US population; instead, it was balanced across gender, education, political ideology, and partisan identification. This approach enhances the internal validity of the study, but limits the generalizability of the findings. Future research could use larger and representative samples. Second, participants may not have had strong views about the claims they saw, most of which were adapted from Rekker (2025). A descriptive analysis showed that 74% of participants view taxation as relevant, followed by immigration at 73%. In contrast, 47% and 43% considered biodiversity (related to climate change) and GMOs relevant, respectively. Future studies could select claims more

salient to the US audience. Third, the claims were used in previous research to represent impact and production science, but are limited in their combinations of research fields and claims. Future research could consider, for example, how participants respond to claims by environmental scientists that emphasize industrial innovation, such as advances in green energy efficiency. Fourth, this study could not account for different consensus levels within claims due to their limited number. Future research could include more claims and account for these differences by, for example, fitting multilevel models. Finally, this study did not assess whether people can distinguish between true and false information or whether they agree with the expert consensus. The claims, particularly the GMOs and taxation ones, may be contested even among experts. The goal is not to determine whether conservatives or liberals have greater epistemic knowledge but to reveal differences in the trust in experts and the accuracy perceptions of claims.

In synthesis, US participants show some political ideology-driven trust in experts, but not the clear alignment with the impact-production distinction that was originally predicted. The findings suggest that trust may depend more on individual experts than on broad expert categories, as evidence for a systematic grouping of experts in participants' trust evaluations is mixed. Liberals trust all experts more than conservatives and perceive all claims as more accurate, except for the taxation claim, which conservatives rate as more accurate than liberals. When evaluating the accuracy of claims, there are no consistent political ideology-driven preferences for expert sources, but attributing some claims to a specific expert increases their perceived accuracy. These findings highlight the potential for political bias to shape perceptions of expert information. Taken together with prior research, the results suggest that political polarization over science is more pronounced at the level of claim evaluation. Recognizing this allows future research to focus on the claim level. Finally, since attributing claims to experts may improve their perceived accuracy, science communicators and policymakers can better frame expert information to improve its reception across both political camps.

Chapter 2

Motivated Reasoning and Polarization Over Science in Europe

This chapter, developed with conceptual input from Peter Achterberg, is under consideration for publication as a first-authored article.

Abstract

This study uses survey data to examine political polarization in public trust in science across 11 European countries (N = 5,410), focusing on the role of motivated reasoning. We distinguish between impact science, which tackles the harmful consequences of industrial activity and policy, and production science, which supports industrial development. In line with previous findings, left-leaning Europeans trust science as a whole and impact science more, while right-leaning Europeans prefer production science. We operationalize motivated reasoning through the complementary interactions of political identification with party closeness and education. Our results suggest that in Europe, motivated reasoning amplifies ideological differences only regarding impact science. However, polarization over science as a whole and production science may only be driven by partisan commitment. We conclude that European skepticism towards science that challenges industrial interests and supports regulatory intervention is better understood as an expression of politically motivated reasoning than as a knowledge deficit.

2.1 Summary

This chapter continues to use the impact-production distinction, shifting the focus to assessing motivated reasoning as a driver of science polarization in Europe. The increasing alignment between trust in science and ideology (Gauchat, 2012, 2023; Suhay & Druckman, 2015) has been frequently attributed to motivated reasoning (Druckman & McGrath, 2019). However, this attribution has relied on evidence of left-right differences alone. In this chapter, I point out that identifying motivated reasoning in science polarization requires observing the combination of partisan commitment and ability (Kunda, 1990; Leeper & Slothuus, 2014). Together, these provide (*i*) the motivation to defend one's political group positions and (*ii*) the ability to reason away dissonant scientific information.

The findings show that Europeans align their preferences with the impact-production distinction and that both closeness to a party and education amplify left-right differences *only* for impact science. This indicates that motivated reasoning, requiring both partisan motivation and the ability to critically engage with information, drives ideological differences among Europeans specifically toward research that challenges industrial interests and supports regulatory intervention. By contrast, polarization over science as a whole and production science is driven only by partisan commitment, with no bigger ideological differences among the more educated.

These results address the research questions concerning preferences for expertise and motivated reasoning. They suggest that Europeans differ from US participants in that their preferences align more closely with the impact-production distinction: the further to the right, the lower the trust in impact science, and the further to the left, the lower the trust in production science. Moreover, skepticism toward science may reflect motivated reasoning primarily for impact science, rather than production science or science in general. This, in turn, is consistent with the idea that contestation and motivated reasoning are more likely to occur when the entanglement between expertise and politics is more salient, as is the case with impact science.

2.2 Introduction

In recent decades, science has become increasingly politicized (Suhay & Druckman, 2015), as trust in research fields and scientific claims is shaped by political identification (Gauchat, 2012, 2023; McCright et al., 2013; Schrøder, 2023). The alignment has been pronounced in the US, where partisan divisions gained attention and have intensified since *The Republican War on Science* (Mooney, 2007). This is evident under the current Trump administration, which has defunded research dealing with gender identity, disability, public health, and the environment (Turnley, 2025).

While the evidence is clear on this issue of science polarization, recent work highlights the importance of distinguishing between impact science and production science (e.g., McCright & Dunlap, 2010; Rekker, 2021, 2025). Whereas the former seeks to identify and mitigate the negative consequences of industry, technology, and public policy, the latter focuses on developing technologies that enhance industrial efficiency and promote economic growth. On the one hand, conservatives should be critical of impact science, as they attempt to protect the industrial capitalist order. By contrast, production science is seen as less problematic by conservatives, as it supports industrial progress. On the other hand, liberals favor impact science, as it challenges the harmful effects of production science on the environment (Schnaiberg, 1980). Hence, there may be different patterns of science polarization depending on the types of research fields and claims being evaluated.

One key question is who is most susceptible to such patterns. In this paper, we test two complementary factors that make up motivated reasoning: the motivation and the ability to reason away dissonant scientific information. The motivation to defend one's political group stems from partisan commitment. This relationship is grounded in social identity and coalitional psychology (Schulz-Hardt et al., 2000; Van Bavel & Pereira, 2018), suggesting that individuals with stronger party ties are more motivated to defend their party's positions. The ability to dissect dissonant information is made possible by background knowledge. Research has shown that political polarization is more pronounced among the highly educated (Gauchat, 2012), those with greater numeracy (Kahan et al., 2012), and the scientifically literate (Drummond & Fischhoff, 2017). These findings suggest that more knowledgeable individuals can process information in ways that serve

their political beliefs. Thus, rather than only examining differences between the left and the right, we assess how these differences vary across levels of partisan commitment and ability. If attributed to motivated reasoning, polarization should be greater among those who are both committed to their group and able to construct identity-consistent justifications.

In short, this paper has two main objectives. First, by distinguishing between general trust in science, trust in impact science, and trust in production science, we examine political polarization across these orientations. Second, we investigate whether political polarization over science can be attributed to motivated reasoning: more pronounced among individuals with stronger ties to a particular political party and among the more educated.

Our first objective partially overlaps with recent work by Rekker (2025), who examined political polarization over science across different levels of trust using the same data as we do. The author found that right-leaning Europeans trust science as a whole less than left-leaning ones and that, after accounting for this generalized trust, both sides tend to be skeptical towards dissonant research fields and scientific claims. We build on these findings by grouping the trust items into theoretical science orientations, distinguishing not only between trust in science as a whole and the rest, but also between *impact* and *production* science. As a result, part of our analysis revisits the patterns observed by the author, noting that we also test two additional groupings.

We address our objectives in a European context by fitting Structural Equation Models (SEMs) to CROss-National Online Survey-2[1] (CRONOS-2) data, fielded in 2023 across 11 countries. This focus is important, as most research on political polarization and polarization over science has centered on the US (Boulianne, 2019; Cologna et al., 2025). Yet, the problem is also evident in Europe (Wagner et al., 2014), where the rise of the populist right has been linked to distrust in science (Rekker, 2021) and skepticism toward environmental protection policies (Huber, 2020). After elaborating on our framework and hypotheses below, we introduce our data and measures, present the results, and end with a discussion of our findings.

[1]https://doi.org/10.21338/ess10e03_2

2.3 Theoretical framework

In the following section, we contextualize political polarization over science, arguing that individuals' political identification shapes whether they favor science linked to industrial activity and economic growth or science focused on environmental protection and public welfare. We then explore how motivated reasoning may drive science polarization via two complementary conditional factors.

2.3.1 Types of science

Political polarization over science can be understood as part of a broader societal transition. During the second half of the twentieth century, the state emerged as a challenger to the industrial elite (Gauchat, 2023), with science playing a central role in legitimizing the state's corporate regulation (Gauchat, 2010). Before this development, science served the industrial order by developing technologies and enhancing industrial output (Beck, 1996; Beck et al., 1994; Shapin, 2008). This orientation—called *production* science (McCright et al., 2013; Schnaiberg, 1980)—conceives nature as external and subordinate to society, to be exploited and controlled for economic gain (Latour, 1993).

Production science created new societal risks, including climate change and public health crises. In response, science takes on a more reflexive orientation, shifting toward *impact* science, focused on assessing and mitigating the consequences of industrial development and public policy (Rekker, 2025). This transformation may help explain why science is increasingly associated with regulation (Gauchat, 2012) and why it is stereotypically perceived as liberal (Altenmüller et al., 2024). Broadly, it reflects a shift in its political position: from being aligned with corporate interests to increasingly aligned with the state.

In response, conservatives have attacked scientific institutions for decades (Dunlap, 2014; Lee, 2021; McCright & Dunlap, 2010). This is not a rejection of science as a whole (Achterberg et al., 2017), but a distrust aimed at impact science. Such skepticism reflects the tendency of impact science to favor regulation over economic growth and free-market principles, which can conflict with conservative values (McCright & Dunlap, 2010; Rekker, 2025). Right-wing parties have amplified this skepticism by positioning

themselves as anti-environmentalist and stewards of the free economy (Gemenis et al., 2012; Oreskes & Conway, 2022). At the same time, conservatives are more inclined to favor production science, which aligns with their emphasis on economic growth, individualism, and system stability (Jost et al., 2003, 2017). Liberals, in turn, favor impact-oriented science because it aligns with their views on environmental protection, social justice, and progressive change (Sterling et al., 2019). Research supports this distinction and the resulting sorting: conservatives are more likely to reject environmental science and sociology (Kahan, 2012; Kennedy et al., 2022; Rekker, 2025; Tromp & Achterberg, 2021), while liberals are more likely to reject nuclear power, hydraulic fracturing, and concerns about genetically modified organisms (Nisbet et al., 2015; Rekker, 2025; Washburn & Skitka, 2018).

In sum, political identification shapes science *preferences*, rather than leading to a rejection of science in general. Conservatives are often portrayed as anti-science because it is increasingly associated with regulation and stereotypically aligned with liberal values. Yet, their skepticism is directed at impact science, while remaining supportive of production science. Liberals, by contrast, prefer impact science over production science. Thus, science polarization can be understood as the alignment between individuals' views on fields and claims and their political beliefs, reflecting a broader trend in which views across multiple issues become correlated and constrained by ideology (Baldassarri & Gelman, 2008; Gauchat, 2023). We now turn to motivated reasoning and how it guides the acceptance or rejection of information, and its contribution to political polarization over science.

2.3.2 Motivated reasoning

Motivated reasoning suggests that people process information in ways that serve specific goals (Kunda, 1990; Leeper & Slothuus, 2014). Individuals may pursue accuracy goals, processing information to reach conclusions that best represent reality, or they may pursue directional goals, processing information in ways that confirm their prior beliefs and serve personal interests.

In the context of politicized information, directionally motivated reasoning leads individuals to favor conclusions aligned with their political identification and to reject those

that do not. In what follows, "motivated reasoning" refers specifically to the pursuit of directional goals. While accuracy goals may be sought when the stakes of reaching a wrong conclusion are high (Kunda, 1990), interpreting information in group-consistent ways yields "considerable emotional, social, and even material benefits" (Williams, 2023, p. 5). Consequently, partisan commitment motivates the pursuit of directional goals.

Humans are group-oriented, showing preference for those who share their beliefs (Van Bavel & Pereira, 2018). Adherence to group beliefs signals loyalty and commitment (Williams, 2022), facilitating cooperation (Blancke, 2023), while dissent can provoke negative social and emotional consequences (Williams, 2023). Political identification functions as an important social category, making it salient in the detection of possible coalitions (Cikara et al., 2017). Consequently, the stronger one's attachment to a political group, the greater the motivation to defend its beliefs.

Having the motivation and a desired conclusion does not guarantee that an individual can reach it. Rather, one must be able to generate supporting evidence in a way that maintains the illusion of objectivity (Kunda, 1990). The key role of ability has received some empirical support (e.g., Joslyn & Sylvester, 2019; Kahan, 2013; Schulz-Hardt et al., 2000; Wagner et al., 2014), showing that more capable individuals are more likely to build self-serving arguments. A similar pattern has been observed in attitudes toward science, where individuals with higher education and greater science literacy tend to be more politically polarized when evaluating controversial scientific information (Drummond & Fischhoff, 2017).

Thus, when evaluating information, directional goals may bias individuals toward desired rather than accurate conclusions, leading them to interpret evidence in line with their political identification. Yet motivation alone does not guarantee a desired conclusion. People must have the resources to reason their way toward it, selectively engaging with evidence to confirm their beliefs. In the context of science polarization, political identification provides the directional motivation, and this motivation must be paired with the ability to construct arguments that dismiss dissonant scientific information.

2.3.3 Commitment and ability

Motivated reasoning requires both partisan commitment and ability, and examining either factor alone provides an incomplete picture of how science polarization emerges. As established earlier, motivation without ability is insufficient, as individuals need to draw from their knowledge to construct supporting arguments in the face of dissonant information.

That said, ability may not be necessary to produce polarized attitudes. People may adopt identity-consistent conclusions simply because they are exposed to sources aligned with their political group (Williams, 2023). Similarly, individuals may strive to form accurate beliefs, yet still appear to diverge because they differ in what they consider to be credible evidence (Druckman & McGrath, 2019). In this view, polarized beliefs may emerge not from biased *reasoning* but due to differences in information sources and credibility evaluations.

Moreover, some experimental evidence suggests that cognitive ability can reduce bias, rather than reinforce it. For example, analytical thinking has been associated with lower belief in fake political news, suggesting that such beliefs may stem from a lack of critical thinking rather than identity-protective reasoning (Pennycook & Rand, 2019). Similarly, cognitive sophistication has been associated with fewer misconceptions about COVID-19 (Pennycook et al., 2022). These findings suggest that without motivation, ability may promote accuracy rather than bias.

Therefore, those committed to a political group have the *motivation* to defend its positions, while those with the ability can *reason* toward identity-consistent conclusions. In other words, insofar as political polarization over science is driven by motivated reasoning, it is most likely to appear among individuals who are both partisans and knowledgeable.

2.3.4 Hypotheses

We apply this motivated reasoning framework to political polarization over science, focusing on trust in research fields and expert claims. When evaluating research fields, people often consider the ideas and values associated with them (Altenmüller et al., 2024; Gligorić et al., 2025; Rekker, 2021, 2025). These fields vary in their alignment with political identifications, depending on whether they aim to mitigate the harms of industrial produc-

tion and public policy (impact science) or advance economic growth (production science). Similarly, scientific claims are aligned with these orientations because of the practical consequences of accepting them (Kahan et al., 2015). Polarization emerges when individuals perceive fields and claims as threatening to their group's beliefs, and those who are partisan and knowledgeable use reasoning not to seek truth but to justify their beliefs and persuade others (Mercier & Sperber, 2011). From this perspective, political polarization over science is not a matter of ignorance or scientific illiteracy but of the defense of one's group.

In sum, trust in science is likely to follow political lines, depending on whether the field or claim aligns with an impact or production orientation. Conservatives tend to be more skeptical of science because it is often perceived as regulatory and liberal-aligned. Specifically, they are expected to be more skeptical of impact science, which tends to call for the regulation of industry and the pursuit of environmental and social justice, and more supportive of production science, which facilitates industrial and technological development. Liberals, by contrast, are expected to show the opposite pattern. In this study, we use left-right identification as a proxy for liberal and conservative orientation (Fuchs & Klingemann, 1990). Consequently, we hypothesize:

H1a - The more to the right someone leans, the less they trust science in general.

H1b - The more to the right someone leans, the less they trust impact science.

H1c - The more to the right someone leans, the more they trust production science.

The hypothesized differences are expected to emerge when individuals engage in motivated reasoning. Observing left-right differences is not sufficient to detect the mechanism (Druckman & McGrath, 2019). Rather, it requires both the partisan commitment and the ability to defend group-consistent beliefs. We operationalize motivated reasoning as two complementary interactions: between closeness to a particular political party and political identification, and between education and political identification.

First, we use closeness to a party because of its connection to engagement with party politics and political trust (Hooghe & Kern, 2015) and affective polarization (Wagner, 2021), and validity in multi-party systems (Paparo & Sio, 2017). Moreover, the measure has been used to observe partisanship (Dassonneville et al., 2023). Such closeness to a

particular political party should motivate individuals to defend its interests.

Second, we use education as a proxy for the ability component because more educated people should be more familiar with science (Bak, 2001; Kunovich, 2024; Noy & O'Brien, 2019; Shapin, 2008). This familiarity endows them with knowledge about the way science works (Archer et al., 2015), giving them the ability to engage with scientific information. Paradoxically, this knowledge also enables them to criticize science and dismiss dissonant information (Drummond & Fischhoff, 2017; Eyal, 2019; Gauchat, 2012). Consequently, we hypothesize:

H2a - The effect of left-right identification on trust in science (general, impact, and production) is more pronounced for individuals who feel closer to a particular party.

H2b - The effect of left-right identification on trust in science (general, impact, and production) is more pronounced for more highly educated individuals.

2.4 Methodology

We followed a similar approach used by Rekker (2025) and leveraged data from the "In Science We Trust" section of the fifth wave of the CRONOS-2 online panel collected between June 2022 and February 2023. We decided to work with these data because they go beyond typical measures of trust in science by including different research fields and claims. Additionally, they offer a perspective on the European context, which has received comparatively less attention than the US in studies of political polarization (Boulianne, 2019; Wagner et al., 2014). These data came from participants who were part of the European Social Survey (ESS) Round 10[2], representing 11 countries: Austria, Belgium, Czechia, Finland, France, the United Kingdom, Iceland, Italy, Portugal, Sweden, and Slovenia. After removing cases with missing information on exogenous variables, we were left with an analytic sample of 5,410 participants.

[2]https://doi.org/10.21338/cron2w5e01

2.4.1 Dependent variables

Our dependent variables were built with data from the fifth wave of the CRONOS-2 questionnaire. In it, participants answered questions about their trust in science as a whole, research fields, and scientific claims. First, they rated their trust in universities, scientists, and the scientific method. Then, they were asked, "How much trust do you have in this scientific discipline: [...]" and rated their trust in sociology, economics, environmental science, and agricultural science. Finally, participants rated their trust in the following statements: "Differences in behavior between men and women are not fixed at birth, but are mainly caused by upbringing of parents and society," "Increasing taxes on the rich will reduce government revenue because rich people will work less," "The Earth's climate is changing as a result of greenhouse gas emissions caused by human activity," and "Genetic modification of plants improves the productivity of farming without posing health risks for consumers." All statements were said to have been made by scientists, though no information was provided about the level of scientific consensus among experts. The questions regarding science as a whole, research fields, and scientific claims were all measured on a scale from 0 ("No trust at all") to 10 ("Complete trust").

We used the responses to the questions above to create three factors. The first factor grouped responses related to *science as a whole*, which included trust in universities, scientists, and the scientific method. Unlike Rekker (2025), who treats disciplines and claims separately, we specify two additional science specific factors. The second factor captured trust in *impact science* and included responses to trust in sociology, environmental science, the gender differences claim, and the climate change claim. The third factor focused on *production science* and combined responses to trust in economics, agricultural science, the taxation claim, and the genetic modification claim. This structure allowed us to test whether Europeans distinguish between these science orientations and whether motivated reasoning shapes their evaluations.

2.4.2 Independent variables

Our independent variables came from both the CRONOS-2 questionnaire and the ESS Round 10 dataset. We used participant's left-right identification from the ESS10 to mea-

sure political identification. Participants were asked to place themselves on a left-right identification scale from 0 ("Left") to 10 ("Right"). Next, to account for their partisan commitment, we used a measure of closeness to a political party from the ESS dataset. Participants were asked whether they felt closer to a particular party than all others ("yes" or "no"). Additionally, we measured participants' ability to engage with scientific information by including their level of education from the CRONOS-2 dataset. This was measured using the harmonized ES-ISCED international classification: "less than lower secondary" (0), "lower secondary" (1), "lower tier upper secondary" (2), "upper tier upper secondary" (3), "advanced vocational, sub-degree" (4), "lower tertiary education, BA level" (5), "higher tertiary education, >= MA level" (6). Finally, to measure motivated reasoning, we build interaction variables between left-right identification and (*i*) closeness to a party, and (*ii*) education, respectively. We assess the effect of left-right identification for three levels of education (0, 3, and 6) and two reference categories of closeness to a political party ("yes" and "no").

2.4.3 Analytical strategy

We tested our hypotheses by fitting five structural equation models (SEM) of increasing complexity to our data. This approach allowed us to perform a factor analysis on the trust items and regress the resulting factor scores on our independent variables simultaneously. Although our sample includes respondents from different countries, we did not specify a multi-level structure due to the limited number of cases per country. The first model (M1) estimated the direct effects of left-right identification, education, and closeness to a party on the trust factors, as well as the effect of trust in science as a whole on trust in impact and production science. Then, we progressively freed interaction terms and controls: first, the interaction between left-right identification and closeness to a party (M2); then, the interaction between left-right identification and education (M3); and finally, the controls for gender (with males as the reference category), age, and country dummies (with Austria as the reference category; M4). After this step, we inspected the modification indices[3] and accounted for covariances that made theoretical sense in a final model (M5). The full

[3]The modification indices indicated that we account for the residual correlations between (*i*) trust in the tax claim and trust in the GMO claim (both potentially related to social justice concerns), (*ii*) trust in environmental science and trust in agricultural science (both fields mention science in their names), and (*iii*) trust in economics and trust in sociology (both are social sciences).

hypothesized model is shown in Appendix 2.1.

We evaluated the fit of each model using scaled RMSEA, CFI, TLI, and unbiased SRMR—standard indices recommended for SEM (Hu & Bentler, 1999), and scaled likelihood-ratio tests (LRT) to compare their relative fit (scaled Chi-squared values). Each model included sampling weights ("c2weight" variable), was estimated with full information maximum likelihood (FIML) to account for missing values in the dependent variables, and its factors' distributions were standardized (their means were set to 0 and standard deviations set to 1).

2.5 Results

The results are presented in four steps. First, we assess the suitability of the data and evaluate the performance of the structural equation models. Second, we interpret the parameters of the simplest model (M1), which includes only the direct effects on the trust factors (H1a, H1b, H1c). Third, we interpret the parameters of the most complex model (M5), which includes interactions between left-right identification and both closeness to a particular party and education (H2a and H2b). For M1 and M5, we report unstandardized estimates for ease of interpretation and 95% confidence intervals (CIs).

2.5.1 Data and model evaluation

The Kaiser-Meyer-Olkin (KMO) value of 0.92 and the near-zero p-value from Bartlett's test indicate that the data are appropriate for factor analysis. The internal consistency of the trust in science as a whole and trust in impact science factors was good, with Cronbach's α values of 0.86 and 0.70, respectively. However, the internal consistency of the trust in impact science factor was low, with a Cronbach's α of 0.51. As shown in Appendix 2.2, each step in model specification significantly improved model fit. The final model (M5)—which includes all interaction terms, control variables, and residual correlations—shows a good fit based on scaled RMSEA and SRMR values (Hu & Bentler, 1999). The scaled CFI is within an acceptable range, though the TLI is low. We follow Ximénez et al. (2022) and report an unbiased SRMR of 0.028, which falls below the accepted threshold (Shi et al., 2018). The factor loadings for the final model (M5) can be

seen in Appendix 2.3.

2.5.2 Political polarization

The results of the direct effects model (M1) are displayed in Figure 2.1. Individuals who lean further to the right tend to trust impact science less ($b = -0.132$, CI = [-0.155, -0.109], $p < 0.001$), trust production science more ($b = 0.107$, CI = [0.080, 0.134], $p < 0.001$), and trust science as a whole less ($b = -0.033$, CI = [-0.050, -0.017], $p < 0.001$). These findings are in line with Rekker's (2025) and support H1a, H1b, and H1c. We observe that the effects of political identification on trust in science as a whole are much smaller than its effects on trust in impact and production science. This shows that impact and production science claims and research fields are more politicized, indicating a stronger political bias in the European public toward these types of science.

Education has a consistently positive effect on all science factors. The more educated people are, the more they trust impact science ($b = 0.044$, CI = [0.013, 0.075], $p = 0.005$), production science ($b = 0.039$, CI = [0.001, 0.077], $p = 0.042$), and science as a whole ($b = 0.128$, CI = [0.105, 0.151], $p < 0.001$). People who feel closer to a particular political party show more trust in science as a whole ($b = 0.203$, CI = [0.130, 0.276], $p < 0.001$), but trust impact science less ($b = -0.143$, CI = [-0.155, -0.109], $p = 0.003$) compared to those who do not feel close to a party. There is no difference between these groups in their trust in production science ($p = 0.909$).

Finally, higher trust in science as a whole is associated with more trust in impact ($b = 1.374$, CI = [1.241, 1.506], $p < 0.001$) and production science ($b = 1.466$, CI = [1.295, 1.638], $p < 0.001$). All factor loadings are significant and positive, though the item measuring trust in the tax claim ($b = 0.170$, CI = [0.105, 0.235], $p < 0.001$) stands out as the lowest. Additionally, the covariance between the trust in impact and production science factors is significant and positive (0.766), indicating that more trust in one orientation corresponds with more trust in the other.

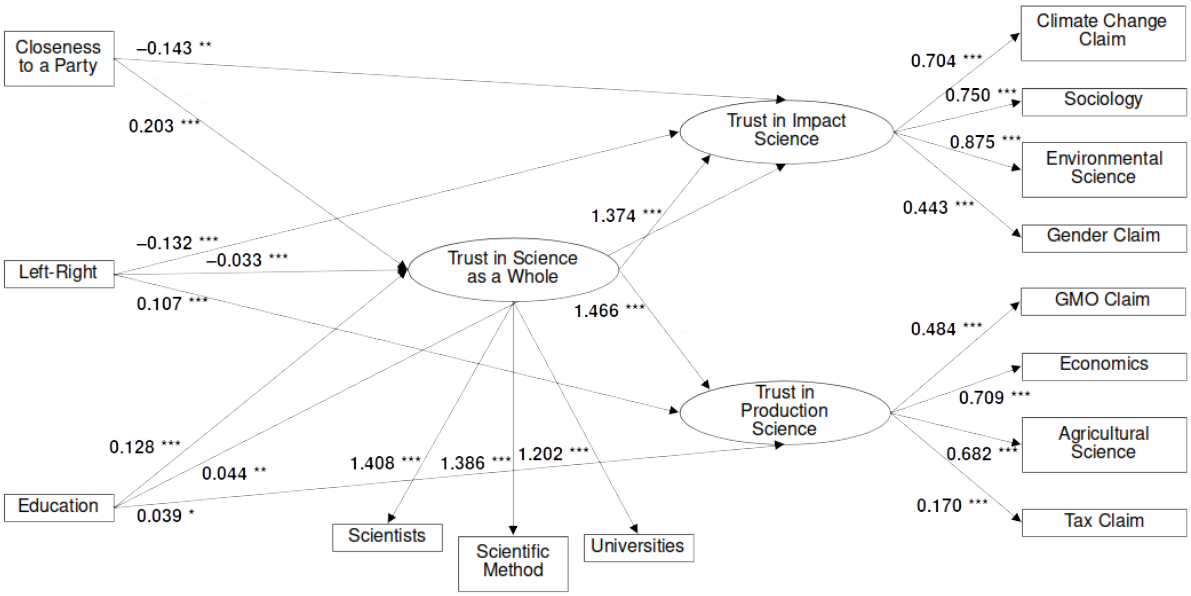


Figure 2.1: **Direct effects model (M1).** Path diagram showing unstandardized effects. The path between closeness to a party and trust in production science is not significant ($p = 0.909$). Covariance between trust in impact and production science (0.766), variances, and errors are not shown for simplicity. $N = 5{,}410$, $\chi^2(219) = 2,431.225$, $p < 0.001$; RMSEA = 0.043, uSRMR = 0.042, CFI = 0.846, TLI = 0.830; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.5.3 Motivated reasoning

Observing left-right differences in trust does not mean that these stem from motivated reasoning. The mechanism entails more than holding group-consistent views: it requires commitment to that group and an active search for justifications that favor it. To assess the role of motivated reasoning, we interpret the effects of our hypothesized interactions on trust in science. The significant paths of the full model (M5) are depicted in Figure 2.2, and Figure 2.3 shows the effect of left-right identification on our trust factors at different levels of party closeness and education.

Europeans who lean further to the right and feel closer to a particular party trust *science as a whole* slightly less ($b = -0.039$, CI = $[-0.077, -0.002]$, $p = 0.041$; H2a) than those who do not feel such closeness. However, the effect of left-right identification does not vary by educational level ($p = 0.685$; H2b). These findings suggest that political identification affects trust in science as a whole only when individuals feel committed to their political group, with those leaning further to the right reporting less trust. Without

such commitment, political identification does not influence their trust.

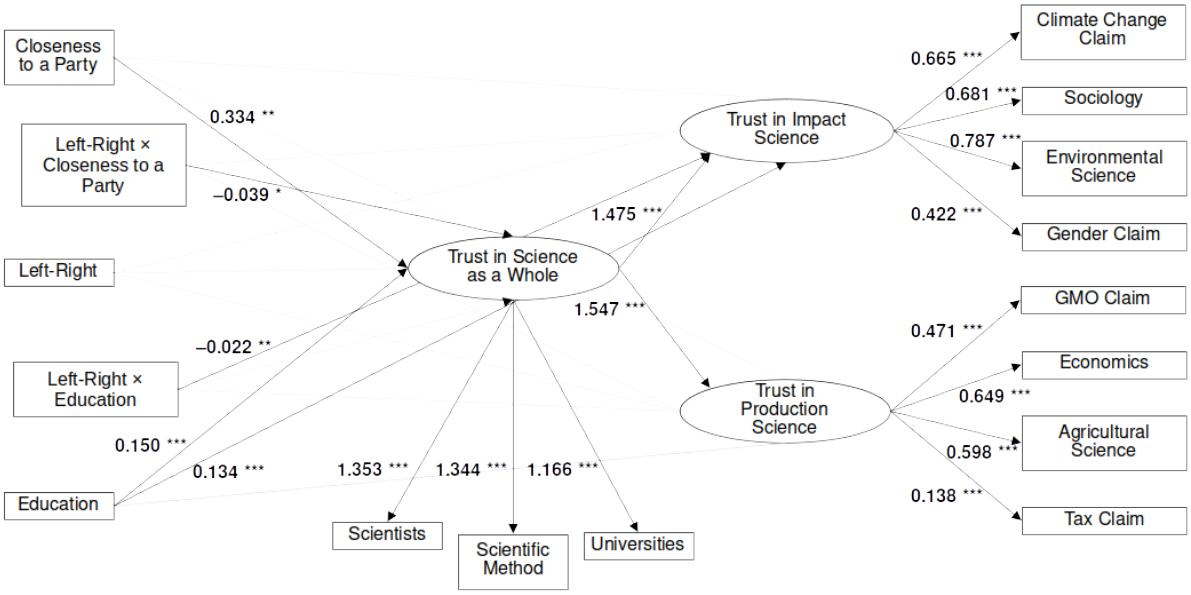


Figure 2.2: **Full model (M5).** Path diagram showing unstandardized effects. The non-significant paths, covariance between trust in impact and production science (0.569), variances, and errors are not shown for simplicity. N = 5,410, $\chi^2(174) = 1,662.205$, p < 0.001; RMSEA = 0.040, uSRMR = 0.028, CFI = 0.897, TLI = 0.856; * p < 0.05, ** p < 0.01, *** p < 0.001.

According to the model parameters in Figure 2.2, the effect of left-right identification on *trust in impact science* does not vary significantly between levels of closeness to a political party ($p = 0.093$; H2a). However, Figure 2.3 reveals that, among Europeans with "less than lower secondary", left-right identification does not significantly affect trust in impact science among those without partisan commitment ($p = 0.460$), but becomes significant for those who feel committed to their party ($b = -0.070$, CI = $[-0.131, -0.009]$, $p = 0.024$). This pattern is indicative rather than conclusive, as it emerges from a comparing two effects within the same group. More broadly, education amplifies the effect of left-right identification across all individuals: the higher their education, the stronger the association between political identification and trust in impact science ($b = -0.022$, CI = $[-0.036, -0.008]$, $p = 0.002$; H2b). In other words, more educated Europeans exhibit greater political polarization in their trust in impact science: those who lean further to the right trust it less, and this divide is stronger the more educated they are. Additionally, we find tentative evidence that this educational effect is stronger among individuals who feel committed to a party.

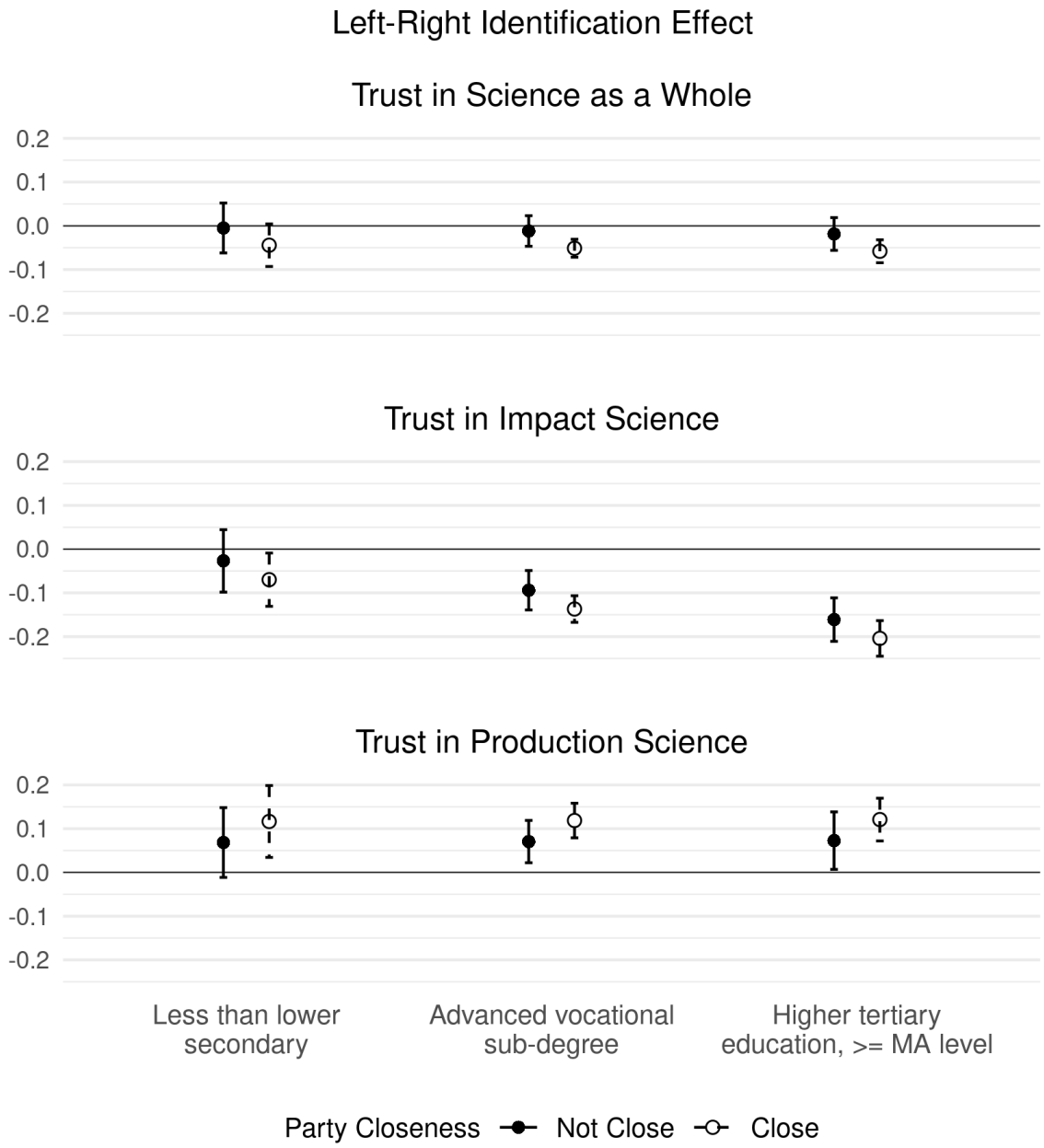


Figure 2.3: **The political polarization of science.** Estimated (M5) effect of left-right identification at different levels of education and closeness to a particular political party on the trust factors. Error bars show 95% confidence intervals. All other variables in the model were set to 0.

Finally, according to the model parameters, the effect of left-right identification on trust in *production science* does not vary by either closeness to a party ($p = 0.110$; H2a) or education ($p = 0.938$; H2b). However, as with trust in impact science, Figure 2.3 reveals a more complex pattern. Among non-committed Europeans, the effect of left-right identification is not significant at the lowest ($b = 0.068$, CI $= [-0.011, 0.148]$, p

= 0.093) and close to zero at the highest (b = 0.073, CI = [0.007, 0.138], p = 0.030) educational levels. The effect becomes significant, however, among those who feel close to a party at both levels (b = 0.116, CI = [0.034, 0.199], p = 0.005; b = 0.121, CI = [0.072, 0.170], $p < 0.001$, respectively). Moreover, education does not significantly interact with left-right identification (p = 0.938). Thus, Europeans who lean further to the right trust production science more, with tentative evidence that this effect is stronger among those committed to a party, but more education is not associated with a stronger political divide.

In sum, we find some tentative evidence supporting H2a and H2b, suggesting that motivated reasoning does not drive *all* political polarization over science. While all trust in science factors are affected by political identification, only trust in *impact science* shows evidence of an amplified political divide in the presence of party commitment and higher education. Both trust in *science as a whole* and, to some extent, *production science* are more strongly influenced by political identification among committed Europeans than among those who are not committed. In the context of science polarization, if motivated reasoning involves both commitment and the ability to reason away dissonant scientific information, then we find evidence for its presence *only* regarding impact science—the type that typically challenges industrial interests and supports regulatory interventions.

2.6 Discussion and conclusion

This study investigates the patterns of political polarization over science, with a focus on the role of motivated reasoning. We distinguish between two orientations of science: impact science, which focuses on assessing the social and environmental consequences of industrial activity and policy, and production science, which tends to support technological and industrial development.

While much of the existing literature on science-related polarization is grounded in the US (Cologna et al., 2025), our analyses of the ESS10 data reveal similar patterns of ideological sorting. Echoing Rekker (2025), we find that left-leaning Europeans exhibit higher levels of trust in science overall, and particularly in impact science, compared to their right-leaning counterparts. By contrast, right-leaning Europeans exhibit higher

trust in production science. These results align with theoretical expectations regarding the influence of ideological congruence on trust in science (Drummond & Fischhoff, 2017; McCright & Dunlap, 2010).

However, in contrast to prior work, our results suggest that Europeans do not consistently group research fields (economics and agricultural science) and claims (taxation and GMOs) into a coherent production science orientation. This is evidenced by two observations: (*i*) low internal consistency of the trust in production science items, and (*ii*) a null effect of both the interactions (closeness to a party and education) on trust in production science. That said, the absence of a clear cluster does not imply the absence of ideological preferences altogether. As Rekker (2025) shows, right-leaning individuals still tend to express greater support for economics, agricultural science, and claims favoring GMOs and opposing taxation. These preferences suggest that even if Europeans do not organize their views consistently across research fields and claims, political identification still plays a role in shaping trust in them.

Motivated reasoning suggests that political polarization over science arises because committed and capable individuals dismiss research fields and claims that contradict their beliefs. In the context of science polarization, this mechanism involves both the motivation and ability to dismiss dissonant scientific information. In this study, we operationalize these as the interactions between left-right identification and (*i*) closeness to a political party, and (*ii*) education, because the former provides the individual with the motivation to defend their party's interests, and the latter with the familiarity with science needed to build arguments against it.

Our findings suggest that motivated reasoning does not drive all science polarization. Trust in science as a whole shows slightly greater polarization among Europeans who feel closer to their party, but education does not amplify this divide. Trust in production science shows tentatively greater polarization among the party-committed, but again without educational amplification. The absence of an educational amplification may reflect alternative mechanisms, such as individuals relying on different sources they perceive as credible rather than actively pursuing directional goals (Druckman & McGrath, 2019). Trust in impact science exhibits the clearest evidence of motivated reasoning: polarization increases and among the more educated, and, tentatively, among the more committed,

supporting the funnel pattern identified by Drummond and Fischhoff (2017) whereby political divides over scientific issues are most pronounced among the highly educated.

In synthesis, we find some evidence for motivated reasoning—requiring both partisan motivation and the ability to selectively engage with scientific information—only in the domain of impact science. The polarization observed in trust in science as a whole and production science seems to be driven by partisan commitment alone, without the active deployment of familiarity with science to dismiss dissonant findings. Moreover, this commitment-driven polarization is modest for science as a whole and remains tentative for production science.

Our study has five main limitations. First, the models specified three latent variables—trust in science as a whole, trust in production science, and trust in impact science—measured with three, four, and four observed items, respectively. These are relatively few indicators per factor. While the internal consistency was good for trust in science as a whole and impact science, it was problematic for production science. Moreover, the lowest factor loading in our model occurred in this factor. These issues likely contributed to the lower TLI fit index in our models. Future research distinguishing these domains may benefit from including more observed items per factor. Second, our trust measures focused on participants' trust in scientific institutions, fields, and claims. However, trust is a multidimensional concept that encompasses perceived accuracy, impartiality, and expertise (Bauer & Freitag, 2017). The measures we used may therefore capture only a subset of public trust, and future surveys could improve on this by including broader indicators. Third, we measured commitment with a single item of closeness to a particular political party, but future research could use multi-item scales to better reflect partisan commitment (e.g., Bankert et al., 2017). Fourth, we used education as a proxy for familiarity with science and scientific knowledge because of data availability. Future research could include more direct measures, such as a science literacy test (e.g., Mede et al., 2025). Finally, our conclusions are based on correlational tests of observational data.

In sum, our study suggests that the political polarization over science is not uniform. In the European context, we find some evidence that motivated reasoning shapes public trust specifically in impact science. By contrast, we do not find support for motivated

reasoning in polarization over trust in science as a whole or in production science. These conclusions should be taken with caution, as we measured motivated reasoning by using proxies (closeness to a party and education) available in our data. As such, we cannot draw definitive conclusions about its presence across science orientations or its causal role in producing science polarization. Nonetheless, this study contributes to the literature by operationalizing motivated reasoning in a way that jointly takes into account partisan commitment and ability. Future experimental research could directly measure and manipulate both to make more definitive claims.

In an era where scientific guidance is increasingly central to policy (Shapin, 2008), understanding the drivers of science polarization is critical. Rather than viewing polarization as a failure of knowledge, it may be more useful to see it as an expression of political identification. Our findings suggest that efforts to build public trust in science should recognize the politically charged nature of impact science. Additionally, public communication strategies in Europe should be sensitive to the fact that conservative audiences may view regulatory-oriented science with skepticism, not due to ignorance, but because of ideological commitments reinforced by both party allegiance and engagement with scientific information that enables selective skepticism.

Despite claims about an ongoing crisis in trust in science (Gauchat, 2010; Nichols, 2017), recent scholarship on the issue concludes that such claims may be overstated (Cologna et al., 2025). In this paper, we show that there are reasons to remain optimistic about the degree to which people oppose and distrust science. Europeans exhibit some politically determined preferences, with the political right favoring certain fields and claims that align with their economic outlooks, while the left favors impact science. However, these patterns should not be interpreted as evidence of outright science rejection. First, the existence of differential preferences shows that individuals of both sides of the ideological aisle can appreciate science, providing a foundation for bridging trust gaps. Second, the political polarization over science as a whole only emerges among committed partisans, while more pronounced divides specifically appear over impact science. Taken all together, general science skepticism remains limited.

Chapter 3

Beliefs and Expectations in the Sharing of Expert Information

This chapter was written by the thesis author. Currently, it is a working paper developed with support in framing and design from Achim Edelmann and Noam Titelman.

Abstract

People often share information without paying attention to its accuracy, and instead prioritize how it interacts with their personal assumptions. This article examines how interactions between expert messages and individuals' beliefs and expectations affect perceptions of surprise and interestingness, and sharing intentions. Drawing from relevance theory and information theory, beliefs and expectations are distinguished as two different types of assumptions, and their roles on information perceptions and sharing are tested separately. US adults (N = 1,772) participated in a pre-registered online experiment, evaluating expert messages about climate change, artificial intelligence, nutrition, migration, and population growth. Results show that both belief contradictions and expectation violations are equally surprising, and both make messages more interesting, with belief contradictions having a stronger effect. However, sharing intentions are mainly driven by the validation of expectations rather than by their violation, while belief reinforcements do not affect sharing. These findings suggest that people react differently to discrepancies and correspondences with beliefs and expectations when deciding to share expect information, and that surprising messages are not necessarily shared more often.

3.1 Summary

When sharing information online, people may prioritize its relevance (Altay et al., 2022) or the informational value it contains (Goebel et al., 2024) rather than its accuracy. This chapter shifts the focus away from epistemic truth toward the interaction between information and the observer's assumptions. Drawing from relevance theory (Wilson & Sperber, 1995, 1998) and information theory (Shannon, 1948), I distinguish between *beliefs*—propositions taken to be true—and *expectations*—anticipations about what is likely to happen. This approach allows me to test how interactions between personal assumptions and information differ depending on whether they involve belief contradictions and reinforcements, and expectation violations and validations. I apply this framework to study the intention to share expert information online.

The results indicate that discrepancies with both assumptions increase how surprising expert information is and belief contradictions increase perceived interestingness more strongly than expectation violations. Reinforcing beliefs does not increase sharing intentions, but validating expectations does. There is also tentative evidence that, when expectations are validated, reinforcing expert information increases sharing intentions.

These findings provide some answers to the question concerning individual assumptions. They suggest that US participants are unlikely to share surprising or interesting information. Instead, they are more inclined to share information that validates an *expectation*. Overall, the results point to differences in how information interacts with beliefs and expectations: belief contradictions shape the interestingness of information more strongly than expectation violations, whereas expectation validations weigh more than belief confirmations in shaping sharing intentions. This implies that the diffusion of expert information on social media may depend less on what users think is true than on whether it aligns with what they *anticipate* will happen in the near future, highlighting the usefulness of information-theoretic accounts in explaining information sharing.

3.2 Introduction

Misinformation has become a pressing challenge, largely amplified by the proliferation of social media platforms (Guess et al., 2018; Rao et al., 2022). The low entry barrier (Zhuravskaya et al., 2020) and reward systems (Ceylan et al., 2023) of social media enable a mediascape with a myriad of outlets that incentivize sharing. This makes these platforms democratic sources of information (Tucker et al., 2017), but also attractive channels for the diffusion of political (Guess et al., 2019) and expert misinformation (Del Vicario et al., 2016; Suarez-Lledo & Alvarez-Galvez, 2021; Wang et al., 2019).

Recent literature suggests that people often do not pay attention to the accuracy of a piece of information (Pennycook et al., 2021), and instead prioritize other factors when evaluating and deciding what to share. For example, users may share information based on its emotional appeal (Tellis et al., 2019), interestingness (Altay et al., 2022), and its alignment with their political attitudes (Buchanan, 2020).

In this study, I focus on expert information, which plays a central role in contemporary societies. It allows citizens and policymakers to deal with the complexity of technology and societal risks (Gauchat, 2023) by relying on scientific knowledge that informs public policy and decision-making (Shapin, 2008). Understanding how expert information is shared and perceived online is therefore important, as many people rely on online platforms as their primary source of information (Cuenca & Pérez, 2025; Wilson et al., 2020).

Drawing from relevance theory (Wilson & Sperber, 1995, 1998) and information theory (Shannon, 1948), this study aims to contribute to understanding the determinants of expert information perceptions and sharing, shifting the focus away from accuracy and epistemic truth toward the interaction between information and individual assumptions. I contribute to the literature in two ways: first, by conceptually distinguishing between *beliefs* and *expectations* as two different types of assumptions; second, by empirically testing how belief contradictions and reinforcements, and expectation violations and validations shape participants' experiences of surprise and interestingness, and their willingness to share expert information.

I test several hypotheses in an online pre-registered[1] experiment, where US participants

[1] https://osf.io/j85gr

were asked to evaluate messages with expert information about artificial intelligence (AI), climate change, migration, nutrition, and population growth. The results suggest that sharing intentions are mainly driven by the validation of expectations rather than the reinforcement of beliefs. This indicates that, when deciding whether to share expert information, individuals react differently depending on how information engages with their different assumptions. Additionally, discrepancies with both beliefs and expectations are perceived as equally surprising, and both increase interestingness, with the effect of belief reinforcements being stronger.

I proceed in four steps. First, I develop a theoretical framework drawing from relevance theory and information theory. Second, I outline the experimental design and procedure, and provide a short description of the sample, measures, and analytical strategy. Third, I describe the results, beginning with the relationships between dependent variables, then turning to perceptions of surprise and interestingness, and finally to information-sharing intentions. I then conclude by discussing the main findings, limitations, and implications for the literature on relevance and surprise.

3.3 Theoretical framework

In the following section, I briefly outline relevance theory and information theory, and their applications to information sharing. I then discuss how beliefs and expectations play distinct roles in the sharing of information.

3.3.1 Relevance theory

Relevance theory, attributed to Wilson and Sperber (1995), posits that human cognition is geared towards improving our knowledge of the world. Individuals aim to acquire knowledge by processing new information that interacts with existing assumptions, understood as "thoughts treated by the individual as representations of the actual world" (p. 2). Because the amount of available information exceeds our cognitive capacity, we selectively process information based on its *relevance*, which is evaluated in terms of cognitive effects and processing effort (Wilson & Sperber, 1998).

On the one hand, cognitive effects refer to the changes in assumptions produced by their interaction with new information. These effects may take the form of reinforcements, contradictions, or newly derived conclusions. On the other hand, processing effort refers to the mental resources required to achieve these effects and depends on how information is presented, as well as on the cognitive resources dedicated to memory recall and imagination used to contextualize it. The greater the cognitive effects of a piece of information and the lower the processing effort required, the more relevant the information is.

Whether new information produces cognitive effects depends on how it relates to an individual's assumptions: it may strengthen them or be strong enough to provoke their revision. Relevance can therefore arise through confirmation or challenge to assumptions, either by contradiction or giving rise to new conclusions. In both cases, what matters is the consequences of processing the information. Even objectively false information may be relevant to an observer if it leads to cognitive gains, as the updating of assumptions "seems more crucial to relevance than [the] truth of the premises" (Wilson & Sperber, 1995, p. 264). In this sense, fictional and objectively false information can produce cognitive effects and be relevant.

Applied to information sharing, Altay et al. (2022) equate the relevance of a piece of information with its *interestingness*[2], in the sense that relevant information is experienced as interesting, and argue that such information is more likely to be shared. Like Wilson and Sperber (1995), they place emphasis on the cognitive effects information may have for an observer, while relegating its epistemic accuracy to a second plane. In this view, relevance, experienced as interestingness, helps explain why individuals share (mis-)information *despite* its inaccuracy (Pennycook et al., 2021).

In synthesis, holding processing effort constant, individuals pay attention to and process information that meaningfully interacts with their assumptions. Information is thus relevant when it generates cognitive gains and, from the observer's perspective, is perceived as interesting regardless of its epistemic accuracy. Because such information carries potential cognitive value, it is also more likely to be shared. Information theory draws a

[2]They distinguish two concepts. First, *interestingness* is equated with relevance and depends on both the *accuracy* and potential cognitive effects of a piece of information. Second, *interestingness-if-true* is independent of accuracy and refers solely to the cognitive effects of a piece of information. For the purpose of their argument, the authors disregard differences in processing effort. When I mention "interestingness" I will be referring to "interestingness-if-true"

comparable conclusion.

3.3.2 Information theory

Information theory, attributed to Shannon (1948), can be used to quantify informational value. According to it, the informational value of an event[3] is captured by its *entropy*, which is a function of its probability distribution. Entropy measures the overall uncertainty associated with an event (Jones, 1979). Accordingly, events with more evenly distributed outcome probabilities have higher entropy than events in which some outcomes are more likely than others (Cover & Thomas, 1991). For example, a fair coin toss has higher entropy than a weighted coin toss in which landing on the heads side is more likely.

While entropy describes the informational value of an event by considering all possible outcomes, information theory allows us to quantify the information given by a particular outcome. This measure is called *surprisal*, or simply surprise (Barto et al., 2013; Tribus, 1961). The amount of information conveyed by an outcome is inversely proportional to its probability of occurring, and this unexpectedness generates surprise. Thus, a coin landing on heads will be more surprising and convey more information when the coin is fair than when the coin is weighted to land on heads. As with entropy, quantifying surprise requires knowledge about the event's probability distribution from the observer's perspective, making it a subjective measure (Baldi, 2002).

Applied to sharing, information theory posits that a message is selected from a set of possible messages because of its informational value (Shannon, 1948). Assuming that individuals continuously form predictions about the world, messages that convey unlikely information carry more informational value and may be more likely to be shared by an observer (Goebel et al., 2024). In this framework, while novelty and surprise are related, they are different concepts. Novelty refers to the degree to which information differs from what is stored in memory, whereas surprise arises from a mismatch between observation and *expectation* (Barto et al., 2013). In other words, novelty is triggered by a single

[3]I use some terms informally for the sake of readability. According to Bertsekas and Tsitsiklis (2008), an experiment is the underlying process that produces exactly one outcome from a set of possible outcomes. A sample space is the set of all possible outcomes of an experiment, and an *event* is a subset of possible outcomes. I use "event" to refer to experiments.

process (i.e., experience not matching memory), whereas surprise requires two phases: an expectation must be formulated, and then that expectation must be violated.

Taken together, informational value is a function of the uncertainty of an event and the likelihood of an outcome. Less probable outcomes convey more information and are more surprising. Applied to information sharing, this implies that messages are shared because they are surprising—that is, because they violate expectations. Both relevance and information theory point to the interaction between a message and the observer's assumptions as key to information sharing, albeit with some differences.

3.3.3 Beliefs and expectations

In relevance theory, the context used to interpret a message consists of a subset of assumptions made salient by incoming information (Wilson & Sperber, 1995). These assumptions include both beliefs about the world and expectations about the future. Accordingly, cognitive effects may arise when new information interacts with assumptions by reinforcing, contradicting, or producing new beliefs *and* expectations. In this framework, beliefs and expectations are treated at the same level, as part of a homogeneous set of assumptions.

Such a broad treatment of assumptions becomes problematic when combined with information-theoretic accounts of surprise in the context of information sharing. According to relevance theory, cognitive effects may arise from contradictions to both beliefs *and* expectations, whereas information theory links informational value specifically to expectation violations. Empirical work reflects this ambiguity. For example, Goebel et al. (2024) operationalize surprise as a reaction to messages that contradict beliefs and find that, although such messages are perceived as surprising, they are not shared more often. This challenges the connection between informational value and sharing.

To address this tension, it is necessary to distinguish between beliefs and expectations as two types of assumptions. Beliefs are propositions taken to be true (Egan, 1986), whereas expectations are predictions about what is likely to occur (Barto et al., 2013). Beliefs can be understood as more stable representations that often inform expectations. For example, if an individual believes that a coin is weighted to land on heads more often, they will expect it to land on heads in a future toss.

Building on this distinction, this chapter considers how information interacts with beliefs by contradicting or reinforcing them, and with expectations by violating or validating them. This framework allows for the examination of how these interactions shape responses to information. Particularly, it helps to clarify why information with the potential to produce cognitive effects, such as belief contradictions that are perceived as interesting, is not necessarily shared (Jonas et al., 2001; Meppelink et al., 2019).

I begin by examining how these interactions affect the perception of information, namely surprise and interestingness. While both belief contradictions and expectation violations could be experienced as surprising in practice, information-theoretic accounts link surprise specifically to expectation violations.

H1 - Messages that violate expectations are more surprising than messages that contradict beliefs.

Beyond surprise, these interactions may also shape whether a message is perceived as interesting. Relevance theory accounts posit that information is interesting when it produces cognitive effects by interacting with existing assumptions, but does not distinguish between different types of cognitive effects in terms of their contribution to perceived interestingness. At the same time, information theory suggests that informational value increases with the unexpectedness of a message. Taken together, these frameworks imply that discrepancies with assumptions produce cognitive effects and carry more informational value than correspondences. Both belief contradictions and expectation violations should therefore be perceived as more interesting than belief reinforcements and expectation validations, respectively.

H2a - Messages that contradict beliefs are more interesting than messages that reinforce them.

H2b - Messages that violate expectations are more interesting than messages that validate them.

Building on these perceptions of information, I turn to how interactions with beliefs and expectations shape sharing intentions. While relevance theory emphasizes the po-

tential cognitive effects of a message, it does not specify which is the primary driver of sharing. Confirmation bias accounts do suggest that individuals prefer belief-consistent information.

H3 - Messages that reinforce beliefs are more likely to be shared than messages that contradict them.

From an information-theoretic perspective, informational value is closely related to unexpectedness. Because expectation violations signal higher informational value and may be experienced as more surprising than validations, they increase the likelihood that a message is shared.

H4 - Messages that violate expectations are more likely to be shared than messages that validate them.

Finally, I consider how engagement with both types of assumptions jointly shapes information-sharing intentions. Beliefs are relatively stable assumptions about what is true, while expectations concern assumptions about what is likely to happen. Beliefs often inform expectations, and expectations in turn shape how incoming information is evaluated in terms of informational value. As a result, interactions between information and these assumptions could jointly influence sharing intentions. Particularly, the combination of belief reinforcement and expectation violation may increase the salience of a message by making it surprising. Therefore, messages that combine these are more likely to be shared.

H5 - Messages that reinforce beliefs and violate expectations are more likely to be shared than messages that reinforce beliefs and validate expectations.

3.4 Methodology

The hypotheses were tested using linear regression models with data from a pre-registered randomized online experiment with US participants.

3.4.1 Experimental design and procedure

The survey experiment was implemented using a Qualtrics questionnaire. After giving their consent, participants reported their familiarity with and knowledge of five topics: climate change, artificial intelligence (AI), nutrition, population growth, and migration.

Participants then indicated their *beliefs* by answering the question, "To the best of your knowledge, is the following statement false or true?" Responses were recorded using a slider from "False 100%" (indicating absolute certainty that the statement was false) to "True 100%" (absolute certainty that the statement was true). Next, participants were asked to formulate an *expectation* by answering to the prompt "According to you, [...]," using a slider from "Will not happen 100%" to "Will happen 100%." These measures captured participants' beliefs and expectations for each of the five topics. All items measuring participants' familiarity, knowledge, beliefs, and expectations were presented in a random order and are shown in Table 3.1.

Responses above 50% toward "True" or "Will happen" were coded as positive responses to the belief and expectation statements, while responses above 50% toward "False" or "Will not happen" were coded as negative responses. If the responses were at exactly 50% or the participant did not provide an answer (possible by indicating "don't know"), they were coded as "no belief" and "no expectation", accordingly. These codings were then used to assign participants to the contradiction/reinforcement and violation/validation conditions described below and match them with the appropriate messages.

Participants were then randomized into one of two versions of the questionnaire. In the *single-report* version, participants read 11 messages (see Table 3.2 for an example) presenting information from fictitious reports that conveyed "new" expert information, one of which was an attention check. Of the remaining ten messages, five were designed to either *contradict* or *reinforce* their stated beliefs, while the other five were designed to either *violate* or *validate* their formulated expectations. In the *dual-report* version, participants read six similarly-framed messages, again including an attention check. The remaining five messages were designed to simultaneously *contradict* or *reinforce* beliefs and *violate* or *validate* their expectations by including information from two fictitious reports presented in a random order within the same message. As such, each questionnaire version had four experimental conditions:

Single-Report:

- Belief contradiction
- Belief reinforcement
- Expectation violation
- Expectation validation

Double-Report:

- Belief contradiction + Expectation violation
- Belief contradiction + Expectation validation
- Belief reinforcement + Expectation violation
- Belief reinforcement + Expectation validation

For example, if a participant in the single-report version indicated that "Industrial activity is the main cause of climate change" with a 75% certainty that the statement was true, they had a 50% chance of being shown a message with information that contradicted that belief and a 50% chance of seeing a message that reinforced that belief. If they assigned a 60% probability that "the United States [will] become overpopulated in the near future", they had a 50% chance of reading a message that violated that expectation and a 50% chance of reading a message that validated that expectation. The randomization and mapping worked the same way for the dual-report condition, but instead of seeing separate messages that interacted with beliefs *or* expectations, this was done simultaneously within the same stimulus. For participants who had no belief or expectation regarding a topic, they were shown a random message. For all possible messages, see Appendices 3.1 and 3.2.

Table 3.1: **Statements shown to participants before the experimental task.** Each participant evaluated all statements and read them in random order.

Measure	Climate Change	AI	Nutrition	Population Growth	Migration
Familiarity	"I constantly hear about the causes of climate change."	"I constantly hear about the the uses of artificial intelligence (AI) in cancer detection."	"I constantly hear about the impact of plant-based foods on people's health."	"I constantly hear about the global population growth."	"I constantly hear about the recent global migration flows."
Knowledge	"How knowledgeable are you about the causes of climate change?"	"How knowledgeable are you about the the uses of artificial intelligence (AI) in cancer detection?"	"How knowledgeable are you about the impact of plant-based foods on people's health?"	"How knowledgeable are you about the global population growth?"	"How knowledgeable are you about the recent global migration flows?"
Belief	"Industrial activity is the main cause of climate change."	"Artificial intelligence (AI) is effective at detecting cancer."	"Eating plant-based meat is healthier than eating animal meat."	"The world's population is constantly growing."	"Immigration to developed countries is at a historic high."
Expectation	"According to you, will winters in the United States become warmer in the near future?"	"According to you, will artificial intelligence (AI) completely replace oncologists in the near future?"	"According to you, will plant-based meat provide nutrition comparable to that of animal meat in the near future?"	"According to you, will the United States become overpopulated in the near future?"	"According to you, will immigration increase pressure on public services in developed countries in the near future?"

Table 3.2: **All possible messages concerning nutrition shown to participants (single-report version).** Messages contradicted or reinforced participants' beliefs and violated or validated their expectations depending on their initial responses. Each message could therefore appear in different experimental conditions. All messages began and ended with a neutral statement.

Belief 1	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat do not outweigh the lack of heme-iron**. Our understanding of the role of plant-based alternatives on health is advancing.
Belief 2	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat outweigh the lack of heme-iron**. Our understanding of the role of plant-based alternatives on health is advancing.
Expectation 1	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that **key nutrients in animal meat will be impossible to produce in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.
Expectation 2	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that **soon, key nutrients in animal meat will be produced in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.

For each report, participants were first asked "How surprising is the above information to you?" and "People often share information with others online. How likely are you to share the above information online (e.g., on Facebook, X/Twitter, Instagram)?" On a second screen, they were asked "How interesting is the above information to you?", among

other measures of unexpectedness, novelty, belief alignment, and relevance. The order of these questions was randomized within each screen.

After indicating their perceptions of the messages, participants were asked to evaluate how true or false they thought the expert claims made in the reports were. These evaluations were measured in the same way as their beliefs, using a slider indicating percentages of certainty. Finally, participants were debriefed on the study's purpose, provided with current expert-supported information on each topic, and thanked for their participation. The full questionnaire, along with the informed consent, can be seen in the OSF materials.

3.4.2 Sample, measures, and analytical strategy

Participants were recruited through Prolific between the 16th and 30th of September 2025. Prolific's prescreening tools were used to ensure that all participants were at least 18 years old and resided in the US at the time of the study. Recruitment aimed for a balance sample in terms of gender, education, and partisan affiliation.

An *a priori* power analysis specifying a conservative effect size, reflecting a 33.3% reduction from a small effect size to account for variance inflation ($d = 0.133$), with a two-tailed test, an $alpha = 0.05$, and a power of 90%, indicated that a minimum of 1,189 observations per cell were required. In the single-report condition, each participant contributed, on average, 2.5 observations per cell (10 pieces of information across 4 possible conditions), implying a required sample of approximately 476 participants. In the dual-report condition, each participant contributed, on average, 1.25 observations per cell (5 pieces of information across 4 possible conditions), implying a required sample of approximately 952 participants.

The initial sample exceeded the targets described above because filters were applied, including: failed attention checks (1.863% and 3.487% in the single- and double-report versions), excessively fast or slow submissions[4] (0.740% and 1.925%), giving the same responses across belief (0.041% and 0.200%), expectation (0.082% and 0.088%), and truth evaluations (0.356% and 0.525%), and missing responses (exactly 50% or "don't know") to the belief and expectation measures (31.795% and 39.050% after the previous filters were applied). A total of 739 participants completed the single-report questionnaire, while

[4]Defined as submissions under three minutes and over 33 minutes and 20 seconds.

1,615 participants completed the double-report questionnaire.

After applying these filters, the analytical sample comprised 564 participants with 4,979 observations for the single-report questionnaire. Of these, 49.823% had a college education, and 50.177% did not; 51.773% were Democrats, and 48.227% were Republicans; 51.950% identified themselves as female, and 48.050% as male.

The double-report questionnaire yielded an analytical sample of 1,208 participants with 4,876 observations. Of these, 54.801% had a college education, and 45.199% did not; 50.414% were Democrats, and 49.586% were Republicans; 50.414% identified themselves as female and 49.586% as male. The combined analytical sample included 1,772 participants and 9,855 observations, distributed across conditions as shown in Table 3.3.

Table 3.3: **Distribution of responses across experimental conditions.** Participants contributed to multiple responses per cell. Observations in which participants expressed neither a belief and nor an expectation (exactly 50% of certainty or chance, or "don't know") were omitted from the analyses. N = 9,855.

Single-Report			
Belief Contradiction	Belief Reinforcement	Expectation Violation	Expectation Validation
1,213 (24.362%)	1,201 (24.122%)	1,269 (25.487%)	1,296 (26.029%)

Double-Report			
Belief Contradiction + Expectation Violation	Belief Contradiction + Expectation Validation	Belief Reinforcement + Expectation Violation	Belief Reinforcement + Expectation Validation
1,194 (24.487%)	1,242 (25.472%)	1,221 (25.041%)	1,219 (25.000%)

Belief *contradictions* and *reinforcements*, as well as expectation *violations* and *validations*, were the manipulated variables and were coded as binary (0/1), with contradictions and violations as the respective reference categories. Perceptions regarding the stimuli,

namely *surprise* and *interestingness*, and *sharing intention* constituted the outcome variables and were measured on 5-point Likert scales from 0 ("not at all") to 4 ("extremely").

Answers to the *familiarity* and *knowledge* statements were included as control variables and measured on 5-point Likert scales from 0 ("strongly disagree" and "not at all") to 4 ("strongly agree" and "extremely"). Topics were also included as control variables, measured using dummy variables, with the AI topic as the reference category. Additional control variables included demographic information provided by Prolific: *education* was measured by distinguishing between those with a college degree (1) and those without (0), *partisan affiliation* distinguished between Republicans (1) and Democrats (0), and *gender* differentiated between females (1) and males (0).

The analysis proceeds in three steps. First, I describe the correlations between the outcomes to provide a general overview of how they are related to one another. Second, I test the hypotheses concerning surprise and interestingness: H1 tests whether expectation violations are more surprising than belief contradictions; H2a and H2b test whether belief contradictions and expectation violations are more interesting than reinforcements and validations. Third, I examine the determinants of sharing intentions: H3 and H4 test whether reinforcements and violations are more likely to be shared than contradictions and validations; H5 tests whether messages that simultaneously reinforce beliefs and violate expectations are more likely to be shared than messages that simultaneously reinforce and validate beliefs and expectations.

For all hypothesis tests, I treat the outcomes of interest as interval scales from 0 to 4, and regress them on the experimental conditions and the control variables: topic, familiarity, knowledge, college education, partisan affiliation, and gender. I specify robust linear regression models with standard errors clustered at the participant level to account for repeated observations per participant. Differences in outcomes between the conditions are also presented as contrasts estimated from these models.

The models testing H1–H4 use the data from the single-report questionnaire, as they test the isolated effects of discrepancies with an individual's assumptions, while the model testing H5 uses data from the double-report questionnaire. Across all models, I report the unstandardized coefficients of the control models accompanied by p-values and 95% confidence intervals (CIs). Results were subjected to robustness checks in the form of

ordinal logistic regression models, and these are reported in Appendices 3.3 and 3.4.

3.5 Results

I start by evaluating the relationships between *surprise*, *interestingness*, and *sharing intention*. Next, I examine the effects of discrepancies on experiences of surprise and interestingness (H1, H2a, and H2b). Finally, I assess the roles of reinforcements, violations, and their combination in shaping sharing intentions (H3–H5).

3.5.1 Overview

As shown in Table 3.4, all outcomes are positively related to each other. Surprise has a medium correlation with interestingness ($r = 0.300$, CI = [0.280, 0.320], $p < 0.001$), meaning that the more surprising a message is perceived by an individual, the more interesting it is perceived to be. Surprise has a small correlation with sharing intention ($r = 0.151$, CI = [0.130, 0.170], $p < 0.001$), which means that, on average, participants are somewhat inclined to share surprising messages, but to a limited extent. Finally, interestingness shows a medium to strong correlation with sharing intention ($r = 0.401$, CI = [0.380, 0.420], $p < 0.001$), meaning that interesting messages are more readily intended to be shared than surprising messages.

Note that these correlations are "raw" and do not control for potential confounders or account for the repeated measures of participants. For a more precise analysis, I turn to the hypothesis testing.

Table 3.4: **Correlations between outcomes.** Simple correlations between surprise, interestingness, and sharing intention. All variables were measured with 5-point Likert scales from 0 ("not at all") to 4 ("extremely") and treated as interval scales. These correlations were computed using a combined dataset from the single- and double-report questionnaire versions. $N = 9{,}855$; $p < 0.001$.

	Surprise	Interestingness	Sharing Intention
Surprise	1.000		
Interestingness	0.300 ***	1.000	
Sharing Intention	0.151 ***	0.401 ***	1.000

3.5.2 Surprise and interestingness

The models testing H1, H2a, and H2b are shown in Table 3.5, while the differences between experimental groups are displayed in Figure 3.1. Participants find messages with reinforcements of beliefs less surprising than those that contradict them ($b = -1.040$, CI $= [-1.523, -1.134]$, $p < 0.001$). Similarly, the validation of expectations is less surprising compared to their violation ($b = -1.001$, CI $= [-1.088, -0.913]$, $p < 0.001$). These results do not support H1, as the effects are comparable in magnitude: the relative increases in surprise are substantial and similar across messages with belief- and expectation-based discrepancies.

Table 3.5: **Models for the tests of H1, H2a, and H2b.** Models are estimated with robust standard errors clustered at the participant level. $N = 564$; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Surprise (H1)		**Interestingness (H2a & H2b)**	
Intercept	1.928 ***	1.909 ***	2.350 ***	2.178 ***
Reinforce (vs. Contradict)	−1.034 ***	−1.040 ***	−0.197 ***	−0.210 ***
Validate (vs. Violate)	−0.997 ***	−1.001 ***	−0.077	−0.080 *
Topic (vs. AI)				
Climate Change	−0.027	−0.091	−0.078 *	−0.417 ***
Migration	0.039	0.021	−0.297 ***	−0.370 ***
Nutrition	0.032	0.011	−0.214 ***	−0.352 ***
Population Growth	0.131 *	0.120 *	−0.281 ***	−0.358 ***
Familiarity		0.033		0.095 ***
Knowledge		0.017		0.166 *
College (vs. No college)		0.084		0.116 ***
Republican (vs. Democrat)		−0.097		−0.395 ***
Female (vs. Male)		−0.096		−0.042
Adjusted R^2	0.141	0.145	0.013	0.089

Regarding interestingness, the results show that belief reinforcements are perceived as less interesting than contradictions ($b = -0.210$, CI $= [-0.288, -0.131]$, $p < 0.001$), and expectation validations are less interesting than violations ($b = -0.080$, CI $= [-0.157,$

−0.004], $p = 0.040$). These differences support H2a and H2b. However, the contrast between belief contradiction and reinforcement seems larger than that between expectation violation and validation. This suggests that belief discrepancies play a stronger role in the experience of interestingness.

Finally, none of the control variables affect surprise[5], save the population growth topic, which is on average more surprising than the AI topic. As for interestingness, messages about every topic are less interesting than messages about AI. Both greater familiarity with and knowledge of the topic being evaluated increase the perceived interestingness of the messages. Additionally, those participants with a college degree and who identified as Democrats find the messages more interesting.

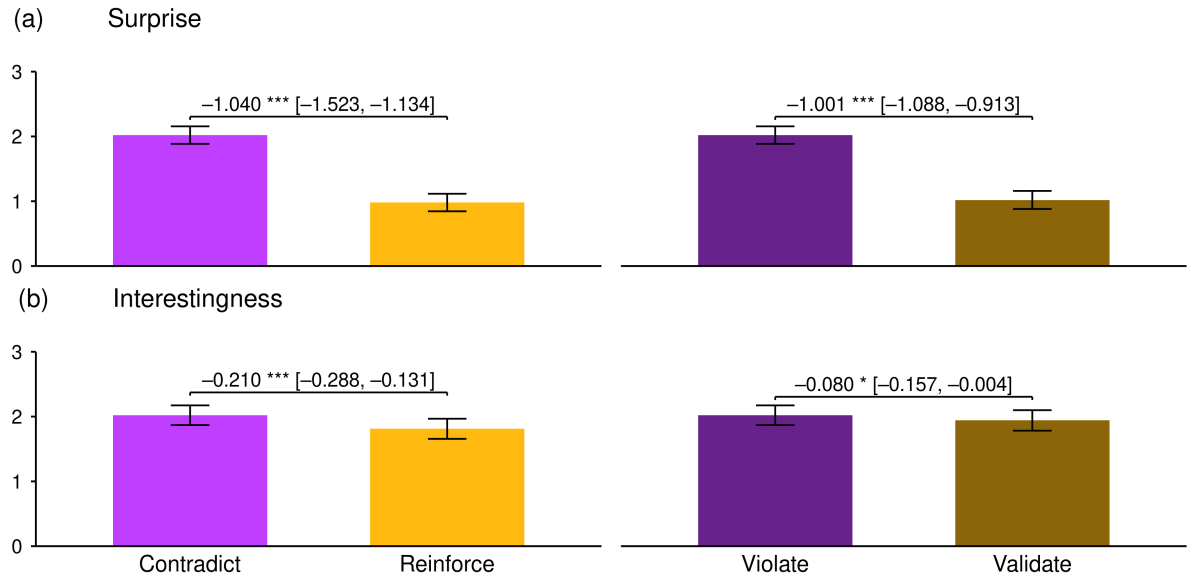


Figure 3.1: **Perceived surprise (a) and interestingness (b) by experimental condition.** Estimated means (error bars show 95% CIs) from robust linear regression models (Table 3.5, control models) with standard errors clustered at the participant level. Values represent responses on a 0–4 scale (0 = "not at all," 4 = "extremely"). Estimated contrasts between conditions are shown above the bars, with 95% CIs in brackets. $N = 564$; *** $p < 0.001$, * $p < 0.05$.

3.5.3 Sharing intention

The models testing the hypothesized effects of the experimental manipulations on participants' sharing intention are shown in Table 3.6, while the estimated differences between

[5]The ordinal logistic regression model does yield significant parameters for education, partisan affiliation, and gender. For details, see Appendix 3.3.

experimental groups are displayed in Figure 3.2. The results do not support H3, as messages that reinforce beliefs are not more likely to be shared than messages that contradict them ($p = 0.198$). Messages that validate expectations are more likely to be shared than those that violate expectations ($b = 0.148$, CI = [0.073, 0.224], $p < 0.001$), even if the effect is small. This result is not in the expected direction and does not support H4.

Table 3.6: **Models for the tests of H3–H5.** Models are estimated with robust standard errors clustered at the participant level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Sharing Intention			
	Single-Report (H3 & H4)		**Double-Report (H5)**	
Intercept	0.828 ***	0.355 ***	0.842 ***	0.270 ***
Reinforce (vs. Contradict)	0.054	0.048	0.032	0.022
Validate (vs. Violate)	0.136 ***	0.148 ***	0.033	0.043
Reinforce × Validate			0.085	0.078
Topic (vs. AI)				
Climate Change	−0.030	−0.524 ***	−0.018	−0.471 ***
Migration	−0.174 ***	−0.304 ***	−0.105 **	−0.211 ***
Nutrition	−0.076 *	−0.280 ***	−0.109 **	−0.284 ***
Population Growth	−0.172 ***	−0.300 ***	−0.168 ***	−0.279 ***
Familiarity		0.144 ***		0.105 ***
Knowledge		0.237 ***		0.249 ***
College (vs. No college)		0.079		0.093
Republican (vs. Democrat)		0.022		0.124 *
Female (vs. Male)		−0.101		0.017
Adjusted R^2	0.005	0.114	0.004	0.103
Participants	$N = 564$		$N = 1{,}208$	

Regarding the control variables, all topics are less likely to be shared on average than the AI one. Familiarity with and knowledge of the message's topic is associated with an increased sharing intention. Education, partisan affiliation, and gender did not affect sharing intention[6].

[6]The ordinal logistic regression model does yield significant parameters for education and gender. For details, see Appendix 3.4.

In the double-report questionnaire, at first glance, the manipulations do not seem to produce differences in sharing intentions. Neither the simple effects of reinforcing beliefs ($p = 0.607$) or validating expectations ($p = 0.319$), nor their interaction term ($p = 0.205$) are statistically significant. However, the contrasts shown in Figure 3.2b indicate a small but significant difference between reinforcements and contradictions when expectations are validated ($b = 0.100$, CI = [0.011, 0.188], $p = 0.028$).

This suggests that, while messages that simultaneously violate expectations and reinforce beliefs do not differ in their likelihood of being shared, messages that *validate* expectations and reinforce beliefs are more likely to be shared than those that also contradict beliefs. These results, however, should be interpreted with caution, as neither the main effects nor the interaction term are significant, and the evidence relies on within-group contrasts with a small effect size. Overall, the evidence does not support H5.

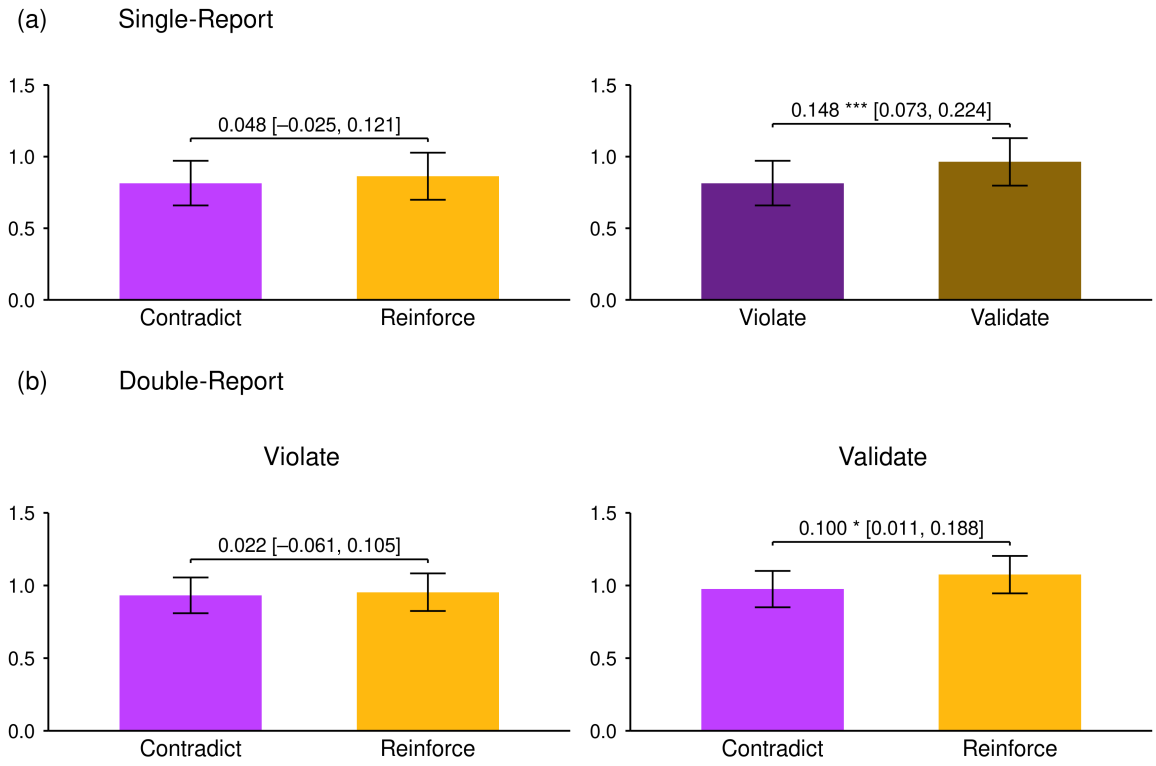


Figure 3.2: **Sharing intention for the single-report (a) and double-report (b) questionnaires by experimental condition.** Estimated means (error bars show 95% CIs) from robust linear regression models (Table 3.6, control models) with standard errors clustered clustered at the participant level. Values represent responses on a 0–4 scale (0 = "not at all," 4 = "extremely"). Estimated contrasts between conditions are shown above the bars, with 95% CIs in brackets. *N(a)* = 564, *N(b)* = 1,208; *** $p < 0.001$, * $p < 0.05$.

As for control variables, the results mirror those of the single-report questionnaire:

messages about AI were the most likely to be shared, and both familiarity with and knowledge of a topic increased the message's sharing intention. While education and gender had no significant effects, Republicans did report higher sharing intentions than Democrats.

In sum, messages that violate expectations are as *surprising* as messages that contradict beliefs. While both belief contradictions and expectation violations are more *interesting* than reinforcements and validations, respectively, belief contradictions seem to activate a channel that plays a stronger role in shaping the experience of interestingness. The validation of expectations is associated with an increase in *sharing intention* relative to their violation, whereas reinforcing beliefs does not increase sharing intention compared to contradicting them. The effect of jointly manipulating interactions with beliefs and expectations on sharing intentions is tenuous, as some evidence suggests that when expectations are validated, reinforcing beliefs increases the sharing intention of a message relative to contradicting them.

3.6 Discussion and conclusion

The spread of misinformation in online environments is a pressing challenge in contemporary societies (Guess et al., 2018; Rao et al., 2022). Many individuals rely on the internet as their primary source of information (Cuenca & Pérez, 2025; Wilson et al., 2020). While users may not intentionally share inaccurate content, they can be distracted from a piece of information's accuracy (Pennycook & Rand, 2021) and instead share it based on other characteristics. This is especially concerning for expert information, which helps to reduce complexity and inform public policy addressing societal risks (Gauchat, 2023; Shapin, 2008).

Since accuracy may not be the most salient factor in sharing intentions, this study focuses on the interaction between information and personal assumptions. Drawing from relevance theory (Wilson & Sperber, 1995, 1998) and information theory (Shannon, 1948), I distinguish between beliefs and expectations to examine how discrepancies with each shape perceptions of surprise and interestingness, and sharing intentions. I hypothesize that expectation violations are more surprising than belief contradictions, discrepancies

with both beliefs and expectations are more interesting than their correspondences, and sharing intentions are driven by confirmation bias (favoring belief-confirming information) and informational value (favoring expectation violations). Finally, I test whether simultaneously reinforcing beliefs and violating expectations increases a message's sharing likelihood compared to reinforcing beliefs and validating expectations. These hypotheses were tested in a pre-registered online experiment with US participants.

The analyses of the experimental manipulations show that *both* belief contradictions and expectation violations are more surprising than reinforcements and validations, respectively. Contrary to what was hypothesized, expectation violations are not more surprising than belief contradictions. In other words, discrepancies with either type of personal assumptions can elicit surprise. This finding is not fully consistent with the definition proposed by Barto et al. (2013), which emphasizes expectation violations as the primary source of surprise. As such, these results suggest that surprise may arise through both channels, originating from discrepancies with both assumptions, and that they may not be easily distinguishable at the phenomenological level.

As hypothesized, both belief contradictions and expectation violations are perceived as more interesting than reinforcements and validations, respectively. However, the belief-related channel seems to play a stronger role than the expectation-related one. This finding helps to bridge the relevance framework of Wilson and Sperber (1995, 1998) to the experience of interestingness (Altay et al., 2022). Overall, the results suggest that belief contradictions may be particularly influential in shaping the characterization of a message as interesting.

The results show that beliefs do not play a role in shaping the sharing intentions of expert information. Contrary to expectations, reinforcing a belief does not affect sharing intention compared to contradicting it, neither in the single- nor in the double-report versions of the questionnaire. Against what was hypothesized, the *validation* of expectations improves a message's sharing intention. When considering messages that simultaneously interact with beliefs and expectations, there is tentative evidence that reinforcing beliefs, compared to contradicting them, improves sharing intentions of expert information, but *only* when also validating expectations. This finding is tentative, with a small effect size and based on within-group contrasts.

The importance of expectation validations in shaping sharing intentions is consistent with accounts that characterize our brains as "proactive": constantly generating predictions about what we will soon experience (Bar, 2007; Brown & Brüne, 2012; Clark, 2013). Within this framework, incoming information is continuously evaluated against these expectations. Messages that validate expectations may be perceived as valuable, as they align with this continuous predictive process, and are consequently shared.

Taken together, these results suggest that discrepancies with both sets of personal assumptions drive perceptions of surprise, and contradicting beliefs is key to the interestingness of a message. Sharing intentions are more strongly shaped by expectation *validations*. Even if expectation violations are theorized to convey greater informational value (Baldi, 2002) and are more surprising, they are not more likely to be shared, a pattern consistent with Goebel et al. (2024). Overall, even if perceptions of surprise and interestingness are shaped by discrepancies with both sets of assumptions, and these may be hard to disentangle as exclusive sources of surprise, participants do react differently to them when evaluating whether to share expert information.

This study has several limitations. First, the sample is restricted to the US and is not nationally representative. Although it was balanced across education, partisan affiliation, and gender, improving internal validity, this design limits the generalizability of the findings. Moreover, the results may not extend to other national contexts. Future research should examine whether the effects of the experimental manipulations replicate in different settings and with representative samples. Second, the experimental manipulations in the double-report questionnaire may have been too long for participants to consider sharing, which may explain why significant differences were only detected through within-group contrasts. Prior research shows that shorter messages are more likely to spread than longer ones (Bojarski et al., 2022; Gligorić et al., 2019). Future research should examine messages that simultaneously interact with beliefs and expectations while accounting for message length. Third, this study focuses on perceptions and sharing intentions of expert information, which primarily involves epistemic claims. As such, the findings should be interpreted within this scope, and future research should examine whether the observed effects extend to other domains, such as political content. Finally, while sharing intentions are related to actual sharing (Mosleh et al., 2020), this study does not examine behavior. Future research should use observational designs to identify how interactions with

assumptions shape sharing behavior with social media data, or employ field experiments to test their effects in social media (e.g., Jones et al., 2023; Munger, 2017).

To conclude, US participants react differently depending on whether expert claims interact with their beliefs or their expectations. Both belief contradictions and expectation violations increase surprise, suggesting that surprise is not uniquely tied to expectations in practice. However, belief contradictions are more strongly associated with interestingness than expectation violations, suggesting that interestingness is more strongly activated by belief-related channels than expectation-related ones. In contrast, sharing intentions are primarily associated with expectation validations rather than violations. Although expectation violations are more surprising and carry more informational value, they do not translate into greater willingness to share, pointing to a disconnect between informational value and sharing intentions that future work should examine.

Chapter 4

Scientific Authority Cues Increase the Spread of Misinformation

This chapter, co-authored with Ismail Harrando and Achim Edelmann, has been accepted for publication in *Proceedings of the National Academy of Sciences*.

Abstract

Despite extensive efforts to curb misinformation, it continues to spread on social media platforms. Research indicates that people rarely share content they know is false, rather, they do so unintentionally due to cognitive biases triggered by contextual cues. Combining a large-scale behavioral data analysis with an experimental study, we investigate how *Scientific Authority Cues*, which we define as mentions of entities that signal scientific expertise or findings, affect the sharing of fact-checked content on social media by granting it credibility. Analyzing 8.7 million tweets linking to fact-checked content shared on Twitter (now X) between 2010 and 2023, we find that tweets with such cues are shared more, especially when linking to low veracity content or posted by right-leaning users. In a pre-registered experiment deployed to US participants (N = 1,187), headlines that attributed claims to scientific entities elicited a greater willingness to share, an effect statistically mediated by perceived accuracy and influenced by trust and familiarity. Together, these findings reveal the double edge of invoking scientific authority: in polarized online spaces, cues that signal scientific credibility can be co-opted to amplify misinformation.

4.1 Summary

Despite extensive efforts to curb it, misinformation continues to spread on social media platforms. Research indicates that people rarely share content they know is false (Pennycook et al., 2021). Rather, they do so unintentionally due to cognitive biases triggered by contextual cues (Ceylan et al., 2023). One of these cues could be the presence of scientific entities as sources of information, who signal authority and give legitimacy to claims (Bode et al., 2021; Vraga & Bode, 2017). In this final chapter, we examine how *Scientific Authority Cues* (SACs), defined as mentions of entities that signal scientific expertise or findings, affect the spread of fact-checked expert information on social media.

The results show that attributing claims to scientific entities increases willingness to share, and this effect is statistically mediated by perceived accuracy. Additionally, the effect of entity attribution on sharing intention is specific to false claims among Republicans.

Together, the findings provide answers to the experts-as-sources question: experts improve the perceived accuracy and sharing of claims. However, science's authority can be exploited online to lend credibility to misinformation. Both the misinformation we observed on Twitter and the false headlines used in the experiment featuring SACs were particularly appealing to Republicans, suggesting that scientific authority is being co-opted to legitimize misinformation that appeals to right-leaning online users. Overall, political identification moderates the influence of expert sources on sharing, with effects concentrated among Republicans when scientific authority is used to lend credibility to misinformation.

4.2 Introduction

Science is widely regarded as essential for democracies (Goldman & Chenoweth, 2025; Jasanoff, 2005). It anchors public debate in shared standards of evidence and provides expertise on issues beyond the public's reach, supporting decision making on challenges from public health to climate change (Algan et al., 2021; Cologna & Siegrist, 2020). Research in sociology and science and technology studies shows that science fulfills this

role not only through the knowledge it produces but also because of the social trust and institutional legitimacy it commands (Gauchat, 2010; Gieryn, 1999; Hendriks et al., 2016; Jasanoff, 2005; Shapin, 1994, 2008). Cross-national evidence further shows that this trust operates at both individual and societal levels: for example, personal trust in science predicts vaccine confidence (Rozek et al., 2021), and societies with higher trust display stronger public consensus (Sturgis et al., 2021).

While science continues to garner trust from the majority of people (Cologna et al., 2025), this trust often divides sharply along political lines. In the United States, trust in science has declined substantially among conservatives since the 1970s (Funk et al., 2019; Gauchat, 2010; Li & Qian, 2022), prompting concerns about a "legitimacy crisis" of science driven by partisan divisions (Furnas et al., 2025; Gauchat, 2023; Kennedy et al., 2022; Lupia et al., 2024). Similar trends appear in other democracies, where partisanship influences acceptance of scientific consensus on climate and related policy domains (Falkenberg et al., 2022; Funk et al., 2020; Merkley, 2020).

This legitimacy crisis is further exacerbated by the increasingly polarized online mediascape where content vies for attention, truth is contested, and misinformation continues to spread (Del Vicario et al., 2016; Ecker et al., 2025; Iyengar & Massey, 2019; Pennycook & Rand, 2021; Suarez-Lledo & Alvarez-Galvez, 2021). Given users' limited capacity to verify information independently, they rely on heuristics—information processing strategies that reduce cognitive load (Chaiken, 1980; Metzger & Flanagin, 2013; Petty & Cacioppo, 1986). Research has linked such reliance on heuristics to the diffusion of misinformation. For instance, social media platforms amplify emotionally charged, identity-signaling, and partisan content, making misinformation not only a cognitive but a social and political phenomenon (Acerbi, 2019; Altay et al., 2023; Bryanov & Vziatysheva, 2021; Guess et al., 2019; McLoughlin et al., 2024; Mosleh & Rand, 2022; Rathje et al., 2023; Scheufele & Krause, 2019). As a result of the reliance on such heuristics, people often inadvertently share misinformation (Ceylan et al., 2023; Pennycook & Rand, 2019; Vellani et al., 2023).

Mentions of scientific entities—such as scientists, universities, or studies—can function as signals that trigger heuristics of authority and recognition, a process that is amplified on digital media (Metzger et al., 2010; Sundar, 2008). Prior work suggests that source and credibility cues can shape how information online is perceived and shared (Bode et al.,

2021; Hendriks et al., 2016; Hu & Shyam Sundar, 2010; Pornpitakpan, 2004; Taddicken & Krämer, 2021; Vraga & Bode, 2017).

Yet little is known about whether cues that invoke scientific authority influence information diffusion at scale. In particular, it remains unclear whether invoking science enhances the reach of truthful information or instead amplifies misinformation that borrows scientific legitimacy.

This study therefore provides the first large-scale evidence, supported by experimental evidence, that references to scientific entities influence the diffusion of content on social media (see Fig. 4.1 for an overview of the study design).

We first analyze a corpus of 8.7 million tweets linking to professionally fact-checked media content shared on Twitter (now X) between 2010 and 2023, drawn from two different databases (*Condor*[1] and *Science Feedback*[2]). By focusing on links that were flagged for professional verification, this data captures the diffusion of contentious and high-stake information prominent in modern digital discourse, thereby reflecting an environment where scientific authority is arguably most frequently invoked to adjudicate hard-to-verify, counter-intuitive, and controversial matters.

Using large language model (LLM) annotations validated by human coders, we identify *Scientific Authority Cues* (SACs), which we define as mentions of *named* or *unnamed* entities in article titles and news headlines that signal scientific expertise or findings. Based on a qualitative examination of the data, we identify three types of scientific entities: individuals (e.g., "Dr. Fauci," "virologists"), institutions (e.g., "Harvard University," "research lab"), and sources (e.g., "The Lancet," "a study"). This approach enables us to estimate how invoking science influences content sharing and whether effects differ between true and false news, and between low and high veracity domains. For a subset of users for whom we can infer a political orientation, we test whether their responsiveness to these cues is affected by the aforementioned politicization of science.

We then complement this analysis with a pre-registered online survey experiment (N = 1,187) designed to support the observational findings in a U.S. sample and identify a

[1] *Condor* (part of *Social Science One*) comprises URLs circulating on Facebook that were flagged to Meta fact-checkers as potentially false (Messing et al., 2020).

[2] Science Feedback contains a curated set of predominantly scientific media content evaluated by a coalition of scientists as part of the *Science Feedback* project.

possible mechanism underlying these patterns. Participants were randomly assigned to view headlines from the observational dataset where references to scientific entities were either inserted or removed, allowing us to isolate the causal impact of attributing claims to such entities on participants' willingness to share. We further test whether this effect operates through perceived accuracy (as a measure of *credibility*) and explore how trust in and familiarity with scientific entities shape sharing intention and perceived accuracy.

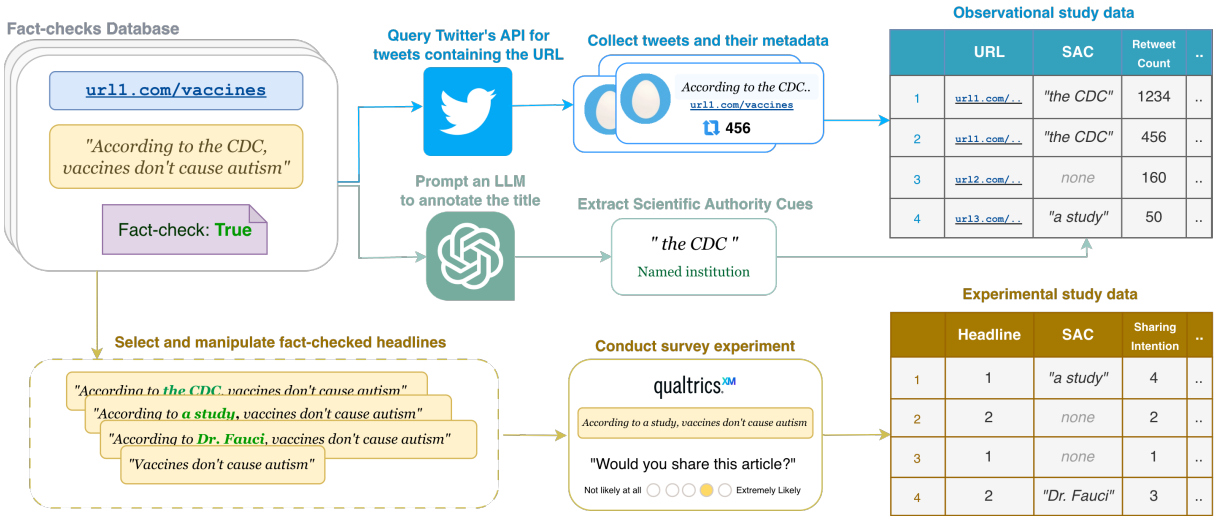


Figure 4.1: **Data pipeline summary.** *Observational study:* Starting from a database of fact-checked URLs (each associated with a title and a veracity rating), we use the Twitter (X) API to retrieve all tweets linking to the URL (either embedded in the tweet or included in its text). A large language model (LLM) is then prompted to identify the presence of Scientific Authority Cues (SACs) within each title. The effect of SACs is then estimated using negative binomial regression models on retweet counts. *Experimental study:* We select a set of fact-checked headlines and systematically manipulate attribution to a scientific entity. Survey participants are randomly shown headlines with or without such attribution and asked to rate their intention to share and perceived accuracy. The effect of scientific attribution is then estimated using robust linear regression models.

4.3 Results

To capture the large-scale impact of SACs on content spread online, we collected all tweets linking to fact-checked content from two different databases: *Condor* and *Science Feedback* (see Appendix 4.1 and 4.2 for details on data collection and annotation, respectively). The final dataset contains 8.7 million original tweets posted by 3 million users, linking to 56,508 URLs and generating 30.9 million retweets. Because our observational sample relies on fact-checking databases, it is not representative of the entirety of information on

Twitter. Rather, it reflects a selective universe of contentious and viral news and claims that warranted professional verification. We explicitly interpret our observational findings regarding SACs within this selected universe.

We refer to tweets linking to an article whose title includes a SAC—such as a mention of a scientist, university, or study—as *cued*, and the rest of the tweets as *uncued*. Because our primary interest lies in characterizing the dynamics of information diffusion, we operationalize the *SAC effect* as the proportional change in expected retweet counts between cued and uncued tweets. This effect is expressed as the incidence rate ratio (IRR) for the binary indicator *cued*, denoting the presence of a SAC. An IRR greater than 1 indicates that SACs are associated with an increase in the expected retweet count, whereas an IRR below 1 indicates a decrease. IRRs are reported alongside their 95% confidence intervals (CIs) and p-values, all estimated using negative binomial regression models. All models also include the tweeter's follower count (log-transformed) as a covariate to adjust for audience size. Any other model specifications or extra covariates are explicitly stated in the analysis and the table captions when used.

Beyond our main specification, we test the robustness of the results to the inclusion of temporal fixed effects (Appendix Table A14), adjusting for bot activity (Appendix Table A15), and accounting for user-level variance via standard errors clustered at the user level (Appendix Tables A27–A29) and mixed-models with random intercepts for users (Appendix Tables A30–A32). We also demonstrate robustness across alternative operationalizations of spread beyond retweet counts, including *sharing likelihood* and *overall spread*, and report the results in Appendix Tables A21–A26. Details on the robustness checks and alternative models are provided in Appendix 4.4. Overall, all robustness checks yield results consistent with the main specification in both direction and significance.

We first test the effect of SACs across the full dataset and for different cue categories (Study 1), and then examine two potential moderators: content veracity (Study 2) and users' political partisanship (Study 3). Full regression parameters are provided in Appendix Tables A1–A13.

Study 1: Scientific Authority Cues (SACs) increase content sharing

We begin by examining the overall impact of SACs on retweet counts across the full dataset. Overall, *cued* tweets, which represent 16.2% of the tweets in the corpus, receive significantly more retweets than those without (IRR = 1.723, CI = [1.711, 1.736], $p < 0.001$). The effect is positive and significant across both fact-checking databases despite differences in their selection criteria and topic coverage (*Condor*: IRR = 1.756, CI = [1.743, 1.769]; *Science Feedback*: IRR = 1.537, CI = [1.519, 1.555]; both $p < 0.001$).

We then test whether the effect is driven primarily by named (e.g., specific scientists, universities, or journals) versus unnamed (generic references such as "researchers" or "study") entities. Against a baseline of uncued tweets, named entity cues have a stronger effect than unnamed ones (*named*: IRR = 1.745, CI = [1.731, 1.758]; *unnamed*: IRR = 1.634, CI = [1.609, 1.658]; both $p < 0.001$).

Finally, disaggregating by entity type reveals that individuals produce the strongest effect, followed by institutions and sources (*individuals*: IRR = 1.662, CI = [1.646, 1.677]; *institutions*: IRR = 1.475, CI = [1.458, 1.493]; sources: IRR = 1.397, CI = [1.380, 1.415]; all $p < 0.001$).

In summary, tweets containing SACs receive substantially more retweets than those without, and this is consistent across both databases (*Condor* and *Science Feedback*). It is present for all cues categories, with the strongest effect observed for cues referencing named entities and individuals. For full regression tables, see Appendix Tables A1–A5.

Study 2: Scientific Authority Cues (SACs) amplify the spread of misinformation

We next assess whether the SAC effect varies by the veracity of the linked content. While every URL in our dataset was reported for fact-checking, they did not all receive a definitive true or false fact-check rating. We therefore rely on three complementary measures of veracity from three independent sources:

1. *Fact-check ratings* at the URL level from our databases (false vs. true).
2. *Factuality levels* from Media Bias/Fact Check (Media Bias/Fact Check, 2024) at the

URL domain level, simplified into three categories (high, mid, and low *factuality*).

3. *Credibility scores* as estimated by Science Feedback (Science Feedback, 2023) at the URL domain level, binned into three terciles (high, mid, and low *credibility*).

Since no veracity measure covers all URLs and domains, each analysis is estimated on the subset of tweets for which that measure is available, and results across subsets should be interpreted as complementary rather than directly comparable (see Appendix Table A33 for the distribution of URLs per veracity label, and Appendix 4.2 for more details on each measure).

For brevity, we collectively refer to URLs receiving a false fact-check rating, originating from a domain of low *factuality* or falling in the lowest *credibility* score tercile as *low veracity content*, and use *high veracity content* for opposite categories in all three measures.

Based on the fact-check ratings, 9.7% of URLs rated true contained a SAC, compared to 26.3% of false ones (see Appendix Table A33 for the full breakdown). The presence of a SAC has contrasting effects depending on veracity, that is, among tweets linking to false URLs, cued tweets receive more retweets (IRR = 1.487, CI = [1.474, 1.500], $p < 0.001$), whereas cued tweets linking to true URLs receive significantly fewer (IRR = 0.754, CI = [0.731, 0.777], $p < 0.001$).

Among tweets linking to URLs with a "neutral" fact-check rating (42.0% of the URLs that were either "not rated" or deemed "not eligible" by fact-checkers), we observe a strong positive effect (IRR = 1.775, CI = [1.747, 1.804], $p < 0.001$). This effect is concentrated in the "mid" category of both factuality levels and credibility scores (see Appendix 4.3, and Appendix Table A16).

For *factuality* levels, cued tweets linking to low factuality domains are shared substantially more often than their uncued counterparts (IRR = 1.948, CI = [1.922, 1.975], $p < 0.001$), whereas the corresponding effect for tweets linking to high factuality domains is positive but significantly smaller (IRR = 1.187, CI = [1.161, 1.213], $p < 0.001$).

Finally, for *credibility*, we observe a similar effect to that of factuality levels: that is, the effect is stronger for low credibility domains compared to high credibility ones (*high*: IRR = 1.309, CI = [1.288, 1.330]; both $p < 0.001$; *low*: IRR = 1.679, CI = [1.656, 1.701]). Using the raw scores, we find a significant negative interaction between the presence of

a SAC and domain credibility scores ($b = -0.492$, CI = [-0.530, -0.454], $p < 0.001$), indicating that SACs amplify the spread of less credible domains.

Together, these findings suggest that SACs consistently amplify the diffusion of low veracity content. For full regression tables, see Appendix Tables A6–A9.

Study 3: Asymmetric effect of Scientific Authority Cues (SACs) by political affiliation

Following existing research (Furnas et al., 2025; Gauchat, 2023; Kennedy et al., 2022; Lupia et al., 2024), we examine whether the effect varies across the political spectrum. We use the estimated political orientation labels (right-leaning or *Republican* versus left-leaning or *Democrat*) of tweeters in our dataset who followed at least one U.S. political elite account, as provided by (Mosleh & Rand, 2022). Because these labels only concern U.S.-American political parties, we restrict the analysis to tweets linking to URLs shared mainly in the U.S.

Across both groups, cued tweets receive more retweets, but the effect is substantially stronger for cued tweets from right-leaning users (*left*: IRR = 1.060, CI = [1.039, 1.089], *right*: IRR = 1.724, CI = [1.705, 1.743]; both $p < 0.001$).

This partisan asymmetry becomes more pronounced when disaggregating by the veracity of the content shared by partisan users. Among tweets posted by left-leaning users, cued tweets linking to URLs fact-checked as true are shared less frequently than uncued ones, whereas the opposite pattern holds for tweets linking to URLs fact-checked as false (*true*: IRR = 0.877, CI = [0.827, 0.931]; *false*: IRR = 1.097, CI = [1.063, 1.132]; both $p < 0.001$). Among tweets from right-leaning users, SACs consistently increase sharing, with a stronger effect observed for false news (*true*: IRR = 1.172, CI = [1.054, 1.303], $p = 0.003$; *false*: IRR = 1.510, CI = [1.490, 1.529], $p < 0.001$). The same pattern emerges when examining domain-level factuality levels and credibility scores: whereas SACs reduce the sharing of high factuality and high credibililty domains among left-leaning users and increase it among right-leaning users, they consistently amplify the spread of low factuality content for both groups (see Appendix Tables A12 and A13).

These findings suggest a partisan asymmetry in reactions to SACs, particularly among right-leaning users. Further robustness checks using users' bios to infer their ideologi-

cal orientation (*self-identified*), and with domain-level political bias labels (*source slant*), yield similar patterns (see Appendix 4.4 for the analysis description and Appendix Tables A17 and A18 for effect estimates). Moreover, supplementary analysis of "neutral" URLs exhibits the same political asymmetry: a significant amplification of diffusion for right-leaning users, but no effect for left-leaning ones (see Appendix Section and Appendix Table A16).

In summary, our observational analysis shows that SACs have a robust, positive effect on content diffusion on Twitter (X), with the strongest amplification observed for URLs fact-checked as false or associated with low *factuality* and low *credibility* domains, and that this amplification is particularly pronounced among right-leaning users.

To test whether these large-scale associations reflect causal effects and to uncover possible mechanisms linking SACs to sharing patterns, we conducted a pre-registered online experiment in the U.S., focusing on the credibility heuristic mechanism (N = 1,187; see pre-registration[3]). Participants were randomly shown variations of fact-checked headlines from our dataset, in which claims where either assigned or not assigned to a scientific entity, allowing us to isolate the causal impact of such attribution on willingness to share and to assess the roles of credibility, trust, and familiarity.

Experimental evidence

Building on the observational results, the experiment tests whether attributing a claim to a scientific entity causally increases people's willingness to share it. Moreover, based on prior misinformation research (Bauer & Clemm von Hohenberg, 2021; Pennycook et al., 2018, 2021), we test whether this effect is driven by a mechanism of credibility, measured as *perceived accuracy*, and which factors moderate it. All analyses followed our pre-registered hypotheses, reflecting the patterns identified in the observational study (H1–H5) and exploring the proposed mechanism (H6–H7) (for detailed hypotheses, see Appendix Section 4.6).

Participants were randomly presented with variations of real news headlines in which Scientific Authority Cues were either inserted or removed. For example, headlines such as "Vaccines are linked to sterilization of women" and "Seas are rising at the fastest rate

[3]https://osf.io/rpk9s

in the last 28 centuries" could appear with an added attribution "according to Harvard University" or "Study says" (see Appendix Table A36 for a complete list of manipulated stimuli). We refer to this manipulation as *Scientific Authority Attribution* (SAA), distinguishing it from *Scientific Authority Mentions* (SAMs), where the scientific entity is not directly linked to an explicit claim in the headline (e.g., "Strikes at Harvard University" or "What the WHO is hiding from you").

The experiment used a 2 x 2 design, crossing *Scientific Authority Attribution* (no attribution vs. attribution) with *claim veracity* (false vs. true). Participants evaluated 12 fact-checked headlines, first indicating their *willingness to share* each and later assessing their *perceived accuracy* ("not at all" to "extremely" likely to share and accurate, respectively). In addition, they rated their level of *trust* in and *familiarity* with the referenced scientific entities ("not at all" to "extremely" trustworthy and familiar with, respectively). Sharing intention and perceived accuracy were measured on 5-point Likert scales, treated as continuous, and analyzed using robust linear models, with results robust to the inclusion of control variables and to an ordinal specification (see Appendix 4.9). Unless stated otherwise in the text, effects are reported as estimated group contrasts derived from model parameters with 95% CIs and p-values from two-tailed tests (see Appendix Tables A38 for hypothesized differences, and A39 and A40 for exploratory differences). Full model estimates are shown in Appendix Tables A41–A45.

While we also ran a complementary analysis of SAMs in the same experimental setting (see Appendix 4.11 and Appendix Table A52), we choose to focus primarily on SAAs in this section for two reasons. First, SAAs allow us to directly test the credibility heuristic mechanism: by explicitly attributing a claim to a scientific entity, we can cleanly isolate the causal effect of scientific attribution on perceived accuracy and willingness to share. Second, unlike SAMs, SAAs are experimentally manipulatable with minimal assumptions, and are more readily susceptible to misuse in the context of misinformation (i.e., easier to add to any claim). Furthermore, we re-run the analyses of the observational studies using only SAA-like cues (which constitute around 75.5% of the SACs) as annotated by an LLM (see Appendix Table A34 for the breakdown of SAAs presence per analysis category, and the "Attribution vs Mention" in Appendix 4.2 for more details about the annotations), and we observe similar effects in significance and direction (see Appendix Table A19).

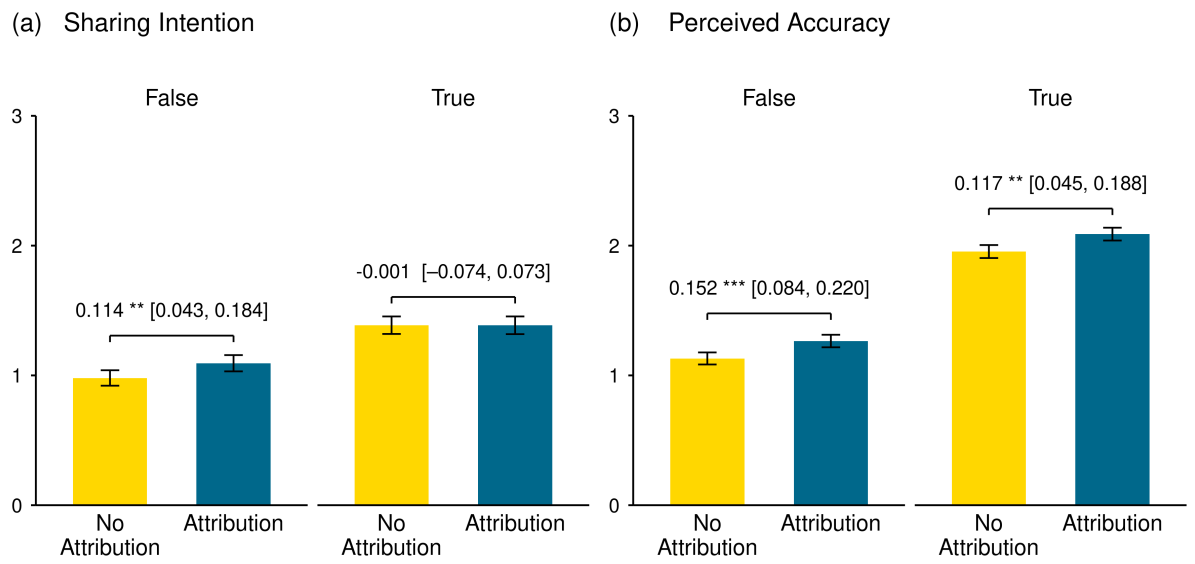


Figure 4.2: **Attributing claims to scientific authority increases both willingness to share and perceived accuracy.** Estimated means (95% CI) from robust linear regression models with participant-clustered standard errors. **(a)** Participants were more willing to share headlines attributed to a scientific entity, but only for false claims. **(b)** Participants rated headlines attributed to a scientific entity as more accurate. Values represent responses on a 0–4 scale (0 = "not at all," 4 = "extremely"). Estimated differences between conditions are shown above the bars, with 95% CIs in brackets. N = 1,187; *** $p < 0.001$, ** $p < 0.01$.

Scientific Authority Attribution effect (SAA) on sharing intentions

Attributing claims to a scientific authority increases participants' willingness to share them relative to no attribution ($b = 0.057$, CI = [0.006, 0.109], $p = 0.030$, H1), whereas the effect does not differ between named and unnamed entities overall ($p = 0.595$, H2). The SAA effect is driven by false claims ($b = 0.114$, CI = [0.014, 0.215], $p = 0.026$, H3; Figure 4.2a) and is non-significant for true claims ($p = 0.987$, H3).

Given the observed asymmetry of response to SACs in the observational study, we next examined whether the effect varies by party affiliation. While Democrats and Republicans respond similarly to attribution overall ($p = 0.836$, H4; Appendix Table A41), Republicans are more willing to share false claims attributed to a scientific entity than those without attribution ($b = 0.124$, CI = [0.021, 0.226], $p = 0.018$, H5).

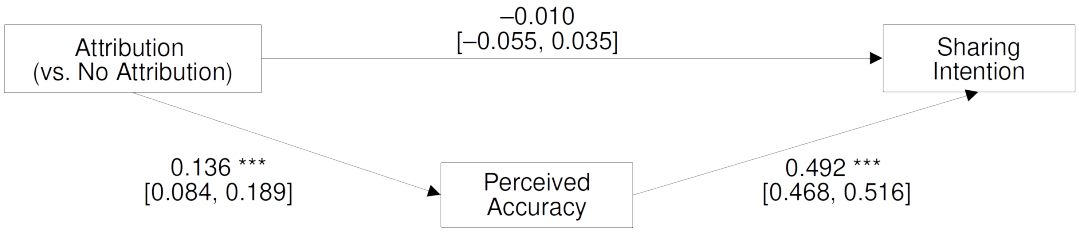


Figure 4.3: **Perceived accuracy statistically mediates the effect of scientific authority attribution on sharing intentions.** Displayed are unstandardized model parameters (95% CI) from a structural equation model with participant-clustered standard errors. The mediation analysis shows that attributing a claim to a scientific entity increases perceived accuracy, which in turn predicts higher willingness to share. The indirect effect ($b = 0.067$, CI = [0.041, 0.093], $p < 0.001$; Appendix Table A43) fully accounts for the total effect, indicating that credibility perceptions serve as mechanism linking authority attribution to diffusion. N = 1,187; *** $p < 0.001$.

Perceived accuracy as a mechanism

If scientific authority cues function as a heuristic for credibility (Bode et al., 2021; Metzger & Flanagin, 2013; Vraga & Bode, 2017), their influence on sharing should operate through an increase in the perceived accuracy of the claim, which in turn motivates the intention to share. We find that attributing claims to a scientific entity increases their perceived accuracy ($b = 0.136$, CI = [0.084, 0.189], $p < 0.001$, H6; Appendix Table A42, see Fig. 4.2b for a breakdown by veracity). Unlike sharing intention, this effect does not vary by claim veracity ($p = 0.475$; Appendix Table A42). A mediation analysis confirmed that perceived accuracy fully accounts for the link between attribution and sharing intention (*indirect* $b = 0.067$, CI = [0.041, 0.093], $p < 0.001$, H7; Appendix Table A43). Once perceived accuracy is included in the model, the direct effect of attribution on sharing disappears ($p = 0.672$; Fig. 4.3), suggesting that change in accuracy perception fully accounts for the effect.

Trust and familiarity

Guided by the heuristic account—which holds that people rely on authority and recognition heuristics to process information efficiently (Bauer & Clemm von Hohenberg, 2021;

Metzger & Flanagin, 2013; Van Der Heide & Lim, 2016)—we explored the roles of trust in and familiarity with the attributed scientific entity in shaping credibility and sharing intention (Appendix Figure A1). When participants trust the attributed entities, they perceive claims as more accurate and are more willing to share them (*indirect* $b = 0.055$, CI = [0.033, 0.077], $p < 0.001$, Appendix Table A43). Familiarity, in contrast, does not significantly increase perceived accuracy ($p = 0.692$, Appendix Table A43) but exerts a small, independent positive effect on willingness to share ($b = 0.071$, CI = [0.035, 0.108], $p < 0.001$; Appendix Table A43).

We next analyzed how trust and familiarity relate to sharing intention and perceived accuracy, focusing on the difference between claims attributed to trusted versus not trusted entities. We find that trust in the scientific entity is positively related to sharing for both true ($b = 0.506$, CI = [0.383, 0.628], $p < 0.001$) and false claims ($b = 0.127$, CI = [0.015, 0.239], $p = 0.027$), while accuracy judgments increase for true claims ($b = 0.450$, CI = [0.344, 0.555], $p < 0.001$) but not false claims ($p = 0.507$) (see Appendix Figure A2).

We then examined whether party affiliation moderated the role of trust for partisan differences across entities. For both groups, claims attributed to trusted entities are more likely to be shared than those attributed to not trusted entities (*Democrats*: $b = 0.272$, CI = [0.149, 0.395]; *Republicans*: $b = 0.471$, CI = [0.337, 0.605]; both $p < 0.001$; Fig. 4.4a). Attribution to a trusted entity increases perceived accuracy (*Democrats*: $b = 0.219$, CI = [0.106, 0.332]; *Republicans*: $b = 0.216$, CI = [0.091, 0.341]; both $p < 0.001$; Fig. 4.4b) equally across party groups ($p = 0.970$; Appendix Table A45).

In sum, the experiment provides causal evidence that attributing claims to scientific entities increases people's willingness to share them, possibly by enhancing their perceived accuracy. This effect is present for false claims and among Republicans. Among headlines with scientific attributions, trust in the referenced authority predicted both higher perceived accuracy and greater intention to share for both Democrats and Republicans. Familiarity showed a weaker pattern of association, whereby recognizing a scientific entity as familiar increases the intention to share a claim without strengthening its perceived accuracy. Together, these findings suggest that the same cues that can strengthen the communication of credible science may also be co-opted to amplify misinformation, functioning as heuristic shortcuts that signal credibility.

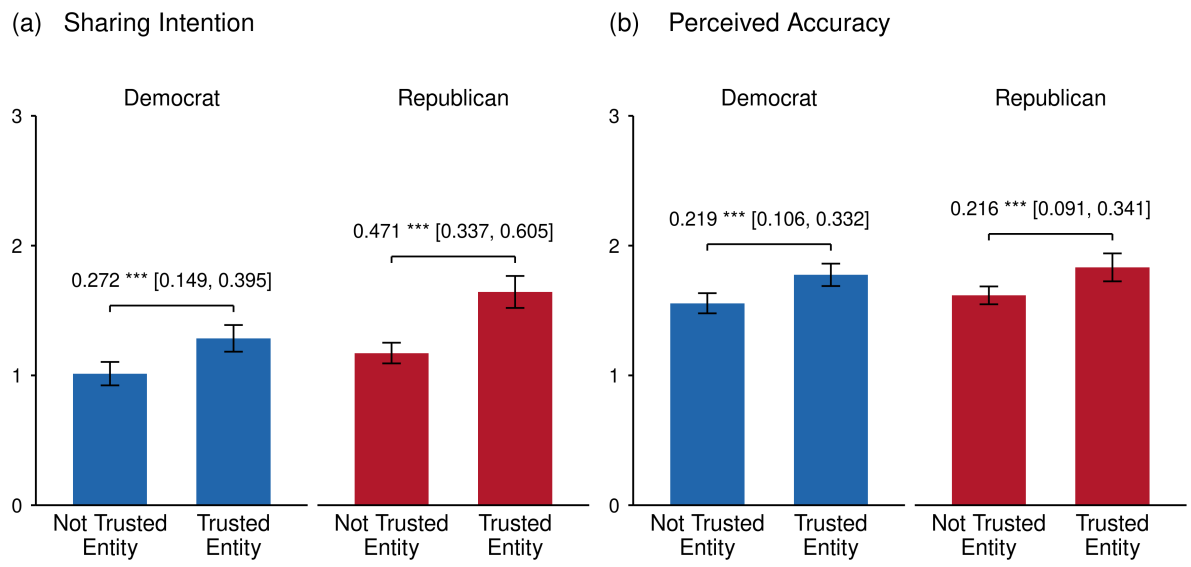


Figure 4.4: **Attributing claims to trusted entities enhances both sharing intention and perceived accuracy.** Estimated means (95% CI) from robust linear regression models with participant-clustered standard errors. **(a)** The increase in sharing intention for not-trusted versus trusted entities is significant for both Democrats and Republicans, but more pronounced for Republicans (for details, see Appendix 4.9, Tables A48 and A51). **(b)** The increase is also significant for perceived accuracy and does not differ between party affiliations. Values represent responses on a 0–4 scale (0 = "not at all," 4 = "extremely"). Estimated differences between conditions are shown above the bars, with 95% CIs in brackets. N = 1,182; *** $p < 0.001$.

4.4 Discussion and conclusion

Our findings reveal the double edge of invoking scientific authority in online communication: the same cues intended to convey expertise and evidence can be co-opted to amplify the spread of misinformation. Combining large-scale behavioral data on fact-checked content sharing on Twitter (X) with a pre-registered experiment administered to U.S. participants, we show that Scientific Authority Cues (SACs) increase the observed spread of content and the reported willingness to share headlines online. The findings in our observational and experimental data strongly suggest that these differentially advantage low-veracity content, especially among right-leaning people.

The experimental results support a credibility heuristic as the mechanism underlying this effect. In attention-demanding online environments, users rely on cognitive shortcuts to make rapid judgments about content credibility (Chaiken, 1980; Metzger & Flanagin, 2013; Petty & Cacioppo, 1986; Sundar, 2008). Our experiment provides causal evidence

that attributing a claim to a scientific entity increases individuals' willingness to share it by increasing its perceived accuracy. This suggests that scientific attributions function as influential signals of epistemic truth. Furthermore, among cued claims, higher trust in the attributed authority predicted both greater perceived accuracy and stronger sharing intention, whereas familiarity played a secondary, direct role by enhancing sharing intention without increasing credibility.

Our results also call attention to a vulnerability in the epistemic function of science within a politically polarized public sphere: instead of anchoring debates in shared standards of evidence beyond partisan boundaries (McNutt, 2024), invoking scientific entities is perceived differently across party lines, which risks deepening the divide between political groups. In our observational data, we find a pronounced asymmetry across political orientation in response to these cues: amplification, especially for false information, was significant among right-leaning users. This somewhat surprising finding adds an important qualification to prior research, which typically associates Republicans with a lower trust in science (Cologna et al., 2025; Gauchat, 2012; Gligorić et al., 2025). Our experimental study suggests that although Republicans have lower trust in scientific entities overall, attributing a claim to entities they trust still increases their willingness to share low-veracity content. This does not imply Republicans are uniquely susceptible to scientific authority; rather, it may reflect that misinformation on social media often criticizes or misconstrues science and more frequently advances conservative viewpoints (Garrett & Bond, 2021). The question, then, is not whether people trust science, but which science they trust (Kahan et al., 2011; Rekker, 2021).

While our experiment links scientific cues to sharing via perceived accuracy and trust partly explains this relationship, the pronounced partisan asymmetry found in our observational data cannot be attributed to this mechanism alone, nor can it be fully adjudicated from our observational design. Several additional explanations deserve consideration. At the structural level, selection effects and unobserved confounders may shape which content reaches right-leaning versus left-leaning users in the first place, and coordinated amplification (including bot activity) may inflate the apparent asymmetries in diffusion patterns (Ferrara et al., 2016; Vosoughi et al., 2018). At the individual level, this asymmetry may reflect broader variations in how individuals engage with science as an institution. For instance, analytic thinking and scientific sophistication (Drummond & Fischhoff, 2017;

Pennycook et al., 2019) can shape the extent to which people can evaluate the legitimacy of a claim, or recognize when scientific language is strategically deployed. These capacities are unevenly distributed and do not map neatly onto partisan identity (Pennycook & Rand, 2019). Additionally, identity-protective cognition (Kahan et al., 2011; Washburn & Skitka, 2018) may drive individuals to selectively embrace scientific authority only when it legitimizes ideologically or culturally congruent narratives. This dynamic intersects with anti-elite dispositions and conspiratorial ideation (Gauchat, 2012; Hofstadter, 2012; Petersen et al., 2023), which can simultaneously produce skepticism toward mainstream scientific consensus and the adoption of alternative "scientific" rhetoric to challenge established authorities (Aupers, 2012; Harambam & Aupers, 2015; Merkley, 2020). While our design cannot directly test these mechanisms or disentangle their contributions to the observed partisan gap, they offer theoretically grounded alternatives that future research could address.

Finally, the high prevalence of SACs we find in low veracity content (in 26.3% of URLs fact-checked as *false*, compared to 9.7% for *true*, 9.8% for *neutral*, see Appendix Table A33) points to another dimension of science's role in misinformation. Two non-exclusive mechanisms may account for this over-representation: on one hand, journalists or content producers may inadvertently report on sensational but questionable or preliminary science, potentially driven by incentives within the scientific community to dramatize findings for media coverage and recognition (West & Bergstrom, 2021); on the other, malicious actors may deliberately leverage scientific authority to legitimize false information for political, social, or financial gain. This pattern is further underscored by findings from the subset of tweets linking to URLs without an explicit rating. Here, the strongest SAC effect emerges at the intermediate levels of (domain) *factuality* and *credibility*, proving most potent in the gray area of potentially *deceptive content*. Unverified or inherently unverifiable, such content can evade professional and systematic fact-checking, making it potentially a greater threat (Allen et al., 2024). In all cases, the epistemic authority of science functions less as a truth signal than as a persuasion rhetoric.

Our study has several limitations. First, by focusing on fact-checked URLs, our observational data capture a selected universe of typically contentious and viral news and claims. Consequently, "true but fact-checked" content likely differs from typical true scientific content (e.g., it may be more counter-intuitive or hard-to-verify), which may explain

why SACs are associated with an attenuation of true content in our data. Additionally, our findings may not generalize to the broader, informal landscape of information sharing on social media. Second, the observational study cannot rule out unobserved confounders, meaning the reported large-scale associations remain descriptive. Moreover, the reliance on retweet counts as a proxy for information spread captures only one dimension of diffusion; other indicators like views, comments, or virality may reflect additional aspects of sharing. Third, our analysis is restricted to Twitter (X), where algorithms and network dynamics may not generalize to other platforms, and users of different ideological leanings may have distinct norms of retweeting. Fourth, the data primarily reflect sharing on Twitter (X) in the U.S., and our measure of political orientation relies on users' following of U.S.-American political accounts, warranting caution when generalizing beyond this context. Fifth, the observational and experimental designs are complementary but not fully comparable. On the observational side, LLMs could not identify claim-like titles with sufficient accuracy, precluding a direct comparison between explicit attributions (SAAs) and baseline claims without any scientific references—the contrast most analogous to our experiment. On the experimental side, while we identify an effect for scientific mentions (SAMs), we focused primarily on direct attributions (SAAs) to isolate the credibility mechanism. Bridging this operational gap remains an important direction for future computational and experimental work. Sixth, although robustness checks show no sensitivity to bot activity among posters and retweeters, we cannot fully rule out the influence of automated or coordinated network amplification on the observed retweet counts. Seventh, the experiment relies on researcher-manipulated stimuli that, while sourced from the observational data, may not fully capture the naturalistic ways scientific claims are framed on social media. Finally, the experiment measures self-reported sharing intentions rather than actual behavior. While this design validates the causal direction of the main effect—that attributions increase sharing via perceived accuracy—it does not capture its magnitude. The larger overall effect and more pronounced partisan gap in the observational data likely reflect platform dynamics and network amplification compounding this baseline individual susceptibility.

Despite these limitations, the complementary evidence from large-scale behavioral data and the experimental study offers the first account of how *Scientific Authority Cues* shape the diffusion of true and false information online, as well as a key potential mechanism

underlying it.

While science remains our most reliable means of understanding the world and addressing questions of public concern, the co-option of its authority to legitimize falsehoods constitutes a significant threat. The tension we explore—between science as a unifying anchor of credibility and science as a legitimizer of falsehood—suggests that simple admonitions to "trust science" and "check sources" may be insufficient, or even counterproductive, as discernment heuristics online. Recognizing this ambivalence is essential for curbing the spread of misinformation.

4.5 Methods

This study combines large-scale observational analyses of social media diffusion with a pre-registered online experiment to identify how scientific authority cues influence the sharing and perception of content online (see Figure 4.1). Across both studies, SACs are measured or manipulated as a property of the title because they are often the first (and sometimes only) piece of information users encounter before deciding to share content online.

4.5.1 Observational study

Data

In our study, we aim to examine the role of *Scientific Authority Cues* (SACs) on the sharing of information on social media. To construct our dataset, we collected fact-check ratings for URLs shared on social media from two primary sources. The first, *Condor*, comprises URLs circulating on Facebook that were flagged to Meta fact-checkers as potentially false. The second, *Science Feedback*, contains a curated set of predominantly scientific media content evaluated by a coalition of scientists as part of the *Science Feedback* project.

For each URL, we extracted the associated title and retrieved all tweets posted between January 2010 and April 2023 that either embedded the URL or included it in its text. Data collection, URL normalization, and metadata extraction were performed using the

Twitter API v2.0 with academic research access in combination with the *minet* (Plique et al., 2019) tool. We manually filtered out URLs with very short, uninformative, or missing title metadata. Tweets containing multiple URLs were also removed to avoid ambiguity in attributing ratings to tweets.

The resulting dataset comprises 8.7 million original tweets, generating 30.9 million retweets and linking to 56,508 distinct URLs with their corresponding titles and fact-check ratings. We simplify ratings into four categories: true, false, neutral, and other. We further augment these URLs with two external domain-level veracity measures: Media Bias/Fact Check factuality levels, and Science Feedback's Consensus credibility Scores. Comprehensive details on pre-processing procedures and the construction of fact-checking categories are provided in the Appendices 4.1 and 4.2.

Detection of Scientific Authority Cues (SACs)

We define *Scientific Authority Cues* (SACs) as mentions of scientific entities in article titles (headlines) such as individuals, institutions, or scientific sources like journals or studies. Entities can be either *named* (e.g., "Dr. Fauci", "CDC", "The Lancet") or *unnamed* (e.g., "Virologists", "German university", "a study"). To avoid ambiguity, we exclude entities that could qualify as scientific authorities but have strong commercial activity (e.g., "Pfizer").

We used a Large Language Model (OpenAI's *gpt-4o-mini*) to annotate each title for the presence of a SAC (see Appendix 4.5 for annotation prompts). To evaluate its output, we compared the model annotations against a manually labeled sample of 400 titles. The model achieved perfect recall (i.e., identified all entities recognized by human annotators) but exhibited lower precision, as it occasionally flagged entities that did not fit our definition.

To establish robust ground-truth labels, we conducted a secondary annotation round in which three independent human annotators re-evaluated the LLM outputs. Annotators were given a definition of scientific authority but also applied their own judgment and used external tools to verify whether the cue identified by the LLMs corresponded to a scientific authority. The reporting in the paper was done on the labels obtained by majority vote (which represents consistency across annotators), whereas results on different degrees of agreements were reported in Appendix 4.13. The description of the annotation process is

shown in Appendix 4.2.

Political orientation labels and analysis

Following (Mosleh & Rand, 2022), users were classified into the two political orientation groups (Democrats and Republicans) using a "political partisanship score." The scores were derived from the distribution of political elites followed by each user, ranging from -1 (only following Democratic political elite accounts) to +1 (only following Republican political elite accounts). They were computed and provided by the authors. Following their operationalization, we classify users as right-leaning or Republican if their score is positive and left-leaning or Democrat if their score is negative.

Given that we only have the scores for 65.2% of the users, we reproduce the analyses in Study 3 using two more external political labels: user political orientation as inferred from their bios, and a political bias label per domain (see Appendix 4.4). When analyzing interactions between the SAC effect and the user political affiliation, we restrict the sample to tweets sharing news primarily circulated in the U.S. (as identified in the metadata of the *Condor* dataset).

Analysis plan

The impact of a given variable on information spread can be measured in multiple ways, depending on data availability, medium characteristics, and assumptions about potential confounders. For instance, the depth of sharing cascades or the number of views can both reflect a content's virality and reach.

Because the Twitter API does not provide information on views, follower-followee networks, or retweet cascades (all retweets originate from the original poster, so no "cascade" information is preserved), we adopt a widely used operationalization from the literature (Bak-Coleman et al., 2022; Li & Chang, 2023; McLoughlin et al., 2024): estimating retweet counts using a Negative Binomial count regression, controlling for log-transformed follower counts. To reduce outlier influence and stabilize model fitting, retweet counts were winsorized at the 99.9th percentile.

Robustness checks

Beyond the main Negative Binomial specification, we tested the robustness of the SAC effect across a range of alternative model specifications and operationalizations, including:

- Logistic regression predicting whether a tweet gets retweeted at all (retweet likelihood)
- Negative Binomal count models aggregated at the URL level (overall spread)
- An expanded specification accounting for commercial entities with scientific activity
- Negative Binomal count models controlling for user-level variance via clustered standard errors and mixed-effects specifications
- Adding year fixed effects to account for temporal variance
- Adding Botometer's scores (Yang et al., 2020, 2023) to account for bot activity, and comparing bot scores for a sample of retweeters, stratified by engagement with cued content, veracity and political leaning

Full details for each specification are provided in Appendix 4.4 and Appendix Tables A14–A32. Across all models, the key observational effects remain consistent in sign and statistical significance, supporting the robustness of our findings.

4.5.2 Experimental study

Participants and recruitment

An analytic sample size of 1,190 participants was determined *a priori* to detect effects of size $d = 0.13$ with 90% power at an $\alpha = 0.05$. The effect size reflects a 33% reduction from a small one ($d = 0.20$) to account for variance inflation due to participant-clustered standard errors.

Participants were recruited via Prolific between 11 and 30 July 2025. All participants were at least 18 years old and resided in the U.S. Recruitment aimed for balance across party affiliation, gender, and education. In total, 1,439 participants completed the Qualtrics questionnaire.

Participants who failed either of two attention checks, gave identical answers to all perceived accuracy questions, or completed the survey unrealistically quickly (under 2.5 minutes) or slowly (over 16.5 minutes) were excluded. The final analytic sample comprised 1,187 participants (49.87% Democrat, 50.13% Republican; 50.25% female, 49.45% male; 50.21% no college, 49.79% college; age M = 43.9 years, SD = 14.3).

Design and stimuli

The experiment used a 2 x 2 within subject design, crossing *Scientific Authority Attribution* (SAA; no attribution vs. attribution) and *claim veracity* (false vs. true). Stimuli were drawn from a pool of 24 fact-checked headlines from our Twitter dataset, selected for topical diversity, current relevance, and a 1:1 false-to-true ratio.

16 headlines expressed factual *claims* that either already contained an SAA or could meaningfully be presented that way by attributing the claim to a scientific entity; 8 additional *news* headlines were not attributed to a scientific entity, half of which mentioned one—a condition we call Scientific Authority Mentions (SAM).

The SAA manipulation applied only to the 16 claims: for these, we randomly removed or inserted scientific entities in the headline as the source of the claim. Entities were of three kinds—individual, institution, or source (such as a journal or a study)—and could be named or unnamed. In the attribution condition, each claim was randomly paired with one of six possible entities, with randomization occurring independently per claim. Claim veracity was determined by independent fact-checking. Complete list of claims (with their variations) and news, are provided in Appendix Tables A36 and A37, respectively.

Procedure

After giving consent, participants viewed 12 headlines randomly drawn from the stimulus pool. The *claims* had a 50% chance of including a SAA and a 50% chance of no attribution. For each, participants rated how likely they were to share the headline online. After evaluating all headlines, they rated the perceived accuracy of each and their trust in and familiarity with the entities mentioned in attributed claims. Participants then completed demographic and political questions, followed by debriefing explaining that some posts

had been fact-checked as false and providing corrected information. The experimental study was approved by the Research Ethics Committee of the Paris Institute of Political Studies (Opinion #2025-080).

Analysis plan

All analyses followed the pre-registered plan. Robust linear models estimated the effects of SAA on sharing intention and perceived accuracy, with claim veracity and political affiliation as moderators. Structural equation models (SEMs) tested mediation paths via perceived accuracy and explored the roles of trust and familiarity as predictors. Participant-clustered standard errors accounted for repeated measures.

Sharing intention, perceived accuracy, trust, and familiarity were measured on a 5-point Likert scales from 0 (“not at all”) to 4 (“extremely") and treated as continuous. Categorical predictors (attribution, veracity, party affiliation, and trusted entity) were coded 0/1 with “no attribution,” “true,” “Republican,” and “not trusted” as reference categories. As robustness checks, we added control variables (Appendix Tables A46–A48) and repeated analyses using ordinal logistic regressions (Appendix Tables A49–A51). Detailed model specifications and hypothesis numbering (H1–H7) are reported in Appendix 4.6.

Conclusion

This thesis set out to move beyond knowledge-deficit accounts that attribute the contestation of expertise to ignorance, focusing instead on the factors that shape individual evaluations of experts and their claims in online, polarized contexts. Drawing from existing literature, I have argued that, even if the current context is sometimes characterized as a crisis of science's legitimacy (Gauchat, 2023; Saltelli & Funtowicz, 2017), science remains highly regarded in contemporary societies (Cologna et al., 2025; Pfänder et al., 2025; Shapin, 2008). Disagreements over science are more targeted at the experts involved in knowledge production and the claims they advance.

The increased use of expertise in governance since the second half of the last century (Dear & Jasanoff, 2010; Eyal, 2019; Gauchat, 2010) has provided grounds for contestation. The expansion of higher education (Archer et al., 2015; Drummond & Fischhoff, 2017; Kahan et al., 2012) and the rise of alternative information sources on the internet and social media (Castells, 2004) have provided citizens with the tools to challenge expertise. In the contemporary polarized environment (Wilson et al., 2020), these dynamics are amplified. Consequently, individuals who oppose specific applications of expertise in policy-making are better equipped to challenge experts and their claims, often along partisan lines. Accordingly, contestation is less about being pro- or anti-science than about preferences for how scientific knowledge is used in policy and governance.

What shapes public perceptions of experts and their claims?

The analyses in this thesis provide some answers, summarized in four main findings: evaluations of expertise are shaped by politically driven preferences, expert sources can lend credibility to claims in online environments, motivated reasoning may not account for all contestation to expertise, and the intention to share expert claims is more strongly

driven by the validation of expectations than by the confirmation of beliefs.

To what extent does political identification shape preferences for expertise along the distinction between impact science and production science?

Political identification shapes preferences for expertise to some extent. These preferences do not completely align with the distinction between impact science, research that addresses the consequences of industry and public policy, and production science, research used to enhance industrial production. The results in Chapter 1 show that, in the US, liberals trust scientists in general as well as both impact and production experts more than conservatives. Nevertheless, the ideological gap in trust is smaller for production than for impact experts. Liberals also perceive most claims as more accurate than conservatives, except for the taxation claim, which is more clearly aligned with conservative positions. The findings in Chapter 2 show a different pattern in Europe: left-leaning individuals express higher trust in science as a whole and impact science than right-leaning individuals. This preference is reversed for production science, where leaning more to the right is associated with higher trust.

Together, the results provide mixed support for the impact-production distinction proposed by Schnaiberg (1980) and McCright and Dunlap (2010). In the US, preferences are expressed as a general ideological gap in trust: liberals trust *all* experts more than conservatives, which aligns with recent evidence (Altenmüller et al., 2024; Gligorić et al., 2025), and suggests a diametrical sorting over expertise. In Europe, by contrast, trust aligns more closely with the theoretical distinction—replicating the results of Rekker (2025)—and may point to a public that is more engaged with expert information (Mejlgaard & Stares, 2013).

The mixed evidence reflects the limits of the theoretical distinction, as the boundary between impact and production science is often blurred in practice: research findings from many disciplines can be leveraged to support regulation *and* to advance industrial production. Additionally, the interpretations I present should be treated with caution, as they are based on a small number of indicators. Future research should continue to examine the impact-production distinction by letting the public evaluate a broader range of experts stereotypically associated with these types of science.

How does political identification shape the way expert sources influence the perception and sharing of expert claims?

Political identification does not affect how expert sources influence the perception of information. The results in Chapter 1 suggest that, regardless of political identification, including an expert as the source of a claim enhances its perceived accuracy. Findings from Chapter 4 similarly show that attributing claims to scientific entities increases their perceived accuracy. However, political identification does shape the relationship between expert sources and sharing intentions: the effect of scientific entity attribution on sharing is mainly driven by Republicans sharing false claims.

The correspondence of the results across studies shows that experts and scientific entities improve perceptions of accuracy, despite differences in design. In Chapter 1, participants evaluated expert claims formulated in neutral language, while in Chapter 4, the analyses focus on claims that appeared on social media and were reported for fact-checking. As such, the findings are robust to differences between controlled and naturalistic (or "viral") language.

The greater willingness of Republicans to share false information with scientific authority cues should not be treated as evidence that they are more easily "fooled" by misinformation than Democrats. Rather, it may reflect a tendency for misinformation circulating on social media to align with conservative views (Garrett & Bond, 2021).

Overall, the results suggest that attributing claims to experts does not drive different perceptions of the accuracy of information across political identifications: it consistently enhances them. By contrast, expert attribution may motivate sharing differently between Democrats and Republicans. Future research should examine whether the effect of expert sources on information perceptions and sharing extends to other national contexts and domains (e.g., political information).

Does motivated reasoning drive political preferences in perceptions of experts and their claims, and does this differ across general, impact, and production science?

Motivated reasoning does seem to shape the perceptions of experts and their claims, but not across all forms of expertise. In Chapter 2, the results suggest that the interac-

tion between political identification, and feeling close to a particular political party and education affects Europeans' trust in *impact* science: the more to the right, the less trust impact experts and claims, and this effect is more pronounced for those who are both closer to a party and more educated. Notably, left-right identification has no significant effect on trust in impact science at the lowest level of education among those who do not feel close to a party, but is significant among those who feel close to a party. Trust in production science and science in general was shaped by political identification, but it was not affected by these interactions in the same way.

This result suggests that there are left-right preferences for expertise, but motivated reasoning does not account for all of them. This interpretation should be treated with caution for two reasons. First, motivated reasoning was operationalized through two complementary interactions: political identification with closeness to a party, and political identification with education. This reflects the idea that motivated reasoning requires both the motivation to defend group beliefs and the ability to do so (Kunda, 1990). Closeness to a party addresses the former, and education serves as a proxy for the latter, capturing familiarity with science and expertise. However, higher education may be indicative of other factors, such as greater awareness of one's party's positions regarding expert claims (Drummond & Fischhoff, 2017). Second, the evidence for motivated reasoning in impact science rests on a specific subgroup comparison: left-right identification has no significant effect among those with low education and no party closeness, but does among those with low education who feel close to a party. Still, this pattern is worth noting. Future research should therefore operationalize this mechanism more directly, for example by using measures of science capital (Archer et al., 2015).

Despite these shortcomings, the results indicate that European attitudes toward impact science are different from attitudes toward production science and science in general. Disputes over expertise are not rooted in a general distrust of science. Instead, they may be *motivated* and directed at forms of science that are most visibly entangled with policy-making and regulation, with such contestation being more pronounced among those with stronger partisan commitments and higher levels of education.

Does the interaction between information and individuals' beliefs and expectations differentially shape the perception and sharing intentions of

expert claims?

The interactions between expert information and individual assumptions shape perceptions and sharing intentions of expert claims in different ways. As shown in Chapter 3, both belief contradictions and expectation violations are perceived as surprising. The perceived interestingness of a message, while shaped by discrepancies with both beliefs and expectations, is more strongly shaped by belief contradictions. Sharing intentions are more closely linked to the validation of expectations, rather than the confirmation of beliefs. Overall, beliefs and expectations seem to interact with information through different channels.

Taken together, the results show that information that captures an observer's attention does not necessarily translate into sharing intentions: surprising and interesting messages are not more likely to be shared, supporting the findings of Goebel et al. (2024). This suggests that the explanatory scope of relevance theory (Wilson & Sperber, 1995, 1998) and information theory (Shannon, 1948) may be better suited to explaining communication, as originally intended, than information sharing. Nevertheless, these frameworks are still useful for distinguishing between beliefs and expectations. Further research should build on these insights by refining how these assumptions are operationalized and by examining actual sharing behavior rather than intentions alone.

Chapter 3 presented the results of a first experiment distinguishing between types of individual assumptions and their roles in information perceptions and sharing. A follow-up experiment examining how the strength of these assumptions shapes responses to information, drawing from a more Bayesian perspective, is underway. The framework separating beliefs from expectations remains open to discussion.

Synthesis

In conclusion, this thesis has examined several factors that shape how expert information is perceived and shared. The results show that experts generally function as credible sources of information, improving the perceived accuracy of factual claims. At the same time, the findings indicate that contestation is not mainly directed at science as a whole, but at how expert authority and specific claims are used in practice. This distinction

complicates the common portrayal of liberals as pro-science and conservatives as anti-science. In the US, expertise evaluations divide more sharply along partisan lines, whereas European preferences for expertise are more dependent on the type of science in question. Specifically, contestation among Europeans appears motivated and concentrated around expertise associated with regulation. These patterns point to a tension: scientific expertise is broadly valued, yet subject to political contestation over specific claims and experts.

The findings also highlight the double-edged role of expert information in contemporary online environments. On the one hand, relying on partisan cues to assess factual claims may make the political and economic interests behind knowledge production more salient, but risks reducing epistemic judgments to political ones. On the other hand, deferring uncritically to expert authority may lead individuals to treat claims as neutral, overlooking the interests they serve. In both cases, perceptions and sharing intentions regarding expert claims are not shaped solely by their factual content but also by how information aligns with individual assumptions.

Broadly, rather than a crisis of science, this thesis suggests a more paradoxical context in which scientific expertise is highly valued, yet expert authority is contested. This contestation is not necessarily alarming, as it may reflect the public's awareness that knowledge production is a social and political process, and that claims can serve private interests or be used to justify political decisions. Skepticism is not always all-out rejection, but a push-back against uncontested authority over facts that reflects a substantive, sometimes critical, engagement with expertise. That said, not all contestation is equally grounded: skepticism toward expert claims about societal and environmental risks can have collective consequences, and the counter-claims it rests on are equally deserving of scrutiny regarding the interests they serve.

Two considerations follow from treating contestation as substantive engagement with expertise. First, acknowledging the entanglement of facts and interests does not imply that widely accepted claims are out of reach, as ongoing contestation and revisions can lead claims to converge on approximate truths over time (Dooremalen et al., 2015). Second, from a liberal naturalist perspective, which holds that common sense is a valid epistemic practice (Bellolio, 2019; Lavazza & Farina, 2022), public contestation can be harnessed through citizen science (den Broeder et al., 2018) rather than framed as an obstacle to

knowledge. In this sense, expertise may be legitimized not despite public contestation, but partly by incorporating it.

That said, my contributions remain limited in scope, focusing on US and European contexts and on online settings. The ways in which experts are evaluated, how individuals engage with expert information, and the meanings attached to expertise may differ in other contexts and real-world settings. Furthermore, while anti-elitist sentiment and the role of higher education in shaping familiarity with science were points of departure for contextualizing the "crisis of expertise", this thesis does not directly test whether these factors drive expert contestation. Similarly, the impact-production distinction rests on the assumption that impact and production science appeal to different political groups because of their associations with regulation and industrial production, which also remains untested.

Additionally, the findings rely heavily on null hypothesis significance testing (NHST), which has well-documented limitations (e.g., Cohen, 1994; Raftery, 1995; Wagenmakers, 2007). For example, p-values reflect the probability of observing data assuming there is no association (Concato & Hartigan, 2016) and, as such, do not indicate the probability of a hypothesis being true or false. Rather, they are measures of the compatibility of the data with a theory-driven model (Wasserstein & Lazar, 2016). Consequently, the empirical findings presented here remain open to critique. Nevertheless, NHST remains a widely used and valuable tool in social science.

As someone aiming to contribute to specialized knowledge, I find the vigilance toward expertise necessary and valuable. Every claim put forward in this thesis can and should be scrutinized, especially if applied beyond academic contexts. Researching expertise has made me more aware that no claim stands on its own: each rests on problem definitions and assumptions that dictate what to study and how. The questions addressed here were shaped by my own commitments and biases, and I hope that the elaboration of methods and reasoning allows the reader to assess the credibility of the results and the plausibility of their interpretation. If the work presented here suggests anything, it is that the entanglement between facts and interests is more evident now than ever: researchers and the public are retying the Knot.

References

Acerbi, A. (2019). Cognitive attraction and online misinformation. *Palgrave Communications*, *5*(1).

Achterberg, P., Koster, W., & van der Waal, J. (2017). A science confidence gap: Education, trust in scientific methods, and trust in scientific institutions in the United States, 2014. *Public Understanding of Science*, *26*(6), 704–720.

Ahler, D. J. (2014). Self-fulfilling misperceptions of public polarization. *The Journal of Politics*, *76*(3), 607–620.

Algan, Y., Cohen, D., Davoine, E., Foucault, M., & Stantcheva, S. (2021). Trust in scientists in times of pandemic: Panel evidence from 12 countries. *Proceedings of the National Academy of Sciences*, *118*(40), e2108576118.

Allen, J., Watts, D. J., & Rand, D. G. (2024). Quantifying the impact of misinformation and vaccine-skeptical content on facebook. *Science*, *384*(6699), eadk3451.

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.

Alper, S., Yelbuz, B. E., Akkurt, S. B., & Yilmaz, O. (2023). The positive association of education with the trust in science and scientists is weaker in highly corrupt countries. *Public Understanding of Science*, *33*(1), 2–19.

Altay, S., Berriche, M., Heuer, H., Farkas, J., & Rathje, S. (2023). A survey of expert views on misinformation: Definitions, determinants, solutions, and future of the field. *Harvard Kennedy School Misinformation Review*, *4*(4), 1–34.

Altay, S., De Araujo, E., & Mercier, H. (2022). "If this account is true, it is most enormously wonderful": Interestingness-if-true and the sharing of true and false news. *Digital Journalism*, *10*(3), 373–394.

Altenmüller, M. S., Wingen, T., & Schulte, A. (2024). Explaining polarized trust in scientists: A political stereotype-approach. *Science Communication*, *46*(1), 92–115.

Archer, L., Dawson, E., DeWitt, J., Seakins, A., & Wong, B. (2015). "Science capital": A conceptual, methodological, and empirical argument for extending bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching*, *52*(7), 922–948.

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

Baier, A. (1986). Trust and antitrust. *Ethics*, *96*(2), 231–260.

Bak, H.-J. (2001). Education and public attitudes toward science: Implications for the "deficit model" of education and support for science and technology. *Social Science Quarterly*, *82*(4), 779–795.

Bak-Coleman, J. B., Kennedy, I., Wack, M., Beers, A., Schafer, J. S., Spiro, E. S., Starbird, K., & West, J. D. (2022). Combining interventions to reduce the spread of viral misinformation. *Nature Human Behaviour*, *6*(10), 1372–1380.

Baldassarri, D., & Gelman, A. (2008). Partisans without constraint: Political polarization and trends in american public opinion. *American Journal of Sociology*, *114*(2), 408–446.

Baldi, P. (2002). A computational theory of surprise. *Information, coding and mathematics: Proceedings of workshop honoring Prof. Bob McEliece on his 60th birthday*, 1–25.

Bankert, A., Huddy, L., & Rosema, M. (2017). Measuring partisanship as a social identity in multi-party systems. *Political Behavior*, *39*(1), 103–132.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289.

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or surprise? *Frontiers in Psychology*, *4*, 907.

Barua, Z., Barua, S., Aktar, S., Kabir, N., & Li, M. (2020). Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. *Progress in Disaster Science*, *8*, 100119.

Bauer, P. C., & Clemm von Hohenberg, B. (2021). Believing and sharing information by fake sources: An experiment. *Political Communication*, *38*(6), 647–671.

Bauer, P. C., & Freitag, M. (2017). Measuring trust. In E. M. Uslaner (Ed.), *The Oxford Handbook of Social and Political Trust* (pp. 15–36). Oxford University Press.

Beck, U. (1996). World risk society as cosmopolitan society? ecological questions in a framework of manufactured uncertainties. *Theory, Culture & Society*, *13*(4), 1–32.

Beck, U., Giddens, A., & Lash, S. (1994). *Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order*. Stanford University Press.

Bellolio, C. (2019). The Quinean assumption. The case for science as public reason. *Social Epistemology*, *33*(3), 205–217.

Berger, J., & Milkman, K. (2010). Social transmission, emotion, and the virality of online content [Working paper, Market Science Institute]. https://marketing.wharton.upenn.edu/files/?whdmsaction=public:main.file&fileID=3461

Berinsky, A. J. (2017). Measuring public opinion with surveys. *Annual Review of Political Science*, *20*, 309–329.

Bernerth, J. B., & Aguinis, H. (2016). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, *69*(1), 229–283.

Berntzen, L. E., Kelsall, H., & Harteveld, E. (2024). Consequences of affective polarization: Avoidance, intolerance and support for violence in the united kingdom and norway. *European Journal of Political Research*, *63*(3), 927–949.

Bertsekas, D., & Tsitsiklis, J. N. (2008). *Introduction to Probability* (Vol. 1). Athena Scientific.

Biagioli, M. (1990a). Galileo the emblem maker. *Isis*, *81*(2), 230–258.

Biagioli, M. (1990b). Galileo's system of patronage. *History of Science*, *28*(1), 1–62.

Blancke, S. (2023). Nothing about collective irrationalities makes sense except in the light of cooperation. *Philosophical Psychology*, *36*(5), 990–1010.

Blank, J. M., & Shaw, D. (2015). Does partisanship shape attitudes toward science and public policy? the case for ideology and religion. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 18–35.

Bode, L., Vraga, E. K., & Tully, M. (2021). Correcting misperceptions about genetically modified food on social media: Examining the impact of experts, social media heuristics, and the gateway belief model. *Science Communication*, *43*(2), 225–251.

Bojarski, P. A., Suchecki, K., & Hołyst, J. A. (2022). Topic selectivity and adaptivity promote spreading of short messages. *Scientific Reports*, *12*(1), 15655.

Boulianne, S. (2019). US dominance of research on political communication: A meta-view. *Political Communication*, *36*(4), 660–665.

Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020). The MAD model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspectives on Psychological Science*, *15*(4), 978–1010.

Breen, R. (2010). Educational expansion and social mobility in the 20th century. *Social Forces*, *89*(2), 365–388.

Brossard, D., & Nisbet, M. C. (2007). Deference to scientific authority among a low information public: Understanding us opinion on agricultural biotechnology. *International Journal of Public Opinion Research*, *19*(1), 24–52.

Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in Human Neuroscience*, *6*, 147.

Bryanov, K., & Vziatysheva, V. (2021). Determinants of individuals' belief in fake news: A scoping review determinants of belief in fake news. *PLoS ONE*, *16*(6), e0253717.

Buchanan, T. (2020). Why do people spread false information online? the effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *PLoS ONE*, *15*(10), e0239666.

Campbell, T. H., & Kay, A. C. (2014). Solution aversion: On the relation between ideology and motivated disbelief. *Journal of Personality and Social Psychology*, *107*(5), 809.

Castells, M. (2000). Materials for an exploratory theory of the network society1. *The British Journal of Sociology*, *51*(1), 5–24.

Castells, M. (2004). Informationalism, networks, and the network society: A theoretical blueprint. In *The Network Society: A Cross-Cultural Perspective* (pp. 3–45). Edward Elgar Publishing, Inc.

Ceylan, G., Anderson, I. A., & Wood, W. (2023). Sharing of misinformation is habitual, not just lazy or biased. *Proceedings of the National Academy of Sciences*, *120*(4), e2216614120.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752.

Chen, J., Yan, Y., & Leach, J. (2022). Are emotion-expressing messages more shared on social media? A meta-analytic review. *Review of Communication Research*, *10*, 59–79.

Cheong, Y., & Haggard, S. (2023). Political polarization in Korea. *Democratization*, *30*(7), 1215–1239.

Chryssochoidis, G., Strada, A., & Krystallis, A. (2009). Public trust in institutions and information sources regarding risk management and communication: Towards integrating extant knowledge. *Journal of Risk Research*, *12*(2), 137–185.

Chung, E., Govindan, P., & Pechenkina, A. O. (2023). The effect of incentives on motivated numeracy amidst COVID-19. *Journal of Experimental Political Science*, *10*(3), 311–327.

Cikara, M., Van Bavel, J. J., Ingbretsen, Z. A., & Lau, T. (2017). Decoding "us" and "them": Neural representations of generalized group concepts. *Journal of Experimental Psychology: General*, *146*(5), 621–631.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Cohen, J. (1994). The earth is round (p<. 05). *American Psychologist*, *49*(12), 997.

Cologna, V., Mede, N. G., Berger, S., Besley, J., Brick, C., Joubert, M., & Metag, J. (2025). Trust in scientists and their role in society across 68 countries. *Nature Human Behaviour*, *9*, 713–730.

Cologna, V., & Siegrist, M. (2020). The role of trust for climate change mitigation and adaptation behaviour: A meta-analysis. *Journal of Environmental Psychology*, *69*, 101428.

Concato, J., & Hartigan, J. A. (2016). P values: From suggestion to superstition. *Journal of Investigative Medicine*, *64*(7), 1166–1171.

Cover, T. M., & Thomas, J. A. (1991). Entropy, relative entropy, and mutual information. In *Elements of Information Theory*. John Wiley; Sons, Inc.

Cuenca, M. C. S., & Pérez, N. E. (2025). The role of social media as a source of information and perception of insecurity. *European Journal on Criminal Policy and Research*, *31*, 377–398.

Dahlstrom, M. F. (2021). The narrative truth about scientific misinformation. *Proceedings of the National Academy of Sciences*, *118*(15), e1914085117.

Dassonneville, R., Fournier, P., & Somer-Topcu, Z. (2023). Partisan attachments in a multidimensional space. *West European Politics*, *46*(4), 678–704.

Davis, B., Goodliffe, J., & Hawkins, K. (2025). The two-way effects of populism on affective polarization. *Comparative Political Studies*, *58*(1), 122–154.

Davis, J. M., & Wilfahrt, M. (2025). Understanding the sensitivity of party identification questions in polarized african contexts. *Political Behavior*, 1–26.

Dear, P., & Jasanoff, S. (2010). Dismantling boundaries in science and technology studies. *Isis*, *101*(4), 759–774.

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

DellaPosta, D. (2020). Pluralistic collapse: The "oil spill" model of mass opinion polarization. *American Sociological Review*, *85*(3), 507–536.

den Broeder, L., Devilee, J., van Oers, H., Schuit, A. J., & Wagemakers, A. (2018). Citizen science for public health. *Health Promotion International*, *33*(3), 505–514.

Dooremalen, H., de Regt, H., & Schouten, M. K. D. (2015). *Exploring Humans: Philosophy of Science for the Social Sciences-A Historical Introduction*. Boom.

Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2022). (Mis) estimating affective polarization. *The Journal of Politics*, *84*(2), 1106–1117.

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, *9*(2), 111–119.

Drummond, C., & Fischhoff, B. (2017). Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. *Proceedings of the National Academy of Sciences*, *114*(36), 9587–9592.

Dunlap, R. E. (2014). Clarifying anti-reflexivity: Conservative opposition to impact science and scientific evidence. *Environmental Research Letters*, *9*(2), 021001.

Ecker, U. K. H., Tay, L. Q., Roozenbeek, J., van der Linden, S., Cook, J., Oreskes, N., & Lewandowsky, S. (2025). Why misinformation must not be ignored. *American Psychologist*, *80*(6), 867–878.

Edelmann, A., Wolff, T., Montagne, D., & Bail, C. A. (2020). Computational social science and sociology. *Annual Review of Sociology*, *46*(1), 61–81.

Egan, O. (1986). The concept of belief in cognitive theory. In *Annals of Theoretical Psychology* (pp. 315–350). Springer.

Ellison, N., & Hardey, M. (2014). Social media and local government: Citizenship, consumption and democracy. *Local Government Studies*, *40*(1), 21–40.

Epstein, Z., Sirlin, N., Arechar, A., Pennycook, G., & Rand, D. (2023). The social media context interferes with truth discernment. *Science Advances*, *9*(9), eabo6169.

Eyal, G. (2019). *The Crisis of Expertise*. Polity Press.

Falkenberg, M., Galeazzi, A., Torricelli, M., Di Marco, N., Larosa, F., Sas, M., Mekacher, A., Pearce, W., Zollo, F., Quattrociocchi, W., & Baronchelli, A. (2022). Growing polarization around climate change on social media. *Nature Climate Change*, *12*(12), 1114–1121.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96–104.

Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., et al. (2020). Political sectarianism in america. *Science*, *370*(6516), 533–536.

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*(1), 563–588.

Fischer, H., Huff, M., & Said, N. (2022). Polarized climate change beliefs: No evidence for science literacy driving motivated reasoning in a us national study. *American Psychologist*, *77*(7), 822.

Forman, P. (2007). The primacy of science in modernity, of technology in postmodernity, and of ideology in the history of technology. *History and Technology*, *23*(1-2), 1–152.

Foucault, M. (2008). Power/knowledge. In S. Seidman & J. C. Alexander (Eds.), *The New Social Theory Reader* (pp. 73–79). Routledge.

Friggeri, A., Adamic, L., Eckles, D., & Cheng, J. (2014). Rumor cascades. *Proceedings of the Onternational AAAI Conference on Web and Social Media*, *8*(1), 101–110.

Frye, T. (2002). The perils of polarization: Economic performance in the postcommunist world. *World Politics*, *54*(3), 308–337.

Fuchs, D., & Klingemann, H. D. (1990). The left-right schema. In M. Jennings & J. van Deth (Eds.), *Continuities in political action: A longitudinal study of political orientations in three western democracies* (pp. 203–234). Walter de Gruyter.

Funk, C., Hefferon, M., Kennedy, B., & Johnson, C. (2019). Trust and mistrust in Americans' views of scientific experts [Accessed: 2025-10-16]. https://www.pewresearch.org/science/2019/08/02/trust-and-mistrust-in-americans-views-of-scientific-experts/

Funk, C., Kennedy, B., & Johnson, C. (2020). Trust in medical scientists has grown in U.S., but mainly among Democrats [Accessed: 2025-10-11]. https://www.pewresearch.org/science/2020/05/21/trust-in-medical-scientists-has-grown-in-u-s-but-mainly-among-democrats/

Furnas, A. C., LaPira, T. M., & Wang, D. (2025). Partisan disparities in the use of science in policy. *Science*, *388*(6745), 362–367.

Gaines, B. J., Kuklinski, J. H., & Quirk, P. J. (2007). The logic of the survey experiment reexamined. *Political Analysis*, *15*(1), 1–20.

Garrett, R. K., & Bond, R. M. (2021). Conservatives' susceptibility to political misperceptions. *Science Advances*, *7*(23), eabf1234.

Gauchat, G. W. (2010). The cultural authority of science: Public trust and acceptance of organized science. *Public Understanding of Science*, *20*(6), 751–770.

Gauchat, G. W. (2012). Politicization of science in the public sphere: A study of public trust in the United States, 1974 to 2010. *American Sociological Review*, *77*(2), 167–187.

Gauchat, G. W. (2023). The legitimacy of science. *Annual Review of Sociology*, *49*(1), 263–279.

Gemenis, K., Katsanidou, A., & Vasilopoulou, S. (2012). The politics of anti-environmentalism: Positional issue framing by the european radical right. *62nd Annual Political Studies Association Conference 2012*.

Gidron, N., Adams, J., & Horne, W. (2023). Who dislikes whom? Affective polarization between pairs of parties in western democracies. *British Journal of Political Science*, *53*(3), 997–1015.

Gieryn, T. F. (1999). *Cultural Boundaries of Science: Credibility on the Line*. University of Chicago Press.

Gligorić, K., Anderson, A., & West, R. (2019). Causal effects of brevity on style and success in social media. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–23.

Gligorić, V., van Kleef, G. A., & Rutjens, B. T. (2025). Political ideology and trust in scientists in the USA. *Nature Human Behaviour*, 1–12.

Goebel, J. T., Susmann, M. W., Parthasarathy, S., El Gamal, H., Garrett, R. K., & Wegener, D. T. (2024). Belief-consistent information is most shared despite being the least surprising. *Scientific Reports*, *14*(1), 6109.

Goldman, G., & Chenoweth, E. (2025). Scientists' role in defending democracy. *Science*, *389*(6761), 667–667.

Goren, P. (2005). Party identification and core political values. *American Journal of Political Science*, *49*(4), 881–896.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on facebook. *Science Advances*, *5*(1), eaau4586.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf

Habinek, J. (2010). *State-building and the origins of disciplinary specialization in nineteenth century germany* (Working Paper) (Previously Affiliated). Center for Culture, Organizations and Politics, UC Berkeley. https://escholarship.org/uc/item/7jv0b9s3

Harambam, J., & Aupers, S. (2015). Contesting epistemic authority: Conspiracy theories on the boundaries of science. *Public Understanding of Science*, *24*(4), 466–480.

Havey, N. F. (2020). Partisan public health: How does political ideology influence support for COVID-19 related misinformation? *Journal of Computational Social Science*, *3*(2), 319–342.

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Trust in science and the science of trust. In B. Blöbaum (Ed.), *Trust and Communication in a Digitized World* (pp. 143–159). Springer.

Hofstadter, R. (2012). *The Paranoid Style in American Politics*. Vintage.

Hooghe, M., & Kern, A. (2015). Party membership and closeness and the development of trust in political institutions: An analysis of the European Social Survey, 2002–2010. *Party Politics*, *21*(6), 944–956.

Hornsey, M. J. (2020). Why facts are not enough: Understanding and managing the motivated rejection of science. *Current Directions in Psychological Science*, *29*(6), 583–591.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Hu, Y., & Shyam Sundar, S. (2010). Effects of online health sources on credibility and behavioral intentions. *Communication Research*, *37*(1), 105–132.

Huang, C., & Kuo, T.-C. (2022). Actual and perceived polarization on independence-unification views in taiwan. *Asian Journal of Communication*, *32*(2), 75–92.

Huber, R. A. (2020). The role of populist attitudes in explaining climate change skepticism and support for environmental protection. *Environmental Politics*, *29*(6), 959–982.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146.

Iyengar, S., & Massey, D. S. (2019). Scientific communication in a post-truth society. *Proceedings of the National Academy of Sciences*, *116*(16), 7656–7661.

Jasanoff, S. (2004a). The idiom of co-production. In S. Jasanoff (Ed.), *States of Knowledge*. Routledge.

Jasanoff, S. (2004b). Ordering knowledge, ordering society. In S. Jasanoff (Ed.), *States of Knowledge*. Routledge.

Jasanoff, S. (2005). *Designs on Nature: Science and Democracy in Europe and the United States*. Princeton University Press.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, *80*(4), 557–571.

Jones, C. M., Diethei, D., Schöning, J., Shrestha, R., Jahnel, T., & Schüz, B. (2023). Impact of social reference cues on misinformation sharing on social media: Series of experimental studies. *Journal of Medical Internet Research*, *25*, e45583.

Jones, D. S. (1979). *Elementary Information Theory*. Oxford University Press.

Joslyn, M., & Sylvester, S. (2019). The determinants and consequences of accurate beliefs about childhood vaccinations. *American Politics Research*, *47*(3), 628–649.

Jost, J. T., Glaser, J., Sulloway, F. J., & Kruglanski, A. W. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, 339–375.

Jost, J. T., Langer, M., Badaan, V., Azevedo, F., Etchezahar, E., Ungaretti, J., & Hennes, E. P. (2017). Ideology and the limits of self-interest: System justification motivation and conservative advantages in mass politics. *Translational Issues in Psychological Science*, *3*(3), e1–e26.

Kahan, D. M. (2012). Cultural cognition as a conception of the cultural theory of risk. In S. Roeser, R. Hillerbrand, & P. Sandin (Eds.), *Handbook of Risk Theory* (pp. 725–759). Springer.

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, *8*(4), 407–424.

Kahan, D. M., Jenkins-Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of Risk Research*, *14*(2), 147–174.

Kahan, D. M., Jenkins-Smith, H., Tarantola, T., Silva, C. L., & Braman, D. (2015). Geoengineering and climate change polarization: Testing a two-channel model of science communication. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 192–222.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732–735.

Kata, A. (2010). A postmodern pandora's box: Anti-vaccination misinformation on the internet. *Vaccine*, *28*(7), 1709–1716.

Kennedy, B., Tyson, A., & Funk, C. (2022, February). *Americans' trust in scientists, other groups declines*. Retrieved April 20, 2025, from https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines/

Khan, G. F., Swar, B., & Lee, S. K. (2014). Social media risks and benefits: A public sector perspective. *Social Science Computer Review*, *32*(5), 606–627.

Kibet, A., & Ward, S. (2018). Socially networked heterogeneity: The influence of whatsapp as a social networking site on polarisation in kenya. *African Journalism Studies*, *39*(4), 42–66.

Kim, H. S. (2015). Attracting views and going viral: How message features and news-sharing channels affect health news diffusion. *Journal of Communication*, *65*(3), 512–534.

Klein, B. D. (2001). User perceptions of data quality: Internet and traditional text sources. *Journal of Computer Information Systems*, *41*(4), 9–15.

Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–progressive beliefs, and communion. *Journal of Personality and Social Psychology*, *110*(5), 675.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498.

Kunovich, R. M. (2024). Work and the public understanding of science. *Public Understanding of Science*, *33*(3), 353–369.

Latour, B. (1993). *We Have Never Been Modern*. Harvard University Press.

Lavazza, A., & Farina, M. (2022). Experts, naturalism, and democracy. *Journal for the Theory of Social Behaviour*, *52*(2), 279–297.

Leach, S., Formanowicz, M., Nikadon, J., & Cichocka, A. (2026). Moral outrage predicts the virality of petitions for change on social media, but not the number of signatures they receive. *Social Psychological and Personality Science*, *17*(2), 194–203.

Lee, J. J. (2021). Party polarization and trust in science: What about democrats? *Socius*, *7*, 1–12.

Lee, S., Rojas, H., & Yamamoto, M. (2022). Social media, messaging apps, and affective polarization in the United States and Japan. *Mass Communication and Society*, *25*(5), 673–697.

Leeper, T., & Slothuus, R. (2014). Political parties, motivated reasoning, and public opinion formation. *Advances in Political Psychology*, *35*(1), 129–156.

Levendusky, M. S. (2009). *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. University of Chicago Press.

Levendusky, M. S., & Malhotra, N. (2016). (Mis) perceptions of partisan polarization in the American public. *Public Opinion Quarterly*, *80*(S1), 378–391.

Lewandowsky, S., & Oberauer, K. (2016). Motivated rejection of science. *Current Directions in Psychological Science*, *25*(4), 217–222.

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA faked the moon landing—therefore,(climate) science is a hoax: An anatomy of the motivated rejection of science. *Psychological Science*, *24*(5), 622–633.

Lewis, J. R. (2010). The science canopy: Religion, legitimacy, and the charisma of science. *Temenos. Nordic Journal of Comparative Religion*, *46*(1), 7–29.

Li, J., & Chang, X. (2023). Combating misinformation by sharing the truth: A study on the spread of fact-checks on social media. *Information Systems Frontiers*, *25*(4), 1479–1493.

Li, N., & Qian, Y. (2022). Polarization of public trust in scientists between 1978 and 2018: Insights from a cross-decade comparison using interpretable machine learning. *Politics and the Life Sciences*, *41*(1), 45–54.

Lin, Y., & Kant, S. (2021). Using social media for citizen participation: Contexts, empowerment, and inclusion. *Sustainability*, *13*(12), 6635.

Lupia, A., Allison, D. B., Jamieson, K. H., Heimberg, J., Skipper, M., & Wolf, S. M. (2024). Trends in US public confidence in science and opportunities for progress. *Proceedings of the National Academy of Sciences*, *121*(11), e2319488121.

Mann, M., & Schleifer, C. (2020). Love the science, hate the scientists: Conservative identity protects belief in science and undermines trust in scientists. *Social Forces*, *99*(1), 305–332.

Mason, L. (2015). "I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science*, *59*(1), 128–145.

McConnell, C., Margalit, Y., Malhotra, N., & Levendusky, M. (2018). The economic consequences of partisanship in a polarized era. *American Journal of Political Science*, *62*(1), 5–18.

McCoy, J., Rahman, T., & Somer, M. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic polities. *American Behavioral Scientist*, *62*(1), 16–42.

McCright, A. M., Dentzman, K., Charters, M., & Dietz, T. (2013). The influence of political ideology on trust in science. *Environmental Research Letters*, *8*(4), 044029.

McCright, A. M., & Dunlap, R. E. (2010). Anti-reflexivity. *Theory, Culture & Society*, *27*(2-3), 100–133.

McCright, A. M., & Dunlap, R. E. (2011). The politicization of climate change and polarization in the American public's views of global warming, 2001–2010. *The Sociological Quarterly*, *52*(2), 155–194.

McLoughlin, K. L., Brady, W. J., Goolsbee, A., Kaiser, B., Klonick, K., & Crockett, M. (2024). Misinformation exploits outrage to spread online. *Science*, *386*(6725), 991–996.

McNutt, M. (2024). Science is neither red nor blue. *Science*, *386*(6723), 707–707.

Mede, N. G., Howell, E. L., Schäfer, M. S., Metag, J., Beets, B., & Brossard, D. (2025). Measuring science literacy in a digital world: Development and validation of a multi-dimensional survey scale. *Science Communication*, 10755470251317379.

Media Bias/Fact Check. (2021). Left vs. right bias: How we rate the bias of media sources [Accessed: 2025-12-05]. https://mediabiasfactcheck.com/left-vs-right-bias-how-we-rate-the-bias-of-media-sources/

Media Bias/Fact Check. (2024). Media Bias/Fact Check methodology [Accessed: 2025-12-05]. https://mediabiasfactcheck.com/methodology/

Mejlgaard, N., & Stares, S. (2013). Performed and preferred participation in science and technology across europe: Exploring an alternative idea of "democratic deficit". *Public Understanding of Science*, *22*(6), 660–673.

Meppelink, C. S., Smit, E. G., Fransen, M. L., & Diviani, N. (2019). "I was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of Health Communication*, *24*(2), 129–140.

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

Merkley, E. (2020). Anti-intellectualism, populism, and motivated resistance to expert consensus. *Public Opinion Quarterly*, *84*(1), 24–48.

Messing, S., DeGregorio, C., Hillenbrand, B., King, G., Mahanti, S., Mukerjee, Z., Nayak, C., Persily, N., State, B., & Wilkins, A. (2020). Facebook privacy-protected full URLs data set. https://doi.org/10.7910/DVN/TDOAPG

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, *59*, 210–220.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, *60*(3), 413–439.

Millstone, E., & van Zwanenberg, P. (2000). A crisis of trust: For science, scientists or for institutions? *Nature Medicine*, *6*(12), 1307–1308.

Mooney, C. (2007). *The Republican War on Science*. Hachette.

Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of Medical Internet research*, *15*(4), e1933.

Morrisson, C., & Murtin, F. (2009). The century of education. *Journal of Human Capital*, *3*(1), 1–42.

Mosleh, M., Pennycook, G., & Rand, D. G. (2020). Self-reported willingness to share political news articles in online surveys correlates with actual sharing on twitter. *PLoS ONE*, *15*(2), e0228882.

Mosleh, M., & Rand, D. G. (2022). Measuring exposure to misinformation from political elites on twitter. *Nature Communications*, *13*(1), 7144.

Motta, M. (2018). The dynamics and political implications of anti-intellectualism in the United States. *American Politics Research*, *46*(3), 465–498.

Munger, K. (2017). Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior*, *39*(3), 629–649.

Newman, T. P., Nisbet, E. C., & Nisbet, M. C. (2018). Climate change, cultural cognition, and media effects: Worldviews drive news selectivity, biased processing, and polarized attitudes. *Public Understanding of Science*, *27*(8), 985–1002.

Nichols, T. M. (2017). *The Death of Expertise: The Campaign Against Established Knowledge and Why It Matters*. Oxford University Press.

Nikolov, D., Oliveira, D. F., Flammini, A., & Menczer, F. (2015). Measuring online social bubbles. *PeerJ Computer Science*, *1*, e38.

Nisbet, E. C., Cooper, K. E., & Garrett, R. K. (2015). The partisan brain: How dissonant science messages lead conservatives and liberals to (dis) trust science. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 36–66.

Noy, S., & O'Brien, T. L. (2019). Science for good? the effects of education and national context on perceptions of science. *Public Understanding of Science*, *28*(8), 897–916.

Nugent, D. (2010). Knowledge and empire: The social sciences and United States imperial expansion. *Identities: Global Studies in Culture and Power*, *17*(1), 2–44.

Oreskes, N., & Conway, E. M. (2022). From anti-government to anti-science: Why conservatives have turned against science. *Daedalus*, *151*(4), 98–123.

Oreskes, N., & Conway, E. M. (2011). *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury Publishing USA.

Paparo, A., & Sio, L. D. (2017). PTV gap as a new measure of partisanship: A panel-data, multi-measure validation showing surprising partisanship stability. *Contemporary Italian Politics*, *9*(1), 60–83.

Pasek, J. (2018). It's not my consensus: Motivated reasoning and the sources of scientific illiteracy. *Public Understanding of Science*, *27*(7), 787–806.

Pelizzo, R., & Babones, S. (2007). The political economy of polarized pluralism. *Party Politics*, *13*(1), 53–67.

Pennycook, G., Bago, B., & McPhetres, J. (2019). Science beliefs, political ideology, and cognitive sophistication. *Journal of Experimental Psychology: General*, *152*(1), 80–97.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865.

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590–595.

Pennycook, G., McPhetres, J., Bago, B., & Rand, D. G. (2022). Beliefs about COVID-19 in Canada, the United Kingdom, and the United States: A novel test of political polarization and motivated reasoning. *Personality and Social Psychology Bulletin*, *48*(5), 750–765.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50.

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, *25*(5), 388–402.

Persson, E., Andersson, D., Koppel, L., Västfjäll, D., & Tinghög, G. (2021). A preregistered replication of motivated numeracy. *Cognition*, *214*, 104768.

Petersen, M. B., Osmundsen, M., & Arceneaux, K. (2023). The "need for chaos" and motivations to share hostile political rumors. *American Political Science Review*, *117*(4), 1486–1505.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (pp. 123–205, Vol. 19). Academic Press.

Pfänder, J., Kerzreho, L., & Mercier, H. (2025). Quasi-universal acceptance of basic science in the United States. *Public Understanding of Science*, 09636625251364407.

Phillips, A. (2010). Old sources: New bottles. In *New Media, Old News: Journalism and Democracy in the Digital Age* (pp. 87–101). SAGE Publications.

Plique, G., Breteau, P., Farjas, J., Théro, H., Descamps, J., et al. (2019). Minet, a web-mining cli tool & library for Python [Accessed: 2025-10-17]. https://github.com/medialab/minet

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, *34*(2), 243–281.

Puryear, C., Vandello, J. A., & Gray, K. (2024). Moral panics on social media are fueled by signals of virality. *Journal of Personality and Social Psychology*, *127*(1), 84.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Rao, A., Morstatter, F., & Lerman, K. (2022). Partisan asymmetries in exposure to misinformation. *Scientific Reports*, *12*(1), 15671.

Rathje, S., Roozenbeek, J., Van Bavel, J. J., & van der Linden, S. (2023). Accuracy and social motivations shape judgements of (mis) information. *Nature Human Behaviour*, *7*(6), 892–903.

Reiljan, A. (2020). 'Fear and loathing across party lines' (also) in Europe: Affective polarisation in european party systems. *European Journal of Political Research*, *59*(2), 376–396.

Rekker, R. (2021). The nature and origins of political polarization over science. *Public Understanding of Science*, *30*(4), 352–368.

Rekker, R. (2025). A four-level model of political polarization over science: Evidence from 10 European countries. *Public Understanding of Science*, *34*(4), 424–445.

Rozek, L. S., Jones, P., Menon, A., Hicken, A., Apsley, S., & King, E. J. (2021). Understanding vaccine hesitancy in the context of COVID-19: The role of trust and confidence in a seventeen-country survey. *International Journal of Public Health*, *66*, 636255.

Saltelli, A., & Funtowicz, S. (2017). What is science's crisis really about? *Futures*, *91*, 5–11.

Sarsfield, R., Moncagatta, P., & Roberts, K. M. (2024). Introduction: The new polarization in latin america. *Latin American Politics and Society*, *66*(2), 1–23.

Savadori, L., Dickson, M. M., Micciolo, R., & Espa, G. (2022). The polarizing impact of numeracy, economic literacy, and science literacy on the perception of immigration. *PLoS ONE*, *17*(10), e0274680.

Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, *116*(16), 7662–7669.

Schiebinger, L. (2004). Feminist history of colonial science. *Hypatia*, *19*(1), 233–254.

Schnaiberg, A. (1980). *The Environment: From Surplus to Scarcity*. Oxford University Press.

Schofer, E., & Meyer, J. W. (2005). The worldwide expansion of higher education in the twentieth century. *American Sociological Review*, *70*(6), 898–920.

Schrøder, T. B. (2023). Don't tell me what I don't want to hear! Politicization and ideological conflict explain why citizens have lower trust in climate scientists and economists than in other natural scientists. *Political Psychology*, *44*(5), 961–981.

Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, *78*(4), 655–669.

Science Feedback. (2023). Consensus Credibility Scores: A comprehensive dataset of web domains' credibility [Accessed: 2025-12-05]. https://science.feedback.org/consensus-credibility-scores-comprehensive-dataset-web-domains-credibility/

Science Feedback. (2025). Process – how science feedback works [Accessed: 2025-12-05]. https://science.feedback.org/process/

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Shapin, S. (1994). *A Social History of Truth: Civility and Science in Seventeenth-Century England*. University of Chicago Press.

Shapin, S. (2008). Science and the modern world. In E. J. Hackett, O. Amsterdamska, M. Lynch, & J. Wajcman (Eds.), *The Handbook of Science and Technology Studies* (pp. 433–448). MIT Press. http://nrs.harvard.edu/urn-3:HUL.InstRepos:3425896

Shi, D., Maydeu-Olivares, A., & DiStefano, C. (2018). The relationship between the standardized root mean square residual and model misspecification in factor analysis models. *Multivariate Behavioral Research*, *53*(5), 676–694.

Shi, F., Shi, Y., Dokshin, F. A., Evans, J. A., & Macy, M. W. (2017). Millions of online book co-purchases reveal partisan differences in the consumption of science. *Nature Human Behaviour*, *1*(4), 0079.

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.

Smith, D. S. (2022). Social scientization and the schooling state in uk parliamentary discourse, 1803–1909. *Social Science History*, *46*(1), 223–254.

Sorell, T. (2013). *Scientism: Philosophy and the Infatuation with Science*. Routledge.

Sterling, J., Jost, J. T., & Hardin, C. D. (2019). Liberal and conservative representations of the good society: A (social) structural topic modeling approach. *Sage Open*, *9*(2), 1–13.

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, *29*(4), 217–248.

Strömbäck, C., Andersson, D., Västfjäll, D., & Tinghög, G. (2024). Motivated reasoning, fast and slow. *Behavioural Public Policy*, *8*(3), 617–632.

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

Sturgis, P., Brunton-Smith, I., & Jackson, J. (2021). Trust in science, social consensus and vaccine confidence. *Nature Human Behaviour*, *5*(11), 1528–1534.

Suarez-Lledo, V., & Alvarez-Galvez, J. (2021). Prevalence of health misinformation on social media: Systematic review. *J Med Internet Res*, *23*(1), e17187.

Suhay, E., & Druckman, J. N. (2015). The politics of science: Political values and the production, communication, and reception of scientific knowledge. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 6–15.

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital Media, Youth, and Credibility* (pp. 73–100). The MIT Press.

Taddicken, M., & Krämer, N. (2021). Public online engagement with science information: On the road to a theoretical framework and a future research agenda. *Journal of Science Communication*, *20*(3), A05.

Tellis, G. J., MacInnis, D. J., Tirunillai, S., & Zhang, Y. (2019). What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence. *Journal of Marketing*, *83*(4), 1–20.

Tribus, M. (1961). Information theory as the basis for thermostatics and thermodynamics. *Journal of Applied Mechanics*, *28*(1), 1–8.

Tromp, P., & Achterberg, P. (2021). Cultural worldviews and lay interpretations of research findings: The role of the scientific consensus. In D. Houtman, S. Aupers, & R. Laermans (Eds.), *Science Under Siege: Contesting the Secular Religion of Scientism* (pp. 133–152). Springer International Publishing.

Tucker, J. A., Theocharis, Y., Roberts, M. E., & Barberá, P. (2017). From liberation to turmoil: Social media and democracy. *Journal of Democracy*, *28*(4), 46–59.

Tufekci, Z. (2020). Keep the parks open [Accessed: 15-10-2023]. https://www.theatlantic.com/health/archive/2020/04/closing-parks-ineffective-pandemic-theater/609580/

Turnley, V. (2025). An assault on science anywhere is an assault on science everywhere. *Nature*, *639*, 7–8.

Van Aelst, P., Toth, F., Castro, L., Štětka, V., De Vreese, C., Aalberg, T., Cardenal, A. S., Corbu, N., Esser, F., Hopmann, D. N., et al. (2025). Does a crisis change news habits? A comparative study of the effects of COVID-19 on news media use

in 17 european countries. In T. Quandt & K. Wahl-Jorgensen (Eds.), *Covering Covid-19* (pp. 10–40). Routledge.

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. *Trends in Cognitive Sciences*, *22*(3), 213–224.

Van Der Heide, B., & Lim, Y.-s. (2016). On the conditional cueing of credibility heuristics: The case of online influence. *Communication Research*, *43*(5), 672–693.

van der Meer, T. G., & Hameleers, M. (2025). Science and the crisis of trust. *Current Opinion in Psychology*, 102202.

Vellani, V., Zheng, S., Ercelik, D., & Sharot, T. (2023). The illusory truth effect leads to the spread of misinformation. *Cognition*, *236*, 105421.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151.

Vraga, E. K., & Bode, L. (2017). Using expert sources to correct health misinformation in social media. *Science Communication*, *39*(5), 621–645.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804.

Wagner, M., Tarlov, J., & Vivyan, N. (2014). Partisan bias in opinion formation on episodes of political controversy: Evidence from Great Britain. *Political Studies*, *62*(1), 136–158.

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, *69*, 102199.

Walter, N., & Tukachinsky, R. (2020). A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it? *Communication Research*, *47*(2), 155–177.

Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic literature review on the spread of health-related misinformation on social media. *Social Science & Medicine*, *240*, 112552.

Washburn, A. N., & Skitka, L. J. (2018). Science denial across the political divide: Liberals and conservatives are similarly motivated to deny attitude-inconsistent science. *Social Psychological and Personality Science*, *9*(8), 972–980.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on statistical significance and p-values. *The American Statistician*, *70*(2), 129–133.

Webster, S. W., & Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the US electorate. *American Politics Research*, *45*(4), 621–647.

West, J. D., & Bergstrom, C. T. (2021). Misinformation in and about science. *Proceedings of the National Academy of Sciences*, *118*(15), e1912444117.

Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on Psychological Science*, *10*(2), 145–158.

Westwood, S. J., Iyengar, S., Walgrave, S., Leonisio, R., Miller, L., & Strijbis, O. (2018). The tie that divides: Cross-national evidence of the primacy of partyism. *European Journal of Political Research*, *57*(2), 333–354.

Williams, D. (2022). Signalling, commitment, and strategic absurdities. *Mind & Language*, *37*(5), 1011–1029.

Williams, D. (2023). The case for partisan motivated reasoning. *Synthese*, *202*(3), 89.

Wilson, A. E., Parker, V. A., & Feinberg, M. (2020). Polarization in the contemporary political and media landscape. *Current Opinion in Behavioral Sciences*, *34*, 223–228.

Wilson, D., & Sperber, D. (1995). *Relevance: Communication and Cognition*. Blackwell Publishers.

Wilson, D., & Sperber, D. (1998). Pragmatics and time. In R. Carston & S. Uchida (Eds.), *Relevance Theory: Applications and Implications* (pp. 169–186). Cambridge University Press.

Wynne, B. (1991). Knowledges in context. *Science, Technology, & Human Values*, *16*(1), 111–121.

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods*, *51*(1), 409–428.

Ximénez, C., Maydeu-Olivares, A., Shi, D., & Revuelta, J. (2022). Assessing cutoff values of sem fit indices: Advantages of the unbiased srmr index and its cutoff criterion based on communality. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(3), 368–380.

Yang, K.-C., Varol, O., Hui, P.-M., & Menczer, F. (2020). Scalable and generalizable social bot detection through data selection. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(01), 1096–1103.

Yang, K.-C., Varol, O., Hui, P.-M., & Menczer, F. (2023). Botometer x [Accessed: 2026-03-05]. https://botometer.osome.iu.edu/

Yearley, S. (2000). Making systematic sense of public discontents with expert knowledge: Two analytical approaches and a case study. *Public Understanding of Science*, *9*(2), 105.

Young, D. J., Ackland, J., Kapounek, A., Madsen, J. K., Greening, L. J., & de-Wit, L. (2026). A new measure of issue polarization using k-means clustering: US trends 1988–2024 and predictors of polarization across the world. *Royal Society Open Science*, *13*(2), 251428.

Zhuravskaya, E., Petrova, M., & Enikolopov, R. (2020). Political effects of the internet and social media. *Annual Review of Economics*, *12*(1), 415–438.

List of Figures and Tables

Figures

1.1 Differences in trust in experts between conservatives and liberals 45

1.2 Perceived accuracy of claims by experimental condition and political ideology, shown for all and specific claims 47

1.3 Perceived accuracy of the taxation and climate change claims by experimental condition and political ideology 50

2.1 Direct effects model (M1) 68

2.2 Full model (M5) 69

2.3 The political polarization of science 70

3.1 Perceived surprise (a) and interestingness (b) by experimental condition . . 93

3.2 Sharing intention for the single-report (a) and double-report (b) questionnaires by experimental condition 95

4.1 Data pipeline summary 104

4.2 Attributing claims to scientific authority increases both willingness to share and perceived accuracy 111

4.3 Perceived accuracy statistically mediates the effect of scientific authority attribution on sharing intentions 112

4.4 Attributing claims to trusted entities enhances both sharing intention and perceived accuracy 114

Tables

1.1 Claims and their variations across experimental conditions 42

1.2 Perceived accuracy of claims depending on the source: specific (impact or production), general, or control (no source) 49

3.1 Statements shown to participants before the experimental task 86

3.2 All possible messages concerning nutrition shown to participants (single-report version) 87

3.3 Distribution of responses across experimental conditions 89

3.4 Correlations between outcomes 91

3.5 Models for the tests of H1, H2a, and H2b 92

3.6 Models for the tests of H3–H5 94

Appendices

Chapter 1

Appendix 1.1: Power calculation and motivation

Since this is a novel approach to evaluating expert claims, no data are available from which we can derive a good estimate of the effect size in the population. Nevertheless, two studies have researched the effects of political ideology on trust in scientists and trust in expert claims. McCright et al. (2013) find significant effects of political ideology on trust in impact and production scientists. The more to the right a person is on the political scale, the less they trust impact science and the more they trust production science. Additionally, Rekker (2025) reports standardized effects which reveal that the more to the right of the political spectrum a person is, the less they trust environmental science ($B = -0.10$) and sociology ($B = -0.11$), and the more they trust agricultural science ($B = 0.08$) and economics ($B = 0.09$).

A conservative approach was taken by assuming that there is a small effect size in the population. This is pertinent considering that I am trying to uncover a moderating effect. The power calculation is based on the test of the four hypotheses above. According to the *pwr* package in R, detecting a small difference (Cohen's $d = 0.2$) between two samples with a desired power of 0.9, α of 0.025 (correcting for two tests), and a one-sided test yields around 527 participants per comparison group. This means that 1054 participants are needed in total.

Appendix 1.2: Exploratory five-factor model

Standardized factor loadings and unstandardized regression coefficients of the exploratory five-factor model. $N = 1054$; $\chi^2(175) = 1381.24$, $p < 0.001$; CFI = 0.98, TLI = 0.97, RMSEA = 0.08, SRMR = 0.04; *** $p < 0.001$.

	Trust in Environ. Scientists	**Trust in Sociologists**	**Trust in Agricultural Scientists**	**Trust in Economists**	**Trust in Scientists**
Create Unbiased Knowledge	0.95***	0.93***	0.92***	0.88***	0.92***
Create Useful Knowledge	0.94***	0.92***	0.91***	0.90***	0.88***
Advise Government Officials	0.93***	0.91***	0.89***	0.86***	0.92***
Inform the Public	0.95***	0.95***	0.91***	0.92***	0.92***
Conservative (vs. Liberal)	−1.39***	−0.95***	−0.79***	−0.43***	−1.39***
Factors Cronbach's α	0.95	0.94	0.92	0.91	0.92

Appendix 1.3: Hypothesized three-factor model

Standardized factor loadings and unstandardized regression coefficients. $N = 1054$; $\chi^2(184) = 5058.45$, $p < 0.001$; CFI = 0.90, TLI = 0.90, RMSEA = 0.16, SRMR = 0.12; *** $p < 0.001$.

	Trust in Impact Experts		**Trust in Production Experts**		**Trust in Scientists**
	Environ. Scientists	**Sociologists**	**Agricultural Scientists**	**Economists**	**Scientists**
Create Unbiased Knowledge	0.93***	0.90***	0.90***	0.83***	0.92***
Create Useful Knowledge	0.92***	0.89***	0.89***	0.85***	0.88***
Advise Government Officials	0.91***	0.89***	0.86***	0.80***	0.92***
Inform the Public	0.93***	0.92***	0.89***	0.86***	0.92***
Conservative (vs. Liberal)	−1.19***		−0.64***		−1.23***
Factors Cronbach's α	0.95		0.91		0.92

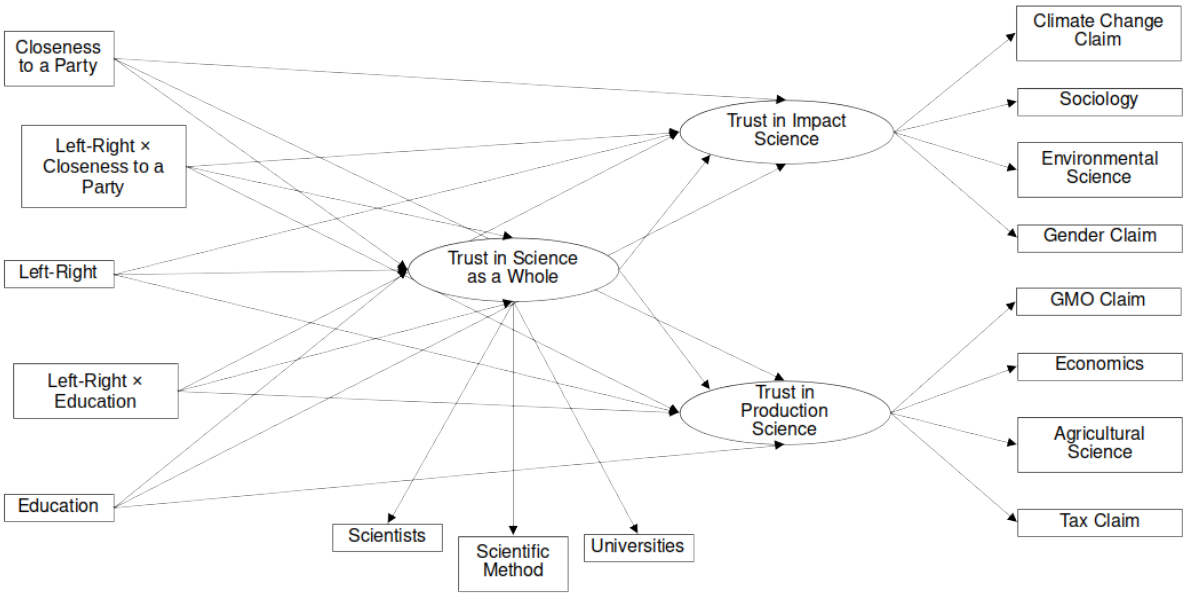
Appendix 1.4: Tests of differences in perceived accuracy

Main (M1) and interaction (M2) unstandardized regression coefficients. $N = 1054$; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Aggregated		**Immigration**		**Taxation**		**GMOs**		**Climate Change**	
	M1	**M2**	**M1**	**M2**	**M1**	**M2**	**M1**	**M2**	**M1**	**M2**
Intercept	2.24^{***}	2.19^{***}	2.43^{***}	2.41^{***}	0.90^{***}	0.81^{***}	2.13^{***}	2.09^{***}	3.49^{***}	3.49^{***}
Conservative (vs. Liberal)	-0.64^{***}	-0.55^{***}	-1.24^{***}	-1.20^{***}	0.32^{***}	0.49^{***}	-0.53^{***}	-0.43^{***}	-1.11^{***}	-1.10^{***}
Production vs. Impact	0.02	0.07	−0.05	−0.10	−0.01	0.17	0.16	0.31^{*}	−0.02	−0.16
Control vs. Impact	-0.15^{**}	−0.06	-0.20^{*}	−0.10	-0.26^{**}	−0.24	0.01	0.01	-0.19^{*}	−0.05
General vs. Impact	0.00	0.05	0.11	0.14	−0.05	0.09	−0.03	−0.02	−0.01	−0.00
Conservative x Production	–	−0.09	–	0.10	–	-0.34^{*}	–	−0.31	–	0.27
Conservative x Control	–	−0.18	–	−0.19	–	−0.05	–	−0.03	–	−0.28
Conservative x General	–	−0.09	–	−0.07	–	−0.30	–	−0.04	–	−0.00
Adjusted R^2	0.06	0.06	0.25	0.25	0.03	0.04	0.05	0.05	0.24	0.25
F-difference	–	–	–	0.81	–	2.00	–	1.05	–	3.38^{*}

Chapter 2

Appendix 2.1: Full hypothesized model



Covariances between factors, controls, and dummy variables are not shown for simplicity.

Appendix 2.2: Comparisons between SEMs

Scaled fit indices and unbiased SRMR for each structural equation model. The degrees of freedom of the LR-tests are determined by the difference in parameters between models. $N = 5410$.

M	Param.	RMSEA	SRMR	CFI	TLI	LRT ($\Delta\chi^2$)	LRT p-value
1	45	0.043	0.042	0.846	0.830	–	–
2	48	0.043	0.042	0.847	0.829	15.725	0.001
3	51	0.044	0.042	0.847	0.827	12.933	0.005
4	87	0.043	0.030	0.879	0.835	501.770	$< .001$
5	90	0.040	0.028	0.897	0.856	148.040	$< .001$

Appendix 2.3: Full model (M5) factor loadings

Completely standardized parameters. $N = 5410$; *** $p < 0.001$.

Trust Items	**Trust in Impact Science**	**Trust in Production Science**	**Trust in Science as a Whole**
Climate Change Claim	0.574***	–	–
Sociology	0.711***	–	–
Environmental Science	0.812***	–	–
Gender Claim	0.360***	–	–
GMO Claim	–	0.356***	–
Economics	–	0.702***	–
Agricultural Science	–	0.718***	–
Tax Claim	–	0.108***	–
Scientists	–	–	0.894***
Scientific Method	–	–	0.878***
Universities	–	–	0.706***

Chapter 3

Appendix 3.1: All possible messages and their variations per topic for the *single-report* questionnaire

Climate Change

Belief 1	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that **an abnormal number of powerful sunspots is driving the rise in average global temperatures**. Our understanding of the world's climate is advancing.
Belief 2	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that **extreme industrial steel production is driving the rise in average global temperatures**. Our understanding of the world's climate is advancing.
Expectation 1	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that **soon, U.S. winters will become colder due to shifting air-pressure dynamics**. Our understanding of the world's climate is advancing.
Expectation 2	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that **soon, U.S. winters will become warmer due to shifting air-pressure dynamics**. Our understanding of the world's climate is advancing.

AI

Belief 1	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **AI works well with practice cases but performs poorly in real-life diagnosis**. Our understanding of the role of human expertise in AI is advancing.
Belief 2	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **AI works well with practice cases and performs in real-life diagnosis**. Our understanding of the role of human expertise in AI is advancing.
Expectation 1	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **oncologists will still be crucial to validate diagnoses even as AI advances**. Our understanding of the role of human expertise in AI is advancing.
Expectation 2	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **soon, oncologists will not be needed to validate diagnoses as AI advances**. Our understanding of the role of human expertise in AI is advancing.

Nutrition

Belief 1	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat do not outweigh the lack of heme-iron**. Our understanding of the role of plant-based alternatives on health is advancing.
Belief 2	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat outweigh the lack of heme-iron**. Our understanding of the role of plant-based alternatives on health is advancing.
Expectation 1	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that **key nutrients in animal meat will be impossible to produce in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.
Expectation 2	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that **soon, key nutrients in animal meat will be produced in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.

Population Growth

Belief 1	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that the **total world population is shrinking due to societal practices**. Our understanding of these demographic trends is advancing.
Belief 2	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that the **total world population is growing due to societal practices**. Our understanding of these demographic trends is advancing.
Expectation 1	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that **the United States' birthrate will remain well below the replacement rate, leading to a decrease in population**. Our understanding of these demographic trends is advancing.
Expectation 2	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that **soon, the United States' birthrate will rise well above the replacement rate, leading to an increase in population**. Our understanding of these demographic trends is advancing.

Migration

Belief 1	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **the number of people moving from developing to developed countries is now a substantial 5% lower than in 1990**. Our understanding of migration and its impact on host countries is advancing.
Belief 2	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **the number of people moving from developing to developed countries is now a substantial 5% higher than in 1990**. Our understanding of migration and its impact on host countries is advancing.
Expectation 1	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **soon, immigration will increase public housing availability in developed countries**. Our understanding of migration and its impact on host countries is advancing.
Expectation 2	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **soon, immigration will increase public housing shortages in developed countries**. Our understanding of migration and its impact on host countries is advancing.

Appendix 3.2: All possible messages and their variations per topic for the *double-report* questionnaire

Reports interacting with beliefs and expectations were presented in random order, such that either could appear first.

Climate Change

Belief 1 + Expectation 1	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that an **abnormal number of powerful sunspots is driving the rise in average global temperatures**. Researchers continue to monitor these trends and their effects. Another report shows that **soon, U.S. winters will become colder due to shifting air-pressure dynamics**. Our understanding of the world's climate is advancing.
Belief 1 + Expectation 2	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that an **abnormal number of powerful sunspots is driving the rise in average global temperatures**. Researchers continue to monitor these trends and their effects. Another report shows that **soon, U.S. winters will become warmer due to shifting air-pressure dynamics**. Our understanding of the world's climate is advancing.
Belief 2 + Expectation 1	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that **extreme industrial steel production is driving the rise in average global temperatures**. Researchers continue to monitor these trends and their effects. Another report shows that **soon, U.S. winters will become colder due to shifting air-pressure dynamics**. Our understanding of the world's climate is advancing.
Belief 2 + Expectation 2	**New Information About Climate Change:** Experts are making strides in understanding Earth's changing climate. A report shows that **extreme industrial steel production is driving the rise in average global temperatures**. Researchers continue to monitor these trends and their effects. Another report shows that **soon, U.S. winters will become warmer due to shifting air-pressure dynamics.** Our understanding of the world's climate is advancing.

AI

Belief 1 + Expectation 1	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **AI works well with practice cases but performs poorly in real-life diagnosis**. Researchers continue testing and refining AI systems. Another report shows that **oncologists will still be crucial to validate diagnoses even as AI advances**. Our understanding of the role of human expertise in AI is advancing.
Belief 1 + Expectation 2	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **AI works well with practice cases but performs poorly in real-life diagnosis**. Researchers continue testing and refining AI systems. Another report shows that **soon, oncologists will not be needed to validate diagnoses as AI advances**. Our understanding of the role of human expertise in AI is advancing.
Belief 2 + Expectation 1	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **AI works well with practice cases and performs in real-life diagnosis**. Researchers continue testing and refining AI systems. Another report shows that **oncologists will still be crucial to validate diagnoses even as AI advances**. Our understanding of the role of human expertise in AI is advancing.
Belief 2 + Expectation 2	**New Information About Artificial Intelligence (AI):** Experts are making strides in understanding AI's capabilities. A report shows that **AI works well with practice cases and performs in real-life diagnosis**. Researchers continue testing and refining AI systems. Another report shows that **soon, oncologists will not be needed to validate diagnoses as AI advances**. Our understanding of the role of human expertise in AI is advancing.

Nutrition

Belief 1 + Expectation 1	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat do not outweigh the lack of heme-iron**. Researchers continue exploring the benefits and drawbacks of meat alternatives. Another report shows that **key nutrients in animal meat will be impossible to produce in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.
Belief 1 + Expectation 2	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat do not outweigh the lack of heme-iron**. Researchers continue exploring the benefits and drawbacks of meat alternatives. Another report shows that **soon, key nutrients in animal meat will be produced in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.
Belief 2 + Expectation 1	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat outweigh the lack of heme-iron**. Researchers continue exploring the benefits and drawbacks of meat alternatives. Another report shows that **key nutrients in animal meat will be impossible to produce in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.
Belief 2 + Expectation 2	**New Information About Plant-Based Alternatives:** Experts are making strides in understanding the health effects of plant-based meat. A report shows that the **health benefits of replacing animal meat with plant-based meat outweigh the lack of heme-iron**. Researchers continue exploring the benefits and drawbacks of meat alternatives. Another report shows that **soon, key nutrients in animal meat will be produced in a lab using vegetable sources**. Our understanding of the role of plant-based alternatives on health is advancing.

Population Growth

Belief 1 + Expectation 1	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that the **total world population is shrinking due to societal practices**. Researchers continue to examine the implications of these demographic shifts. Another report shows that **the United States' birthrate will remain well below the replacement rate, leading to a decrease in population**. Our understanding of these demographic trends is advancing.
Belief 1 + Expectation 2	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that the **total world population is shrinking due to societal practices**. Researchers continue to examine the implications of these demographic shifts. Another report shows that **soon, the United States' birthrate will rise well above the replacement rate, leading to an increase in population**. Our understanding of these demographic trends is advancing.
Belief 2 + Expectation 1	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that the **total world population is growing due to societal practices**. Researchers continue to examine the implications of these demographic shifts. Another report shows that **the United States' birthrate will remain well below the replacement rate, leading to a decrease in population**. Our understanding of these demographic trends is advancing.
Belief 2 + Expectation 2	**New Information About Demographic Trends:** Experts are making strides in understanding global population trends. A report shows that the **total world population is growing due to societal practices**. Researchers continue to examine the implications of these demographic shifts. Another report shows that **soon, the United States' birthrate will rise well above the replacement rate, leading to an increase in population**. Our understanding of these demographic trends is advancing.

Migration

Belief 1 + Expectation 1	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **the number of people moving from developing to developed countries is now a substantial 5% lower than in 1990**. Researchers continue to examine the effects of migration policies. Another report shows that **soon, immigration will increase public housing availability in developed countries**. Our understanding of migration and its impact on host countries is advancing.
Belief 1 + Expectation 2	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **the number of people moving from developing to developed countries is now a substantial 5% lower than in 1990**. Researchers continue to examine the effects of migration policies. Another report shows that **soon, immigration will increase public housing shortages in developed countries**. Our understanding of migration and its impact on host countries is advancing.
Belief 2 + Expectation 1	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **the number of people moving from developing to developed countries is now a substantial 5% higher than in 1990**. Researchers continue to examine the effects of migration policies. Another report shows that **soon, immigration will increase public housing availability in developed countries**. Our understanding of migration and its impact on host countries is advancing.
Belief 2 + Expectation 2	**New Information About Migration:** Experts are making strides in understanding global migration flows. A report shows that **the number of people moving from developing to developed countries is now a substantial 5% higher than in 1990**. Researchers continue to examine the effects of migration policies. Another report shows that **soon, immigration will increase public housing shortages in developed countries**. Our understanding of migration and its impact on host countries is advancing.

Appendix 3.3: Ordinal logistic regression models for the tests of H1 and H2

Model parameter estimates are in log odds. N = 4,979; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Surprise (H1)		**Interestingness (H2)**	
Not at all | Slightly	−1.316 ***	−1.294 ***	−2.426 ***	−2.227 ***
Slightly | Somewhat	−0.431 ***	−0.405 ***	−1.261 ***	−1.008 ***
Somewhat | Very	0.549 ***	0.578 ***	−0.002	0.333 ***
Very | Extremely	1.886 ***	1.919 ***	1.692 ***	2.113 ***
Reinforce (vs. Contradict)	−1.461 ***	−1.473 ***	−0.313 ***	−0.341 ***
Validate (vs. Violate)	−1.448 ***	−1.456 ***	−0.131 *	−0.125 *
Topic (vs. AI)				
Climate Change	−0.004	−0.085	−0.119	−0.719 ***
Migration	0.117	0.089	−0.475 ***	−0.622 ***
Nutrition	0.121	0.096	−0.325 ***	−0.581 ***
Population Growth	0.233 **	0.218 **	−0.436 ***	−0.592 ***
Familiarity		0.052		0.158 ***
Knowledge		0.012		0.297 ***
College (vs. No college)		0.133 *		0.188 ***
Republican (vs. Democrat)		−0.124 *		−0.618 ***
Female (vs. Male)		−0.157 **		−0.076
AIC	14,408.290	14,389.740	15,144.560	14,731.770
BIC	14,473.420	14,487.430	15,209.690	14,829.470
Log Likelihood	−7,194.150	−7,179.870	−7,562.280	−7,350.890

Appendix 3.4: Ordinal logistic regression models for the tests of H3–H5

Model parameter estimates are in log odds. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Sharing Intention			
	Single-Report (H3 & H4)		**Double-Report (H5)**	
Not at all | Slightly	0.345 ***	1.258 ***	0.271 **	1.263 ***
Slightly | Somewhat	1.066 ***	2.048 ***	1.094 ***	2.152 ***
Somewhat | Very	1.956 ***	3.010 ***	1.995 ***	3.116 ***
Very | Extremely	3.152 ***	4.263 ***	3.257 ***	4.428 ***
Reinforce (vs. Contradict)	0.050	0.049	0.025	0.031
Validate (vs. Violate)	0.186 **	0.239 ***	0.046	0.094
Reinforce × Validate			0.160	0.141
Topic (vs. AI)				
Climate Change	−0.051	−0.948 ***	−0.003	−0.790 ***
Migration	−0.291 **	−0.587 ***	−0.181	−0.404 ***
Nutrition	−0.133	−0.523 ***	−0.170	−0.490 ***
Population Growth	−0.282 **	−0.557 ***	−0.267 **	−0.493 ***
Familiarity		0.279 ***		0.195 ***
Knowledge		0.429 ***		0.436 ***
College (vs. No college)		0.129 *		0.115
Republican (vs. Democrat)		0.006		0.198 ***
Female (vs. Male)		−0.139 *		−0.003
AIC	11,646.440	11,092.740	11,741.200	11,275.680
BIC	11,711.570	11,190.440	11,812.610	11,379.560
Log Likelihood	−5,813.220	−5,531.370	−5,859.600	−5,621.840
Observations	N = 4,979		N = 4,876	

Chapter 4

Appendix 4.1: Observational study data

Fact-checking databases

To ensure diversity in data sources, we constructed our tweet corpus by integrating two fact-checking databases, which rely on distinct selection criteria and are maintained for different institutional purposes: Meta's Condor and Science Feedback.

Condor Meta's Condor database is part of the *Social Science One*[4] dataset, which contains URLs that were "shared (as an original post or a reshare) with *public* privacy settings more than 100 times" Messing et al., 2020.

Some URLs in Condor were flagged and referred to third-party fact-checkers, who may have assigned ratings to evaluate the veracity of the linked content. If a URL was sent for review but no verdict was returned, its label is recorded as "not rated." Our Condor subset includes all URLs that were referred to fact-checkers, and were posted between 1 January 2011 and 5 July 2022 (N = 48,651).

Science Feedback Science Feedback is a fact-checking initiative that evaluates the accuracy of scientific claims circulating online. It focuses on articles or statements that are both verifiable and widely shared on social media. According to its methodology Science Feedback, 2025, editors select claims or articles for review if they have gone viral on social media platforms or originate from sources with large readerships, provided they contain scientifically verifiable content.

On 22 December 2022, we requested all appearances of fact-checked claims that were published on the database between 1 January 2000 and 22 December 2022 using the main endpoint for Science Feedback's Appearances API[5] (N = 10,928). Subsequently, we iterated through each appearance's notice to collect metadata for each appearance. Another call was made in September 2023 for the same URLs using the new end point

[4]https://socialscience.one/request-proposals-facebook-urls-dataset
[5]https://api.feedback.org/appearances

which provided more metadata fields.

Tweet collection and processing

Data collection was conducted in two rounds: the first in March 2023, ahead of the planned shutdown of the Twitter API, and the second in April 2023, after the API remained temporarily available.

Tweets and associated metadata were retrieved using the *minet* tool (Plique et al., 2019), which enabled the extraction of both tweet-level and user-level information. We collected all tweets, retweets, replies, and quote tweets that contained any URL from our dataset, either directly embedded or included in the tweet text. Prior to the retrieval phase, manual verification of the URLs and their titles was performed by different annotators to discard invalid queries (e.g. URLs pointing to TLDs such as *youtube.com* and *twitter.com*, URLs with uninformative or empty titles).

In this work, we excluded quote tweets, replies, and tweets that contained multiple URLs to ensure unambiguous attribution between tweets and content ratings. Retweet counts were winsorized at the 99.9th percentile to limit the influence of extreme outliers and stabilize statistical model fitting.

The final tweets corpus contains 8,729,332 tweets shared between January 2010 and April 2023, linking to 56,508 unique URLs from 13,217 distinct domains. The tweets were posted by 2,987,722 users, ranging in followers between 0 and 134,605,299. The most active user shared 55,096 tweets (a journalist), but the median number of tweets per user is 1. 85.5% of the tweets did not receive any retweets, whereas the average tweet receives 3.05 retweets.

Appendix 4.2: Observational study annotations and labeling

Fact-check ratings

Both *Condor* and *Science Feedback* provide a set of fact-check ratings that do not map directly onto a True/False dichotomy. To facilitate comparisons between categories and between databases, we standardized all fact-check ratings into four unified labels: *True*, *False*, *Neutral*, and *Other*.

Condor The Condor database includes nine distinct fact-check ratings. We applied the following mapping scheme:

- "Fact checked as true" → True
- "Fact checked as false", "missing context", "mixture or false headline", "altered media" → False
- "Not rated", "Not eligible" → Neutral
- "Opinion", "Satire" → Other

Science Feedback The Science Feedback database contains 36 ratings, ranging from "Very high credibility" to "Pants on fire." However, 99.9% of the rated entries also include a numerical accuracy score ranging from 0 (lowest accuracy) to 5 (highest accuracy), with 3 denoting partial or neutral credibility. We averages scores if a URL had more than one.

We therefore mapped the ratings as follows:

- Score > 3 → True
- Score < 3 → False
- Score $= 3$ → Neutral

Combining ratings Some URLs received multiple, and occasionally inconsistent, fact-checks. This could occur because a URL was (a) collected multiple during our dataset construction, (b) re-flagged following renewed virality, or (c) represented by multiple variants of the same normalized link (due to URL shortening, redirection, or normalization artifacts). Additionally, URLs appearing in both databases received at least one rating

from each source. In the combined dataset, we resolved such cases by collapsing the simplified labels using the following rules:

- If a URL had any False rating (even if mixed with Neutral), it was labeled False.
- If a URL had any True rating (even if mixed with Neutral), it was labeled True.
- URLs combining True, False, and/or Other ratings were classified as Other.
- URLs with only Neutral ratings remained Neutral.

Domain-level veracity measures

To qualify the effect beyond individual ratings which may contain errors or reflect some biases, and to cover parts of our unrated or neutral URLs, we incorporated two external domain-level annotation schemes to complement the URL-level fact-checks. For the full distribution of URLs per category, see Table A33.

Factuality levels Media Bias/Fact Check (MB/FC) provides domain-level ratings (Media Bias/Fact Check, 2024), which classify outlets into six categories of factuality: *Very High, High, Mostly Factual, Mixed, Low, and Very Low.*

To align with the fact-check rating analysis (which primarily compares the diffusion of true versus false URLs), we report in the main text only two aggregated factuality categories to maintain consistency with this True/False distinction. Specifically, we merge *Very High* and *High* into a single "High" category, combine *Very Low* and Low into a "Low" category, and group *Mixed* and *Mostly Factual* into a "Mid" category (with results reported in Table A7). In our dataset, 33.2% of URLs and 60.1% of tweets have an associated MB/FC factuality level.

Credibility scores Science Feedback provides a *Consensus Credibility Score* (Science Feedback, 2023) for more than 20k domains. These scores are computed by aggregating publicly available credibility assessments from a broad range of expert sources, including academic researchers, media professionals, and civil society organizations. To align with the fact-check rating analysis, we create three categories or *credibility levels* based on the distribution of the scores in our data into three terciles: *Low*, *Mid* and *High*. In the main text we report on the *Low* and *High* categories, and provide details for the *Mid* category

alongside the other two in Table A8. Within our dataset, 54.2% of URLs and 54.0% of tweets were associated with an estimated credibility score.

Scientific Authority Cues (SACs)

To identify the presence of Scientific Authority Cues (SACs) in article titles, we implemented a multi-step semi-automatic annotation procedure combining large language model (LLM) annotations with human validation. We started by first manually removing all unusable URLs from our collection (very short titles, links to images, empty titles). We also removed any URL that had different titles in different metadata fields (the title metadata from fact-checking datasets, page titles from different collection campaigns, etc.).

Based on our close reading and annotation of a subset of titles from the datasets, we identified 5 categories that can be considered scientific authorities:

- *Individuals*: Specific named persons who can be considered a scientific authority, e.g. "Dr. Fauci."
- *Roles*: Unnamed individuals that indicate a scientific authority (e.g. professions, titles), e.g. "scientists."
- *Institution*: Organizations such as universities, public research centers and agencies, e.g. "Oxford."
- *Companies*: Private commercial entities directly involved in scientific research or industry, e.g. "Pfizer."
- *Other*: Words or phrases strongly suggesting a scientific source or authority, e.g. "study."

After evaluating several existing pre-trained fine-grained Named Entity Recognition models, we found that they reliably identified well-known entities but struggled to detect less prominent ones. As expected, the models also performed poorly when encountering unnamed or generic entities.

We therefore employed a general-purpose large language model (OpenAI's *gpt-4o-mini*) and followed prompt-engineering guidelines to automatically annotate the titles associated with each URL in our dataset (see Prompt 1 in Appendix 4.5).

The annotation followed the following procedure:

1. *Automatic candidate extraction*: We prompted OpenAI's *gpt-4o-mini* through the API to extract potential mentions of scientific authorities from each title (previously translated to English through another API call), using a predefined list of entity types (individuals, institutions, roles, and sources). This step produced a list of *candidate entities* indicating scientific authority mentions.
2. *Sample annotation*: To assess extraction quality, we manually annotated a random sample of 400 titles for scientific authority mentions and compared them with the LLM outputs. The LLM exhibited perfect recall but a low recall. Based on this finding, we adopted a conservative strategy: retaining all automatically extracted candidates and subsequently removing false positives through human validation.
3. *Human filtering*: Three independent annotators reviewed the extracted entities and labeled each as Scientific Authority, Not Scientific Authority, or Uncertain, based on a simple operational definition of the concept. When encountering unfamiliar names, particularly less-known individuals, institutions or acronyms, annotators verified their status online to the best of their ability. It is also at this stage that we identified the presence of references to some unnamed institutions ("a lab") and named publications ("The Lancet"). The annotators thus indicated if an entity was named or unnamed, and retyped them into four categories: *individual*, *institution*, *source*, or *business*. Annotators had the possibility to rate a candidate mention as "maybe" if they could not squarely put it in one category or another.
4. *Scoring and aggregation*: Each mention received a numerical score: +1 for *Yes*, 0.5 for *Maybe*, and 0 for *No* (from most to least indicative of scientific authority). From these, we derived three levels of annotation agreement:
    - *Unanimity*: score > 2 2.5
    - *Majority*: score ≥ 2
    - *Any*: score < 2

Inter-annotator agreement was assessed using the average pairwise Cohen's κ, which was 0.634 across the three annotators. The annotators reached unanimous agreement on 72.5% of the mentions, classifying them as either correct or incorrect. A majority

agreement (at least two annotators) was achieved for 54.7% of the mentions identified as Scientific Authority Cues (SACs). Most disagreements among annotators concerned entities that appeared infrequently in the corpus.

In the main text, we report results based on the majority agreement level, as it offers a practical balance between accuracy and coverage. Results for all agreement levels are provided in the appendix tables to show the consistency of the observed effects.

Attribution vs. Mention To distinguish Scientific Authority Attributions (SAAs) from Scientific Authority Mentions (SAMs) in the observational data (paralleling the experimental manipulation) we prompted the LLM to classify whether each title contains an attributable claim (e.g., "Emergency COVID-19 Vaccines May Cause Massive Side Effects") or not (e.g., "George W. Bush Military Tribunal: Day 2"). When the SAC was present, we prompted the model to take into consideration to assess attribution, and we also asked the model to extract the claim when the title contains one (see Prompts 3 and 4 in Appendix 4.5).

To evaluate model performance on this task, three annotators independently labeled a balanced sample of 400 titles, assessing whether each contained a claim attributable to a scientific authority and operationalized as titles that can be meaningfully reframed as "*According to [SAC], [claim]*" or "*Study shows [claim].*" The sample was stratified by LLM label (claim vs. no claim), SAC presence, fact-check rating (true vs. false), and political slant (as measured by the average partisanship score of users tweeting the URL).

Inter-annotator agreement was moderate (Mean pairwise Cohen's $\kappa = 0.598$), signaling the difficulty of the task even for human annotators. However, it exceeded the agreement between the majority-vote human label and the LLM (Cohen's $\kappa = 0.425$). Crucially, disagreement was concentrated in the uncued subset of titles (LLM vs. human $\kappa = 0.21$), while agreement was substantially higher for cued titles ($\kappa = 0.64$). This pattern indicates that both human annotators and the LLM reliably distinguish SAAs from SAMs when a SAC is present (label agreement = 82.0%), but diverge on what constitutes an attributable claim in the absence of a SAC.

This asymmetric error structure renders a one-to-one observational reconstruction of the experimental design unreliable, as the high misclassification rate in the negative class

would systematically bias any SAA-restricted analysis. Agreement among annotators and with the LLM were comparable for the other categories (Left vs. Right and True vs. False). Nonetheless, we re-estimated our main observational models restricting the effect to SAAs (SAC on claim-like titles), and report results in Table A19, with further interpretation in Appendix 4.4 (Scientific Authority Attribution only analysis).

User partisanship

For a subset of users, we estimate political partisanship using the scores provided by the authors (Mosleh & Rand, 2022). These scores are derived from the proportion of political elite accounts followed by each user across parties, ranging from -1 (users who follow only Democratic political elite accounts) to $+1$ (users who follow only Republican ones). Consistent with the original operationalization, we classify users with positive scores as right-leaning and those with negative scores as left-leaning, representing 24.0% and 28.0% of our user base, respectively.

As a robustness check, we further annotated available user biographies using an instruction-fine-tuned large language model (Qwen2.5-7B-Instruct), which classifies users based on their bios into four categories: Left, Right, Other, and None (see Prompt 2 in Appendix 4.5). The analysis is described in Appendix 4.4 (under self-identification) and the results of this complementary classification are presented in Table A17.

Bot screening

To assess whether automated accounts confound our main findings, we queried Botometer X's API (Yang et al., 2020, 2023) for all users in our tweets dataset (N = 2,987,722), and retrieved a score for 99% of them.

Botometer scores range from 0 to 1, where higher values indicate greater likelihood of automated or bot-like behavior, with 0.5 commonly used as a classification threshold. We re-estimated our main observational models on the full dataset including Botometer score as a continuous covariate, to test whether the SAC effect persists after adjusting for the degree of bot-like activity of the posting user.

To assess the presence of bots among "spreaders", we have similarly scored a strat-

ified subsample of users who retweeted some tweet in our dataset (N = 16,000), sampled equally the relevant analysis categories (SAC presence in the original tweet's URL, factcheck rating, and political leaning). We compared mean Botometer scores across all relevant subgroups to test whether automated accounts are differentially concentrated in any cell of our design.

The analyses are described in Appendix 4.4 (bot sensitivity analysis), and the results are reported in Tables A15 and A35.

Appendix 4.3: Observational study extended results

Operationalization and modeling choice

Estimating information spread can take several forms depending on (meta)data availability and modeling assumptions. Based on the fields returned by the Twitter API at the time of collection, certain dimensions of diffusion could not be studied directly. Specifically, we lacked (a) cascade structures, as retweet chains were unavailable – only an "original tweeter" is assigned for each tweet, (b) influence networks, since follower–followee relationships are not included, and (c) impression counts (views), which the API provides only for tweets posted after December 2022, corresponding to only the last four months of our collection period.

Because retweeting (excluding quote tweets) is widely accepted as a behavioral signal of sharing intent and the principal mechanism of information diffusion on Twitter, most prior studies operationalize higher spread as higher retweet counts. To model this statistically, we use count regression, with the key explanatory variable being the presence of a SAC (and controlling for log-transformed follower count, which is strongly correlated with retweet volume). While Poisson regression is occasionally used in similar contexts, retweet counts in our data were overdispersed, making the Negative Binomial regression the appropriate specification. Another theoretically more appropriate model for this setting is the Zero-Inflated Negative Binomial (ZINB), which explicitly accounts for the two processes underlying tweet diffusion: (1) the likelihood of being retweeted at all (logistic component) and (2) the number of retweets conditional on being shared (count component). In our tests, ZINB models, though slower to converge due to their increased parameterization and mixture structure, produced nearly identical results to the simpler Negative Binomial model in terms of effect sizes, confidence intervals, p-values, and model fit, so we only included the results of the negative binomial models.

To test the robustness of our findings to modeling choices (i.e., beyond modeling retweet counts), we also study it under two other operationalizations: **Retweet Likelihood** and **Overall Spread**, both detailed in Appendix 4.4. Because of the potential noise in the annotation process and the inherent subjectivity in identifying Scientific Authority Cues (SACs), we run every experiment described in the paper and in the Appendix for all

annotation agreement level: (*Unanimity*, *Majority*, and *Any*).

Finally, for each parameter in the model, we report the exponentiated regression coefficients, which represent the incidence rate ratio, or IRR, with 95% confidence intervals (CIs), and corresponding p-values. In these experiments, the IRR quantifies the proportional change in the expected number of retweets associated with the presence of a Scientific Authority Cue (SAC). Specifically, the IRR compares the expected retweet count for cued tweets to that for uncued tweets. An IRR greater than 1 indicates that the presence of SACs is associated with an increase in expected retweet counts, whereas an IRR below 1 indicates a decrease.

In the following subsections, we provide detailed commentary on the experimental results summarized in the main text. We elaborate on how the findings vary across annotator agreement levels, and we document additional supporting experiments, implementation precisions, and data-processing details was not be reported in full in the main article. These supplementary explanations are intended to offer greater methodological transparency and to aid in the interpretation and replication of our analyses.

Study A.1 In this study, we investigate the presence of the overall effect of SAC on the spread of information over all collected tweets, and how it varies by cue category (“named” versus “unnamed”, and “individual” versus “institution” versus “source”).

For all level of annotation agreement, the presence of a SAC significantly increases the predicted number of retweets compared to its absence, as indicated by the incidence rate ratios in Table A1. The estimated coefficients remain consistent across agreement levels, suggesting that the effect is robust to variations of the subjective perception of "scientific authority" among annotators.

The effect remains robust when the dataset is divided into its two components and is positive and statistically significant in both cases. The magnitude of the effect is larger for Condor (Table A2) than for Science Feedback (Table A3), likely reflecting differences in how the URLs are chosen to be fact-checked in each case: *Condor* includes URLs from a wide range of domains that are particularly viral (i.e., based on their spread Facebook), whereas *Science Feedback* focuses almost exclusively on scientific misinformation. These results indicate that the observed effect is present even when varying the selection criteria

used to determine what content gets fact-checked.

In Table A4, we see that the effect is stronger for named cues than for unnamed cues, although the effect remains generally comparable. We note that titles containing both named and unnamed entities were tagged both as *cued (named)* and *cued (unnamed)*.

The breakdown by entity type (Table A5) shows that individuals are associated with the strongest effect, followed by institutions and then sources, across two levels of annotator agreement ("Unanimity" and "Majority"). When considering the most inclusive level or agreement, institutions drop to third place, reflecting the fact that they account for most of the mentions on which annotators disagreed.

Study A.2 In this study, we examine the relationship between SACs and the veracity of the content being shared.

We begin by analyzing the effect within each subset of tweets linking to URLs of given fact-check rating (True, False, Neutral, Other). To move beyond URL-level assessments, we also examine how the effect varies when using domain-level measures of factuality and credibility. Specifically, we rely on the Media Bias/Fact Check (Media Bias/Fact Check, 2024) Factuality levels and the Science Feedback Consensus Credibility Scores (Science Feedback, 2023). For this purpose, URLs were mapped to their corresponding domains by extracting the top-level domain (TLD). For brevity, we refer to URLs with a false fact-check rating, or originating from a domain with a "Low" factuality level or a low credibility score as *low veracity content* (similarly for *high veracity content* referring to true fact-checks and high factuality or credibility).

In Table A6, we find that cued tweets linking to content fact-checked as False or Neutral receive a significant boost in retweets, whereas those linking to content rated as True or categorized as Other (satire, opinion, or mixture of true and false ratings) show a negative effect. A similar pattern is observed when examining domain-level factuality levels (Table A7) and credibility score (Table A8): the higher the associated factuality, the weaker the SAC effect. This trend is further confirmed by the interaction between SAC presence and domain credibility score (Table A9), which is consistently negative and significant for all levels of annotator agreement. Finally, unlike for content fact-checked as True, we observe that SACs are associated with a positive effect for high factuality

and credibility domains. Zooming on the Neutral category (Table A16), we see that the effect is concentrated in the "mid" factuality and credibility range, suggesting that SACs particularly boost content that sits at the gray area of fact-checking, where credibility signals become more potent.

In summary, these results suggest that SACs are particularly associated with increased sharing of low veracity content, whether it is evaluated using URL-level or domain-level measures.

Study A.3 In this study, we examine the relationship between SACs and the estimated political orientation (*partisanship*) of users sharing the URLs, as well as the interaction between partisanship and content veracity. In this analysis, we use on the political ideology estimates based on the political accounts that each user follows, as provided by (Mosleh & Rand, 2022). For robustness, we also conduct the same experiments using two external political labels, which we describe in Appendix 4.4 (study A.3). We report these complementary results in Tables A17 and A18.

Across all levels of annotator agreement, we find that cued tweets posted by left-leaning users receive a smaller amplification than those posted by right-leaning users, relative to their respective baselines (Table A10). When focusing on tweets linking to content with a True/False fact-check ratings (Table A11), High/Low domain-level factuality level (Table A12), or High/Low credibility score (Table A13), a clear asymmetry emerges. SACs increase the diffusion of low veracity content posted by both left-leaning and right-leaning users, but their effect on high veracity content is weaker for right-leaning users and negative for content shared by left-leaning users. This pattern helps explain the overall difference observed in Study A.2: while right-leaning users experience amplification from SACs regardless of content veracity, left-leaning users show a boost only for low-veracity content, with high-veracity content showing the opposite effect, which nearly cancels out when aggregating across veracity categories.

Appendix 4.4: Observational study robustness checks and alternative models

Alternative models

In this section, we examine alternative ways of operationalizing the spread of content on Twitter/X. We reproduce the main results from the three observational studies: (A.1) the overall boosting effect of SACs, (A.2) how this effect varies across different veracity measures, and (A.3) how partisanship influences the overall effect and interacts with veracity. To do so, we propose two alternative operationalizations of content spread:

- **Retweet Likelihood**: defined as whether a tweet containing a SAC receives at least one retweet. This is an signal, as the majority of tweets receive no retweets at all. Retweet Likelihood also provides a complementary perspective to Retweet Counts, as it is not influenced by the self-exciting dynamics of retweeting—tweets that begin to accumulate retweets are subsequently more likely to receive additional ones due to increased exposure. Conceptually, this measure is also closer to the design of our survey experiment.
- **Overall Spread**: defined as the total number of retweets a URL receives aggregated across all distinct accounts that shared it. This reflects whether content with a SAC achieves greater overall diffusion, but it does not account for differences in the properties of the users who posted the URL (e.g., follower count, partisanship). Such URL outcome, while it has the advantage of abstracting away tweet- or user-specific factors that cannot be directly observed or controlled for by providing an alternative aggregation, exhibits higher variance due to the smaller number of observations and introduces new sources of bias related to website audience size, as URLs from high-traffic domains are naturally posted more frequently.

We model Retweet Likelihood using logistic regression (with log-transformed follower count as a covariate) and report the exponentiated coefficients as odds ratios (ORs), which capture the change in the likelihood of a tweet being retweeted. For Overall Spread, we use negative binomial regression (with no additional covariates) and report exponentiated coefficients as incidence rate ratios (IRRs), which indicate the multiplicative change in

expected total retweet counts. We briefly summarize the observations in the following subsections and report the ORs and IRRs corresponding to the *cued* indicator variable (associated with the presence of a SAC at a given annotator agreement level), along with their 95% confidence intervals and p-values, in Tables A21–A26.

In addition to the alternative operationalizations above, we report two further robustness checks that address user-level variance in the main Negative Binomial models used in the main observational study:

- **Clustered Standard Errors (CSE)**: we re-estimate all main models with standard errors clustered at the user level, to account for the non-independence of tweets originating from the same poster. This specification leaves point estimates unchanged but corrects for the underestimation of uncertainty that arises when treating tweets from the same user as independent observations.
- **Mixed-effects Model (ME)**: we re-estimate all main models as mixed-effects Negative Binomial models with user-level random intercepts, to account for stable baseline differences in engagement across posters (e.g., users with systematically more engaged follower networks). This specification estimates a *conditional* (within-user) effect of SACs rather than the *population-average* effect reported in the main analyses.

We report the results in Tables A27–A32.

Study A.1 In Table A21, we find that the presence of SACs has a positive and significant effect on the likelihood that a tweet is retweeted. This effect is robust across all levels of annotator agreement and for both subsets of fact-check databases. For Overall Spread (Table A24), the effect of SACs is positive and significant for Condor and for the combined dataset, but not for Science Feedback. This null result may reflect the smaller number of observations in the Science Feedback subset, as well as the fact that the content curated in this database is less characteristically "viral", compared to that of Condor.

The SAC effect on retweet counts is robust to correcting for within-user dependence: clustering standard errors at the user level leaves the overall effect and statistically significant (Table A27). Finally, the SAC effect remains positive and statistically significant,

though attenuated relative to the main specification (Table A30). This attenuation is expected: mixed-effects models estimate the *conditional* (within-user) effect of SACs, accounting for between-user differences in baseline engagement. This attenuation means that users who systematically share cued content tend to differ from those who do not in ways that generate higher retweet counts independently of SAC presence. This between-user component is part of the diffusion process we quantify in the main model specification (population-average estimate), but is absorbed by the random intercept.

Study A.2 Consistent with the main text analysis, SACs are associated with a higher retweet likelihood for URLs rated as false, and a slightly lower likelihood of sharing for true. The results are also consistent when considering domain level veracity measures: lower factuality and veracity are associated with a higher increase due to the presence of SAC (Table A22).

A similar pattern emerges for *Overall Spread* (Table A25): URLs with a true fact-check rating and a SAC exhibit less overall diffusion than those without, whereas URLs with a false rating are significantly boosted. Consistent with the main study, the effect of SACs increases as the factuality of the domain decreases. Finally, for credibility scores, URLs from high credibility domains show a positive SAC effect, while for low credibility domains the effect is not significant, likely because domains with low credibility scores tend to generate the least overall spread.

In the CSE estimates (Table A28), the effect remains statistically significant for all key subsets except for domains from High credibility domains, suggesting that users within this category exhibit the most variance in expected retweets. Indeed, the top 3 users in the High Credibility subset accounts for 3.6% of tweets, compared to 0.7% for the Low Credibility category. The ME variant (Table A31)shows the same statistically significant patterns observed by the main model (stronger effect for False, Low Factuality and Low Credibility), even when adjusted for between-user variance.

Study A.3 Regarding political orientation, we also observe a pattern consistent with the main analysis: while the presence of SACs increases the likelihood of retweeting for content shared by users across the spectrum, the effect is stronger among right-leaning users. For all veracity measures, SACs are associated with a higher increase of retweet

likelihoods as veracity decreases (Table A23).

For the Overall Spread (Table A26), we assign a political-leaning label to each URL by averaging the estimated political scores of all users who posted it from our data. We consider a URL to be *left-leaning* if its average poster score is negative, and right-leaning if it is positive (as political scores range from -1 for extreme left to +1 for extreme right). Overall, we observe patterns consistent with those reported in the main paper: left-leaning URLs exhibit reduced overall spread when containing a SAC, whereas right-leaning URLs are significantly boosted. At the domain level, credibility scores do not yield statistically significant differences, but factuality labels show the same pattern as before, with low-factuality domains receiving the strongest boost from SACs.

For the CSE variant of the NegBin model, the effect is only significant for Right-leaning users, meaning that the (weak) effect observed on Left-leaning users was driven by specific (highly-active) users (Table A29). The Mixed-Effects estimates (Table A32) show a significant effect for nearly all categories, which means that even on the user-level, the presence of SACs is particularly associated with Right-leaning users and for low veracity/credibility content.

The consistency of the results across all three modeling approaches (*Retweet Count* with its three variants, *Retweet Likelihood*, and *Overall Spread*) and across all annotator agreement thresholds (*Unanimity*, *Majority*, and *Any*) provides strong evidence of the robustness of the effect. Regardless of modeling choice or labeling strictness, the estimated effect remains comparable across all the main dimensions of analysis.

Political orientation alternative measures

Beyond the political orientation inferred from followed political-elite accounts which we use in the main paper (Study A.3), we conduct two robustness checks to further assess the relationship between SACs and political identity. Specifically, we rely on the following alternative measures:

- **Self-identification**: Using users' profile bios, we leverage an instruction-finetuned language model to classify each user into one of four categories: left, right, other, or none (see Prompt 2 in Appendix 4.5), and the results in Table A17.

- **Source Slant**: Using domain-level ratings provided by Media Bias/Fact Check Media Bias/Fact Check, 2021, we substitute the user's political leaning by the political bias label of the domains they share. We group "Left" and "Center-Left" domains into a unified *left* category, and likewise merge "Right" and "Center-Right" into a *right* category (see results in Table A18).

Overall, we observe nearly identical patterns when using these alternative methods to assess users' political leaning: cued tweets receive more retweets than uncued ones for right-leaning users, and when aggregated by veracity, cued high veracity content shows a weak or negative effect for left-leaning users, whereas low veracity content is consistently boosted by the presence of SACs for both users and domains of both leanings.

The consistency of these results demonstrates that our main findings are robust to different operationalizations of political identity, reinforcing the conclusion that the boosting effect of SACs is indeed moderated by political identity.

Year fixed effects

To assess whether the SAC effect reflects temporal trends in scientific relevance or platform-wide engagement over our 2010–2023 observation window, we re-estimate all main models including year fixed effects. For brevity, Table A14 reports only the SAC effect estimates, as year dummies capture baseline differences in retweet counts across years rather than the effect of interest. The direction and pattern of all main findings are preserved: SACs increase overall content spread, with stronger amplification for low-veracity content and among right-leaning users. Point estimates are attenuated relative to the main specification, which is expected given that year fixed effects absorb a portion of the SAC effect that is temporally concentrated, most notably the surge in scientific discourse surrounding COVID-19, during which both the prevalence of scientific authority mentions and platform-wide engagement increased simultaneously. That the effect persists confirms that SAC amplification is not reducible to a COVID-era phenomenon.

Bot sensitivity adjustment

To assess whether automated or bot-like accounts account for the observed SAC effects, we re-estimate all main models including each user's Botometer score as a continuous covariate. Results are reported in Table A15. The direction and pattern of all main findings are preserved across all specifications, and point estimates are consistent with the main findings, and in many cases make the asymmetries (for veracity and partisanship) more pronounced. This pattern is the opposite of what one would expect if bot activity were inflating our results: controlling for the degree of automated behavior either leaves the SAC effect unchanged or amplifies it, suggesting that bots do not drive the observed associations and that the effect is concentrated among human users.

Further corroboration comes from a targeted analysis of retweeters, whose behavior directly constitutes the diffusion outcome we model. Since annotating all retweeters is prohibitively expensive at the scale of our dataset, we retrieved Botometer scores for a stratified subsample (N=16,000) covering all relevant subgroups (see Table A35). Bot score distributions and the proportion of likely automated accounts (Botometer score > 0.5) were strikingly consistent across subgroups, with potential bots representing approximately 3% of retweeters in all subsets. This uniformity confirms that differential bot presence across subgroups cannot account for the overall effect or the partisan and veracity asymmetries we report.

These results provide strong evidence that coordinated or automated activity does not account for our main findings.

Scientific Authority Attribution only

To assess whether the SAC effect is driven specifically by titles that attribute claims to a scientific entity (SAA) – as manipulated in the experiment – we re-estimate all main models restricting cues to titles classified as claims by the LLM, as described in the annotation section. Results are reported in Table A19. The overall effect and its veracity and partisan asymmetries are preserved and statistically significant across nearly all subsets (with the exception of the domain factuality analysis where the effect reverses). The attenuation of point estimates relative to the main SAC models further suggests that

SAMs (non-attribution mentions) independently contribute to the overall diffusion effect, a pattern consistent with the experimental results (see Table A49).

Inclusion of commercial entities with scientific activity

Our main analyses exclude entities with strong commercial activity (e.g., pharmaceutical companies) from the SAC definition, on the grounds that their authority signals may operate through different mechanisms than those of academic or institutional science. To assess the impact of this decision, we re-estimate all main models extending the SAC definition to include commercial entities. Results are reported in Table A20. The overall SAC effect increases substantially (IRR $>$ 2 in several subsets), with stronger amplification for low-veracity content and among right-leaning users, and all asymmetries more pronounced than in the main specification. This suggests that commercial entities with scientific activity play a similar role to SACs in amplifying content diffusion, and that their exclusion leads to conservative estimates of the overall effect. That said, the mechanism through which commercial entities confer credibility may differ from the credibility-based heuristic tested in the experiment. We retain the exclusion in the main analyses to maintain conceptual precision, but consider commercial entities with scientific activity as a theoretically distinct (and empirically relevant) class of authority cues warranting future investigation.

Appendix 4.5: Observational study annotation prompts

Prompt 1: Prompt used for Scientific Authority Cues extraction.

```
# Task Description:
You will be given titles and headlines from news articles shared
   ↪ online.
Your goal is to extract any mentions of words or names implying **
   ↪ academic or scientific authority**.

A mention can fall into one of the following categories:

   1) Individual: Specific named persons who can be considered a
      ↪ scientific authority, e.g. "Dr. Fauci", "Marie Curie".
   2) Institution: Organizations such as universities, public
      ↪ research centers, and space agencies, e.g. "Oxford", "
      ↪ NASA".
   3) Company: Private commercial entities directly involved in
      ↪ scientific research or industry, e.g. "Pfizer", "Moderna
      ↪ ".
   4) Role: Unnamed individuals that indicate a scientific
      ↪ authority (e.g. professions, titles), e.g. "Doctor", "
      ↪ scientists".
   5) Other: Words or phrases strongly suggesting a scientific
      ↪ source or authority, e.g. "study", "research", "science",
      ↪  "data".

For each given title in the list, output a valid JSON list
   ↪ containing:
- "mention": The extracted term as it appears in the title.
- "category": One of the five categories above.

# Processing Instructions:
- Only include mentions of scientific and academic authority.
```

```
  - Exclude general authority figures such as "officials," "
    ↪ government leaders," or "CEOs."
- Ensure correct mention extraction and categorization:
  - If the category is Individual, extract both first and last
    ↪ names (if both are mentioned).
  - If the mention includes a title (e.g. "Dr.", "MD"), keep it.
  - If the mention includes an adjective or a modifier, keep only
    ↪ the mention e.g. "Study" instead of "New Study" and "
    ↪ Scientists" instead of "German Scientists".
  - If the category is Role, make sure the context implies
    ↪ scientific authority. For example, "experts" in military
    ↪ experts" do not count as scientific authority.
  - If the category is Other, make sure the context implies
    ↪ scientific authority. For example, "data" in "data storage
    ↪ " or "paper" in "paper printers" do not count as
    ↪ scientific authorities.
  - If an acronym is widely known as a scientific entity (e.g. "
    ↪ WHO" for the World Health Organization), include it.
- If the title contains no clear mention of any scientific entity,
   ↪  the "mentions" field should be an empty list.

# Example:

## Input:
[
  { "id": 0, "title": "Immunologist Bruce Beutler says the MRNA
    ↪ Vaccine is a mistake, as New Data Show Accumulation in
    ↪ Reproductive Organs" },
  { "id": 1, "title": "Fintan O'Toole: Trial runs for fascism are
    ↪ in full flow" },
  { "id": 2, "title": "Newly Released Transcripts Show Michael
    ↪ Flynn Betrayed the United States" },
  { "id": 3, "title": "Pfizer Vaccine Confirmed To Cause
    ↪ Neurodegenerative Diseases - Harvard Study" },
```

```
  { "id": 4, "title": "California doctors say they've seen more
     ↪ deaths from suicide than coronavirus since lockdowns" }
]

## Output:
[
  {
    "id": 0,
    "mentions": [{"mention": "Immunologist", "category": "Role"},
       ↪ {"mention": "Bruce Beutler", "category": "Individual"},
       ↪ {"mention": "Data", "category": "Other"}]
  },
  {
    "id": 1,
    "mentions": []
  },
  {
    "id": 2,
    "mentions": []
  },
  {
    "id": 3,
    "mentions": [{"mention": "Pfizer", "category": "Company"}, {"
       ↪ mention": "Harvard", "category": "Institution"}, {"
       ↪ mention": "Study", "category": "Other"}]
  },
  {
    "id": 4,
    "mentions": [{"mention": "doctors", "category": "Role"}]
  }
]

## Input:
```

Prompt 2: Prompt used for user bios partisanship classification.

```
Task: Analyze a Twitter bio to classify its political alignment.

Instructions:
1. Analyze the language, hashtags, and self-descriptions in the
   ↪ bio
2. Consider explicit political statements, affiliations, or
   ↪ ideological indicators
3. Look for implicit signals through commonly used political
   ↪ phrases or symbols
4. Classify the bio into exactly one of these categories:
   - Left: Clear indicators of left-wing, progressive, or liberal
      ↪ alignment
   - Right: Clear indicators of right-wing, conservative, or
      ↪ traditionalist alignment
   - Other: Political alignment present but not clearly Left or
      ↪ Right (e.g., libertarian, centrist)
   - None: No clear political alignment indicators
5. Even if the bio is in a different language, only output one of
   ↪ the previous labels.

Classification Guidelines:
- Only classify as Left/Right when there are clear, explicit
   ↪ indicators
- Avoid over-interpreting ambiguous statements
- If in doubt between political categories, use "Other"
- If in doubt between political and non-political, use "None"
- Only output Left, Right, Other, or None, followed by a period
   ↪ "."

Examples:
Bio:
"Progressive | BLM | She/Her | Fighting for social justice"
Classification:
```

```
Left.

Bio:
"America First | Constitutional Conservative | 2A supporter"
Classification:
Right.

Bio: "Politics junkie | Independent thinker | Both sides have
   ↪ points"
Classification:
Other.

Bio: "Dog lover | Coffee addict | Travel enthusiast"
Classification:
None.

Bio:
```

Prompt 3: Prompt used for classifying SACs into SAA / SAM.

```
Task: You are an annotator for a study on how scientific authority
   ↪  affects information spread online.

Your task is to analyze each title and a list of scientific
   ↪ entities, and determine if the title corresponds to a claim
   ↪ made or reported by any of the entities, or they were merely
   ↪  mentioned in the title.

For each title, return a JSON object with the following fields:
- "id": the id of the title
- "contains_claim": true if the title contains a claim that is
   ↪ stated by any of the entities in the list (regardless of
   ↪ whether the claim is correct or not); false all entities are
   ↪  merely mentioned.
- "claim": If the title contains a claim, extract it without
```

```
   ↪ mentioning the scientific authority.

Examples:

## Input:
[
    {
        "id": 1,
        "title": "CDC Director Julie Gerberding Admits Vaccines
           ↪ Can Cause Autism-Like Symptoms",
        "scientific_authorities": ["CDC Director", "Julie
           ↪ Gerberding"],
    },
    {
        "id": 2,
        "title": "Interview with Dr. Byram Bridle",
        "scientific_authorities": ["Dr. Byram Bridle"],

    },
    {
        "id": 3,
        "title": "According to a new study, herpes infection
           ↪ possibly linked to COVID-19 vaccine",
        "scientific_authorities": ["study"],
    },
    {
        "id": 4,
        "title": "Joe Rogan: Dr. Pierre Kory Treated 200 Members
           ↪ of Congress With Ivermectin",
        "scientific_authorities": ["Dr. Pierre Kory"],
    },
    {
        "id": 5,
        "title": "Chinese Doctors Confirmed African Blood Genetic
```

```
            ↪ Composition Resist Coronavirus After Student Cured",
        "scientific_authorities": ["Doctors"],
    }
]

## Outputs:
[
    {
        "id": 1,
        "contains_claim": true,
        "claim": "Vaccines can cause autism-like symptoms"
    },
    {
        "id": 2,
        "contains_claim": false,
        "claim": ""
    },
    {
        "id": 3,
        "contains_claim": true,
        "claim": "herpes infection possibly linked to COVID-19
            ↪ vaccine"
    },
    {
        "id": 4,
        "contains_claim": false,
        "claim": "",
    },
    {
        "id": 5,
        "contains_claim": true,
        "claim": "African blood genetic composition resists
            ↪ coronavirus",
    },
```

```
]

## Input:
"""
```

Prompt 4: Prompt used for to identify claim-like titles in the absence of SACs.

```
You are an annotator for a study on how scientific authority
   ↪ affects information spread online.

Your task is to analyze titles of articles shared online, and
   ↪ determine if the title corresponds to a claim that could be
   ↪ made by or attributed to a scientific authority or not, i.e.
   ↪  the title can be rewritten as "A study shows that <title>"
   ↪ or "According to <scientific authority>, <title>"

For each title, return a JSON object with the following fields:
- "id": the id of the title
- "contains_claim": true if the title contains a claim (regardless
   ↪  of whether the claim is scientifically correct or not);
   ↪ false otherwise.
- "claim": If the title contains a claim, extract it.

Examples:

## Input:
[
    {
        "id": 1,
        "title": "Gaetz calls for DOJ to ban Chinese-made drones
           ↪ used to enforce social distancing guidelines",
    },
    {
        "id": 2,
        "title": "The BBC said there is HIV in the Covid-19
```

```
            ↪ vaccines.",
    },
    {
        "id": 3,
        "title": "Pfizer Covid vaccine side effects: More than one
            ↪  in 10 people suffer joint pain",
    },
    {
        "id": 4,
        "title": "Cat infected with coronavirus: 'One should wash
            ↪ hands before contact with one's pet'",
    },
    {
        "id": 5,
        "title": "Putin has banned the Rothschilds from Russia",
    }
]

## Outputs:
[
    {
        "id": 1,
        "contains_claim": false,
        "claim": ""
    },
    {
        "id": 2,
        "contains_claim": true,
        "claim": "There is HIV in the Covid-19 vaccines."
    },
    {
        "id": 3,
        "contains_claim": true,
        "claim": "Pfizer Covid vaccine causes joint pain."
```

```
    },
    {
        "id": 4,
        "contains_claim": true,
        "claim": "One should wash hands before contact with one's
            ↪ pet",
    },
    {
        "id": 5,
        "contains_claim": false,
        "claim": "",
    },
]

## Input:
"""
```

Appendix 4.6: Experimental study detailed analysis plan and hypotheses

In this section, we provide an extended explanation of the experimental methods. The analysis proceeded in four steps. First, we addressed H1, predicting that claims attributed to scientific entities would be more likely to be shared than those without attribution, and H2, expecting that claims attributed to named scientific entities would be shared more than those attributed to anonymous entities. These hypotheses were tested using two robust linear regressions. In the first model, the "no attribution" condition served as the reference category; in the second model, the "anonymous" attribution condition served as the reference category.

Second, we tested H3, which predicted that the effect of SAA (Scientific Authority Attribution) on sharing intention would be stronger for false than for true claims, and H4, proposing an interaction between attribution and participants' party affiliation. In the robust linear model testing H3, "true" claims served as the reference category, whereas in the model testing H4, "Republicans" served as the reference category. We then tested H5, hypothesizing that the effect of attribution on sharing intention would differ jointly by claim veracity and party affiliation. For this test, we fitted three robust linear models: one that included only direct effects, one that included two-way interactions, and another that included the three-way interaction effect.

Third, we examined the role of perceived accuracy, trust, and familiarity in shaping sharing intention. We tested H6, expecting that SAA would increase perceived accuracy using a robust linear model with "no attribution" as the reference category, and explored whether this effect differed for false versus true claims. To test H7, which hypothesized that perceived accuracy mediates the effect of attribution on sharing intention, we fitted a structural equation model (SEM) to estimate the indirect path from attribution to sharing intention via perceived accuracy.

Finally, for claims in the attribution condition, we explored the roles of trust in and familiarity with scientific authorities by fitting an SEM with these variables as predictors of both sharing intention and perceived accuracy, including indirect effects. As a complementary analysis, we examined whether attribution to a trusted scientific author-

ity moderated party and claim veracity differences in sharing intention and perceived accuracy by fitting robust linear models.

In all models, the experimental manipulation (no attribution vs. attribution), claim veracity (false vs. true), and party affiliation (Democrat vs. Republican) were operationalized as dummy variables, coded 0 (reference category) or 1 (comparison category). Sharing intention, perceived accuracy, trust, and familiarity were all measured on a 5-point Likert scale from 0 ("Not at all") to 4 ("Extremely"), and treated as continuous. In the additional analyses, we created a dummy measure of trusted authority, coded as 1 when participants reported being "Very" or "Extremely" trusting of an authority and 0 when they reported "Not at all", "Not very", and "Somewhat". All robust linear models and SEMs included participant-clustered standard errors to account for our repeated measures.

To illustrate mean differences relevant to our hypotheses, we report pairwise comparisons of key conditions in the Tables A36 and A37. For analyses involving interaction terms (H3–H5, exploratory analyses of differences in perceived accuracy, trusted entity × party affiliation, and trusted authority × fact-checked veracity) or mediation paths (H7, exploratory analysis of trust and familiarity), we assess the added explanatory power by comparing models with and without these terms (shown in the corresponding tables). Robustness checks are provided by control models for hypothesis tests (Tables A43 and A44) and complementary analyses (A45), as well as ordinal logistic regression models (Tables A46, A47, and A48).

Appendix 4.7: Experimental study stimuli explanation

In this section, we briefly describe the experimental stimuli. The stimuli pool comprised 24 headlines, of which each participant was randomly shown 12. For *claims* (manipulated, see Table A36), each participant saw a randomly assigned version of the 16 possible headline. For *news* (not manipulated, see Table A37), there was only one version of the 8 possible headlines.

Appendix 4.8: Experimental study extended results

In this section, we present an extended version of our experimental result. For the parameters of the models testing our hypotheses, see Tables A38– A40.

Contrasts

Effects are reported as estimated group contrasts derived from model parameters, with 95% CIs obtained from two-tailed tests. These contrasts were calculated using the *emmeans* and *contrast* functions from the *emmeans* R-package, and represent differences in the estimated means of the groups relevant to our hypotheses (Table A38) and exploratory tests (Tables A39 for sharing intention, and A40 for perceived accuracy).

Structural equation models

We report the parameter estimates of the SEMs testing H7 and exploring the roles of trust in and familiarity with an attributed scientific entity on sharing intention (Table A40). Additionally, we present the path diagram of the latter relationships in Figure A1. This model shows that trust in the attributed entity increases perceived accuracy ($b = 0.118$, CI = [0.072, 0.164], $p < 0.001$) and sharing intention ($b = 0.085$, CI = [0.041, 0.130], $p < 0.001$). Familiarity with the attributed entity does not increase perceived accuracy ($p = 0.692$), but it does increase sharing intention ($b = 0.071$, CI = [0.035, 0.108], $p < 0.001$).

Trusted entities, veracity, and political affiliation

In this subsection, we show the results of the additional analysis of the dummy measure of trusted authority. In Figure A2 we see that the increase in sharing intentions for trusted versus non-trusted entities is significant for both false ($b = 0.127$, CI = [0.015, 0.239], $p = 0.027$) and true claims ($b = 0.506$, CI = [0.383, 0.628], $p < 0.001$). Also, the increase is significant for perceived accuracy, but only for true claims. Additionally, the parameter estimates of the models used to produce the group means displayed in Figure A2 are shown in Table A42.

Appendix 4.9: Experimental study robustness checks

In this section we present two types of robustness checks for our analyses. First, we add control variables to all robust linear regression models in Tables A43, A44, and A45. Second, we specify ordinal logistic regression models in Tables A46, A47 and A48.

We note that the effect of attribution to a trusted entity on *sharing intention* is more pronounced for Republicans than Democrats (Table A42) but this interaction effect disappears once control variables are added to the model (Table A45). However, according to the ordinal logistic regression model, this same effect is significant even when control variables are added to the model (Table A48). Given the ambiguity of the effect, we decided not to report it in the Results section.

Appendix 4.10: Experimental study political affiliation differences in trust

In this section, we examine differences in trust between Democrats and Republicans for each scientific entity evaluated in our experiment. We dichotomized trust by coding responses of "Extremely" or "Very" trustworthy as 1, and all other responses as 0. For each scientific entity, we then calculated the difference by subtracting the proportion of Democrats who indicated trust from the proportion of Republicans who did so as well. The results of this descriptive analysis are shown in Figure A3. As shown, most entities are more trusted by Democrats than by Republicans.

Appendix 4.11: Experimental study analysis of Scientific Authority Mentions (SAMs)

In this section, we provide a complementary interpretation of the analyses of the effect of Scientific Authority Mentions (SAMs) in *news* (headlines without manipulation). We find that mentioning a scientific entity in a headline (without attribution) increases its sharing intention (Table A49). This effect is larger than the effect of attributing a *claim* to a scientific entity (SAA). However, we note that this comparison should be interpreted cautiously, as the analyses of SAA are nested within headlines and experimentally manipulated, whereas the analysis of SAMs is *between* headlines that contain a mention and those that do not.

Appendix 4.12: Experimental study figures

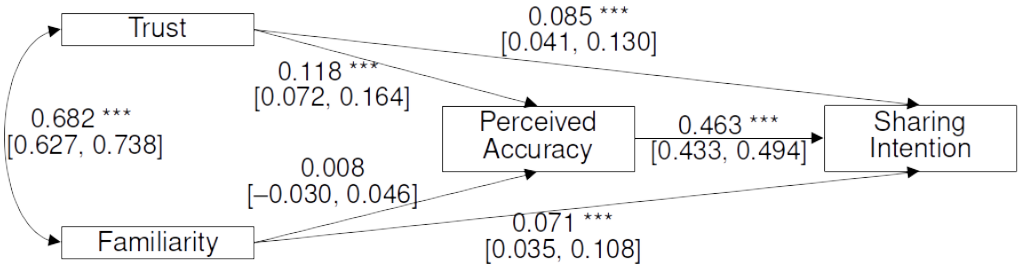


Figure A1: Trust in and familiarity with an attributed entity affect sharing intention. Unstandardized model parameters from a structural equation model (N = 1,182) with participant-clustered standard errors with CIs in brackets. *** $p < 0.001$.

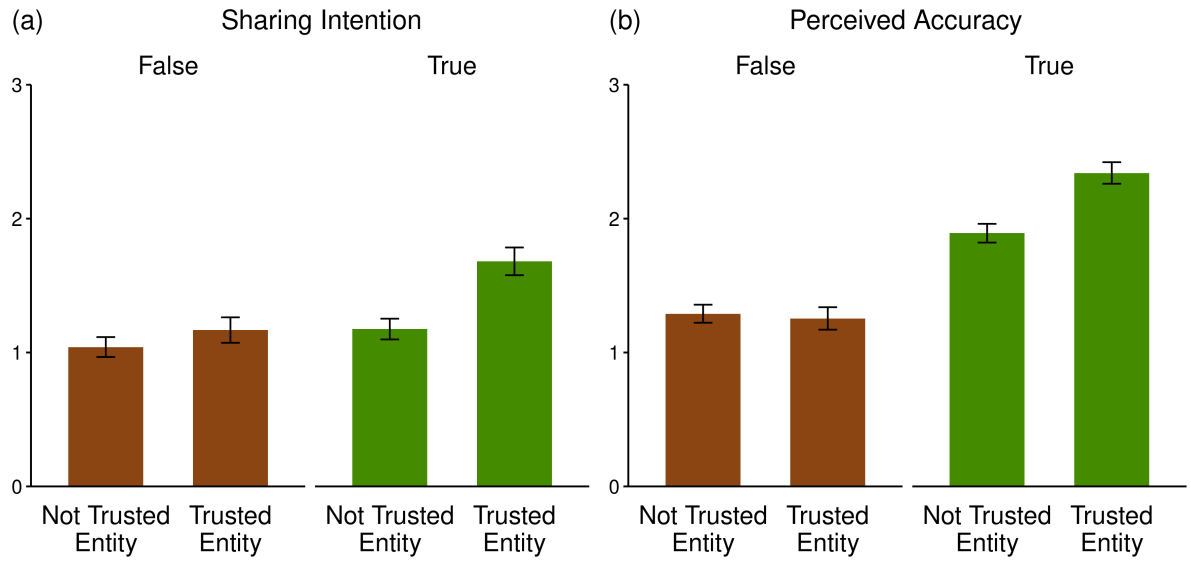


Figure A2: Attributing claims to trusted entities enhances both sharing intention and perceived accuracy. Estimated means (95% CI) from robust linear regression models (N = 1,182) with participant-clustered standard errors.

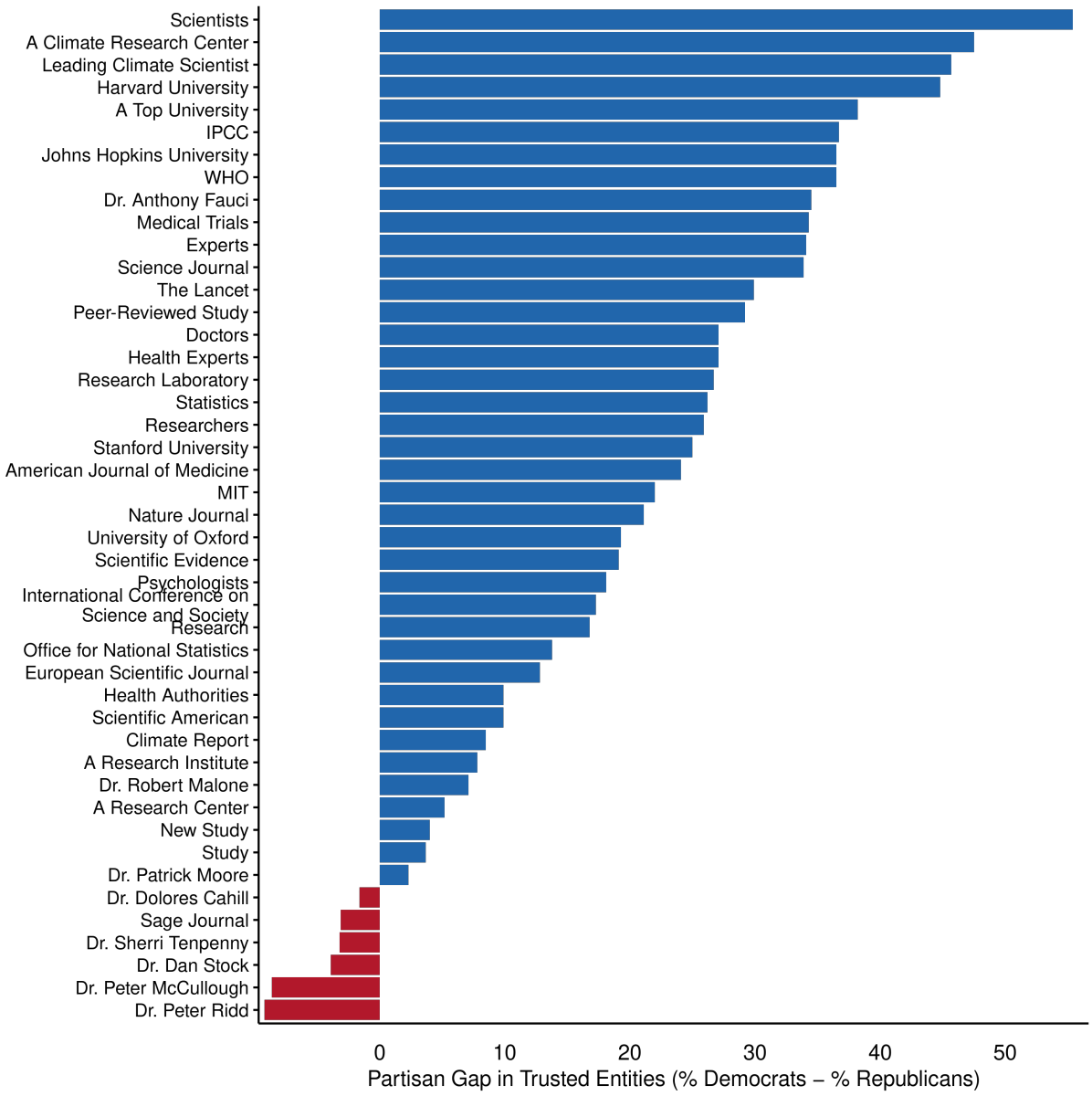


Figure A3: Partisan differences in trust across entities. Positive values indicate a higher trust percentage among Democrats, while negative values indicate a higher percentage among Republicans.

Appendix 4.13: Observational study tables

Tables A1 to A13 report the full parameters of negative binomial regression models estimating the effect of Scientific Authority Cues ("Cued") on retweet counts for three levels of annotator agreement (*Unanimity*, *Majority*, and *Any*). All models include the log of the author's follower count as a covariate (with extra covariates indicated when present). Reported values are exponentiated to obtain Incidence Rate Ratios (IRRs), accompanied by 95% confidence intervals and p-values, where IRRs greater than 1 indicate an increase in expected retweet counts. For each analysis, the corresponding subset size is reported, reflecting the portion of the data for which the relevant property is true. Tables A14 to A32 contain robustness check results for the main effect, and report only the estimates for the main effect (*Cued*) for clarity. Underlined parameters denote the specific estimates reported in the main text.

Table A1: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts for the overall dataset.

Agreement	**Parameter**	**IRR**	**CI**	***p***
Unanimity	**Cued**	1.714	[1.701, 1.726]	< 0.001
	Followers (log)	2.084	[2.082, 2.086]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Majority	**Cued**	1.723	[1.711, 1.736]	< 0.001
	Followers (log)	2.085	[2.082, 2.087]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Any	**Cued**	1.647	[1.636, 1.658]	< 0.001
	Followers (log)	2.086	[2.083, 2.088]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001

Table A2: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts for tweets linking to URLs from the *Condor* database.

Agreement	**Parameter**	**IRR**	**CI**	***p***
Unanimity	**Cued**	1.745	[1.732, 1.758]	< 0.001
	Followers (log)	2.080	[2.078, 2.083]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Majority	**Cued**	1.756	[1.743, 1.769]	< 0.001
	Followers (log)	2.081	[2.078, 2.083]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Any	**Cued**	1.670	[1.659, 1.683]	< 0.001
	Followers (log)	2.082	[2.079, 2.084]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001

Table A3: Negative binomial regression estimates of the effect of Scientific Authority Cues on retweet counts for tweets linking to URLs from the *Science Feedback* database.

Agreement	**Parameter**	**IRR**	**CI**	***p***
Unanimity	**Cued**	1.550	[1.532, 1.569]	< 0.001
	Followers (log)	2.067	[2.062, 2.072]	< 0.001
	Intercept	0.005	[0.005, 0.005]	< 0.001
Majority	**Cued**	1.537	[1.519, 1.555]	< 0.001
	Followers (log)	2.067	[2.062, 2.072]	< 0.001
	Intercept	0.005	[0.005, 0.005]	< 0.001
Any	**Cued**	1.508	[1.491, 1.526]	< 0.001
	Followers (log)	2.067	[2.062, 2.072]	< 0.001
	Intercept	0.005	[0.005, 0.005]	< 0.001

Table A4: Negative binomial regression estimates of the effect of *named* versus *unnamed* Scientific Authority Cues (SACs) on retweet counts.

Agreement	**Parameter**	**IRR**	**CI**	***p***
Unanimity	**Cued (named)**	1.727	[1.714, 1.741]	< 0.001
	Cued (unnamed)	1.659	[1.634, 1.684]	< 0.001
	Followers (log)	2.084	[2.082, 2.086]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Majority	**Cued (named)**	1.745	[1.731, 1.758]	< 0.001
	Cued (unnamed)	1.634	[1.609, 1.658]	< 0.001
	Followers (log)	2.084	[2.082, 2.087]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Any	**Cued (named)**	1.653	[1.641, 1.665]	< 0.001
	Cued (unnamed)	1.618	[1.594, 1.642]	< 0.001
	Followers (log)	2.086	[2.083, 2.088]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001

Table A5: Negative binomial regression estimates of the effect Scientific Authority Cues (SACs) on retweet counts, by entity type (*institution*, *institution*, and *source*).

Agreement	**Parameter**	**IRR**	**CI**	***p***
Unanimity	**Cued (Individual)**	1.667	[1.651, 1.682]	< 0.001
	Cued (Institution)	1.446	[1.428, 1.463]	< 0.001
	Cued (Source)	1.403	[1.385, 1.420]	< 0.001
	Followers (log)	2.083	[2.081, 2.085]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Majority	**Cued (Individual)**	1.662	[1.646, 1.677]	< 0.001
	Cued (Institution)	1.475	[1.458, 1.493]	< 0.001
	Cued (Source)	1.397	[1.380, 1.415]	< 0.001
	Followers (log)	2.083	[2.081, 2.086]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001
Any	**Cued (Individual)**	1.638	[1.624, 1.652]	< 0.001
	Cued (Institution)	1.308	[1.293, 1.322]	< 0.001
	Cued (Source)	1.431	[1.414, 1.449]	< 0.001
	Followers (log)	2.084	[2.082, 2.087]	< 0.001
	Intercept	0.004	[0.004, 0.004]	< 0.001

Table A6: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by the fact-check rating of the shared URL (*True*, *False*, *Neutral*, and *Other*).

Fact-check Rating	**Subset Size**	**Agreement**	**Parameter**	**IRR**	**CI**	***p***
True	$N = 703{,}732$	Unanimity	**Cued**	0.707	[0.686, 0.729]	< 0.001
			Followers (log)	2.122	[2.115, 2.130]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
		Majority	**Cued**	0.754	[0.732, 0.777]	< 0.001
			Followers (log)	2.122	[2.115, 2.130]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
		Any	**Cued**	0.886	[0.861, 0.911]	< 0.001
			Followers (log)	2.121	[2.114, 2.129]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
False	$N = 3{,}873{,}614$	Unanimity	**Cued**	1.468	[1.455, 1.481]	< 0.001
			Followers (log)	2.054	[2.050, 2.057]	< 0.001
			Intercept	0.006	[0.006, 0.006]	< 0.001
		Majority	**Cued**	1.487	[1.474, 1.500]	< 0.001
			Followers (log)	2.054	[2.051, 2.058]	< 0.001
			Intercept	0.006	[0.006, 0.006]	< 0.001
		Any	**Cued**	1.385	[1.373, 1.397]	< 0.001
			Followers (log)	2.055	[2.051, 2.058]	< 0.001
			Intercept	0.006	[0.006, 0.006]	< 0.001
Neutral	$N = 3{,}746{,}129$	Unanimity	**Cued**	1.812	[1.782, 1.841]	< 0.001
			Followers (log)	2.144	[2.141, 2.148]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
		Majority	**Cued**	1.775	[1.747, 1.803]	< 0.001
			Followers (log)	2.144	[2.140, 2.148]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
		Any	**Cued**	1.732	[1.708, 1.757]	< 0.001
			Followers (log)	2.145	[2.141, 2.149]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
Other	$N = 405{,}857$	Unanimity	**Cued**	0.749	[0.714, 0.786]	< 0.001
			Followers (log)	2.092	[2.081, 2.103]	< 0.001
			Intercept	0.004	[0.004, 0.005]	< 0.001
		Majority	**Cued**	0.718	[0.685, 0.752]	< 0.001
			Followers (log)	2.093	[2.082, 2.103]	< 0.001
			Intercept	0.004	[0.004, 0.005]	< 0.001
		Any	**Cued**	1.065	[1.020, 1.111]	0.004
			Followers (log)	2.091	[2.081, 2.102]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001

Table A7: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by the factuality level of the shared URL's domain (*High*, *Mid*, and *Low*).

Factuality Label	**Subset Size**	**Agreement**	**Parameter**	**IRR**	**CI**	***p***
High	$N = 791{,}477$	Unanimity	**Cued**	1.183	[1.157, 1.210]	< 0.001
			Followers (log)	2.124	[2.117, 2.131]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Majority	**Cued**	1.187	[1.161, 1.213]	< 0.001
			Followers (log)	2.124	[2.117, 2.131]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Any	**Cued**	1.166	[1.142, 1.191]	< 0.001
			Followers (log)	2.124	[2.117, 2.131]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
Mid	$N = 1{,}389{,}834$	Unanimity	**Cued**	1.456	[1.432, 1.481]	< 0.001
			Followers (log)	2.15	[2.145, 2.156]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Majority	**Cued**	1.449	[1.425, 1.473]	< 0.001
			Followers (log)	2.15	[2.144, 2.155]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Any	**Cued**	1.448	[1.425, 1.471]	< 0.001
			Followers (log)	2.152	[2.147, 2.158]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
Low	$N = 1{,}213{,}228$	Unanimity	**Cued**	1.971	[1.944, 1.998]	< 0.001
			Followers (log)	2.046	[2.040, 2.051]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Majority	**Cued**	1.948	[1.922, 1.975]	< 0.001
			Followers (log)	2.045	[2.040, 2.051]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Any	**Cued**	1.935	[1.909, 1.961]	< 0.001
			Followers (log)	2.046	[2.040, 2.051]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001

Table A8: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by the credibility score tercile of the shared URL's domain (*High*, *Mid*, and *Low*).

Credibility Score	**Subset Size**	**Agreement**	**Parameter**	**IRR**	**CI**	***p***
High	N = 1,572,942	Unanimity	**Cued**	1.304	[1.283, 1.326]	< 0.001
			Followers (log)	2.062	[2.057, 2.066]	< 0.001
			Intercept	0.004	[0.004, 0.005]	< 0.001
		Majority	**Cued**	1.309	[1.288, 1.330]	< 0.001
			Followers (log)	2.062	[2.057, 2.066]	< 0.001
			Intercept	0.004	[0.004, 0.005]	< 0.001
		Any	**Cued**	1.24	[1.221, 1.259]	< 0.001
			Followers (log)	2.061	[2.056, 2.066]	< 0.001
			Intercept	0.005	[0.004, 0.005]	< 0.001
Mid	N = 1,355,979	Unanimity	**Cued**	1.491	[1.465, 1.517]	< 0.001
			Followers (log)	2.164	[2.158, 2.170]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Majority	**Cued**	1.487	[1.462, 1.513]	< 0.001
			Followers (log)	2.164	[2.158, 2.170]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Any	**Cued**	1.199	[1.180, 1.218]	< 0.001
			Followers (log)	2.164	[2.158, 2.170]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
Low	N = 1,790,094	Unanimity	**Cued**	1.637	[1.616, 1.66]	< 0.001
			Followers (log)	2.035	[2.030, 2.040]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Majority	**Cued**	1.679	[1.656, 1.701]	< 0.001
			Followers (log)	2.037	[2.032, 2.042]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Any	**Cued**	1.739	[1.717, 1.762]	< 0.001
			Followers (log)	2.038	[2.033, 2.043]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001

Table A9: Negative binomial regression models estimating the overall effect of the interaction between the presence of Scientific Authority Cues (SACs) and credibility scores on retweet counts.

Agreement	**Parameter**	b	**CI**	**p**
Unanimity	Credibility Score	−0.097	[−0.107, −0.087]	< 0.001
	Cued	0.717	[0.677, 0.757]	< 0.001
	Cued x Credibility Score	−0.318	[−0.359, −0.278]	< 0.001
	Followers (log)	0.734	[0.732, 0.735]	< 0.001
	Intercept	−5.446	[−5.458, −5.435]	< 0.001
Majority	Credibility Score	−0.082	[−0.093, −0.072]	< 0.001
	Cued	0.892	[0.855, 0.930]	< 0.001
	Cued x Credibility Score	−0.492	[−0.530, −0.454]	< 0.001
	Followers (log)	0.734	[0.733, 0.736]	< 0.001
	Intercept	−5.462	[−5.474, −5.450]	< 0.001
Any	Credibility Score	−0.086	[−0.097, −0.076]	< 0.001
	Cued	0.901	[0.865, 0.938]	< 0.001
	Cued x Credibility Score	−0.567	[−0.604, −0.530]	< 0.001
	Followers (log)	0.735	[0.733, 0.736]	< 0.001
	Intercept	−5.46	[−5.472, −5.449]	< 0.001

Table A10: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, based on the inferred political orientation (partisanship) of the tweet poster.

Partisanship	Subset Size	Agreement	Parameter	IRR	CI	p
Left-Leaning	$N = 1{,}551{,}200$	Unanimity	**Cued**	1.068	[1.047, 1.089]	< 0.001
			Followers (log)	2.315	[2.309, 2.322]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Majority	**Cued**	1.06	[1.039, 1.080]	< 0.001
			Followers (log)	2.315	[2.309, 2.322]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Any	**Cued**	1.023	[1.005, 1.041]	0.012
			Followers (log)	2.315	[2.309, 2.322]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
Right-Leaning	$N = 2{,}383{,}478$	Unanimity	**Cued**	1.731	[1.711, 1.751]	< 0.001
			Followers (log)	2.087	[2.082, 2.091]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
		Majority	**Cued**	1.724	[1.705, 1.743]	< 0.001
			Followers (log)	2.087	[2.082, 2.091]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
		Any	**Cued**	1.665	[1.647, 1.683]	< 0.001
			Followers (log)	2.085	[2.081, 2.090]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001

Table A11: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by inferred poster partisanship (Left, Right) and shared URL fact-check ratings (True, False)

Partisanship × URL rating	**Subset Size**	**Agreement**	**Parameter**	**IRR**	**CI**	***p***
Left × True	$N = 157{,}200$	Unanimity	**Cued**	0.870	[0.820, 0.924]	< 0.001
			Followers (log)	2.284	[2.265, 2.303]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
		Majority	**Cued**	0.877	[0.827, 0.931]	< 0.001
			Followers (log)	2.284	[2.265, 2.303]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
		Any	**Cued**	0.850	[0.803, 0.899]	< 0.001
			Followers (log)	2.284	[2.265, 2.303]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
Left × False	$N = 402{,}798$	Unanimity	**Cued**	1.099	[1.065, 1.135]	< 0.001
			Followers (log)	2.086	[2.075, 2.097]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Majority	**Cued**	1.097	[1.063, 1.132]	< 0.001
			Followers (log)	2.086	[2.075, 2.097]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
		Any	**Cued**	1.050	[1.020, 1.082]	0.001
			Followers (log)	2.086	[2.075, 2.098]	< 0.001
			Intercept	0.003	[0.003, 0.003]	< 0.001
Right × True	$N = 64{,}349$	Unanimity	**Cued**	0.964	[0.861, 1.079]	0.525
			Followers (log)	2.377	[2.344, 2.411]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Majority	**Cued**	1.172	[1.054, 1.303]	0.003
			Followers (log)	2.381	[2.348, 2.415]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Any	**Cued**	1.214	[1.102, 1.336]	< 0.001
			Followers (log)	2.384	[2.351, 2.418]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
Right × False	$N = 1{,}584{,}988$	Unanimity	**Cued**	1.517	[1.497, 1.536]	< 0.001
			Followers (log)	2.055	[2.049, 2.060]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Majority	**Cued**	1.510	[1.49, 1.529]	< 0.001
			Followers (log)	2.055	[2.050, 2.061]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Any	**Cued**	1.454	[1.436, 1.472]	< 0.001
			Followers (log)	2.054	[2.049, 2.060]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001

Table A12: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by inferred poster partisanship (Left, Right) and shared URL's domain factuality levels (High, Low)

Partisanship × Domain Factuality	**Subset Size**	**Agreement**	**Parameter**	**IRR**	**CI**	***p***
Left × High Factuality	$N = 320{,}921$	Unanimity	**Cued**	0.87	[0.835, 0.906]	< 0.001
			Followers (log)	2.362	[2.349, 2.376]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Majority	**Cued**	0.865	[0.831, 0.901]	< 0.001
			Followers (log)	2.363	[2.349, 2.376]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Any	**Cued**	0.857	[0.825, 0.890]	< 0.001
			Followers (log)	2.362	[2.349, 2.376]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
Left × Low Factuality	$N = 69{,}269$	Unanimity	**Cued**	2.067	[1.945, 2.196]	< 0.001
			Followers (log)	2.28	[2.251, 2.311]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Majority	**Cued**	2.023	[1.904, 2.149]	< 0.001
			Followers (log)	2.277	[2.248, 2.307]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Any	**Cued**	1.973	[1.858, 2.095]	< 0.001
			Followers (log)	2.278	[2.249, 2.309]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
Right × High Factuality	$N = 97{,}100$	Unanimity	**Cued**	1.767	[1.678, 1.861]	< 0.001
			Followers (log)	2.325	[2.300, 2.350]	< 0.001
			Intercept	0.002	[0.001, 0.002]	< 0.001
		Majority	**Cued**	1.744	[1.656, 1.836]	< 0.001
			Followers (log)	2.324	[2.299, 2.350]	< 0.001
			Intercept	0.002	[0.001, 0.002]	< 0.001
		Any	**Cued**	1.758	[1.672, 1.849]	< 0.001
			Followers (log)	2.327	[2.302, 2.352]	< 0.001
			Intercept	0.002	[0.001, 0.002]	< 0.001
Right × Low Factuality	$N = 720{,}597$	Unanimity	**Cued**	2.179	[2.141, 2.217]	< 0.001
			Followers (log)	2.035	[2.027, 2.042]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Majority	**Cued**	2.148	[2.111, 2.186]	< 0.001
			Followers (log)	2.035	[2.027, 2.042]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001
		Any	**Cued**	2.089	[2.053, 2.125]	< 0.001
			Followers (log)	2.032	[2.025, 2.040]	< 0.001
			Intercept	0.005	[0.005, 0.005]	< 0.001

Table A13: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by inferred poster partisanship (Left, Right) and shared URL's domain credibility score terciles (High, Low)

Partisanship × Domain Credibility	**Subset Size**	**Agreement**	**Parameter**	**IRR**	**CI**	***p***
Left × High Credibility	$N = 603{,}312$	Unanimity	**Cued**	0.931	[0.905, 0.957]	< 0.001
			Followers (log)	2.356	[2.346, 2.365]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Majority	**Cued**	0.926	[0.900, 0.952]	< 0.001
			Followers (log)	2.355	[2.346, 2.365]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
		Any	**Cued**	0.88	[0.858, 0.904]	< 0.001
			Followers (log)	2.355	[2.345, 2.365]	< 0.001
			Intercept	0.001	[0.001, 0.001]	< 0.001
Left × Low Credibility	$N = 198{,}868$	Unanimity	**Cued**	1.424	[1.340, 1.513]	< 0.001
			Followers (log)	1.898	[1.881, 1.915]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
		Majority	**Cued**	1.426	[1.345, 1.513]	< 0.001
			Followers (log)	1.896	[1.879, 1.913]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
		Any	**Cued**	1.446	[1.365, 1.530]	< 0.001
			Followers (log)	1.895	[1.878, 1.912]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001
Right × High Credibility	$N = 212{,}938$	Unanimity	**Cued**	1.668	[1.609, 1.729]	< 0.001
			Followers (log)	2.26	[2.244, 2.276]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
		Majority	**Cued**	1.667	[1.609, 1.729]	< 0.001
			Followers (log)	2.26	[2.244, 2.275]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
		Any	**Cued**	1.659	[1.602, 1.719]	< 0.001
			Followers (log)	2.262	[2.246, 2.278]	< 0.001
			Intercept	0.002	[0.002, 0.002]	< 0.001
Right × Low Credibility	$N = 720{,}597$	Unanimity	**Cued**	1.358	[1.334, 1.383]	< 0.001
			Followers (log)	2.043	[2.036, 2.050]	< 0.001
			Intercept	0.004	[0.004, 0.005]	< 0.001
		Majority	**Cued**	1.352	[1.328, 1.376]	< 0.001
			Followers (log)	2.044	[2.036, 2.051]	< 0.001
			Intercept	0.004	[0.004, 0.005]	< 0.001
		Any	**Cued**	1.376	[1.352, 1.400]	< 0.001
			Followers (log)	2.043	[2.036, 2.050]	< 0.001
			Intercept	0.004	[0.004, 0.004]	< 0.001

Table A14: Negative Binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts for the key findings reported in the paper. Models include year fixed effects (omitted for brevity) to account for temporal trends over the 2010–2023 observation window. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Condition	Value	Subset Size	Agreement	IRR	CI	p
Source Dataset	Condor	$N = 8{,}142{,}088$	Unanimity	1.209	[1.078, 1.355]	0.001
			Majority	1.223	[1.094, 1.368]	< 0.001
			Any	1.169	[1.052, 1.299]	0.004
	Science Feedback	$N = 1{,}903{,}546$	Unanimity	1.107	[1.031, 1.189]	0.005
			Majority	1.100	[1.024, 1.182]	0.009
			Any	1.090	[1.015, 1.171]	0.018
	Combined	$N = 8{,}729{,}332$	Unanimity	1.207	[1.198, 1.216]	< 0.001
			Majority	1.220	[1.211, 1.229]	< 0.001
			Any	1.172	[1.163, 1.180]	< 0.001
URL Fact-check	True	$N = 703{,}732$	Unanimity	0.673	[0.569, 0.797]	< 0.001
			Majority	0.709	[0.602, 0.835]	< 0.001
			Any	0.829	[0.692, 0.993]	0.042
	False	$N = 3{,}873{,}614$	Unanimity	1.121	[1.013, 1.240]	0.027
			Majority	1.143	[1.033, 1.265]	0.010
			Any	1.063	[0.959, 1.179]	0.244
Domain Factuality	High	$N = 791{,}477$	Unanimity	1.106	[0.893, 1.370]	0.356
			Majority	1.111	[0.900, 1.372]	0.328
			Any	1.101	[0.887, 1.366]	0.384
	Low	$N = 1{,}213{,}228$	Unanimity	1.480	[1.365, 1.605]	< 0.001
			Majority	1.471	[1.357, 1.594]	< 0.001
			Any	1.471	[1.358, 1.593]	< 0.001
Credibility Score	High	$N = 1{,}572{,}942$	Unanimity	1.244	[0.852, 1.817]	0.258
			Majority	1.249	[0.866, 1.802]	0.235
			Any	1.183	[0.855, 1.638]	0.311
	Low	$N = 1{,}790{,}094$	Unanimity	1.164	[1.076, 1.259]	< 0.001
			Majority	1.207	[1.106, 1.318]	< 0.001
			Any	1.265	[1.161, 1.378]	< 0.001
User Leaning	Left	$N = 1{,}551{,}200$	Unanimity	0.855	[0.769, 0.951]	0.004
			Majority	0.850	[0.766, 0.944]	0.002
			Any	0.808	[0.732, 0.891]	< 0.001
	Right	$N = 2{,}383{,}478$	Unanimity	1.253	[1.171, 1.340]	< 0.001
			Majority	1.255	[1.174, 1.341]	< 0.001
			Any	1.233	[1.154, 1.318]	< 0.001

Table A15: Negative Binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts for the key findings. Models include Botometer's *bot score* as a continuous covariate to adjust for automated or bot-like account activity. Follower count and intercept, and estimates for the "Any" agreement level are omitted for clarity, and only *cued* indicator estimates are reported.

Condition	Value	Subset Size	Parameter	IRR	CI	p
Source Dataset	Combined	$N = 8{,}142{,}088$	Unanimity	1.709	[1.696, 1.721]	< 0.001
			Bot score	0.704	[0.693, 0.715]	< 0.001
			Majority	1.718	[1.705, 1.730]	< 0.001
			Bot score	0.707	[0.697, 0.718]	< 0.001
URL Fact-check	True	$N = 703{,}732$	Unanimity	0.705	[0.684, 0.727]	< 0.001
			Bot score	0.910	[0.866, 0.956]	< 0.001
			Majority	0.753	[0.731, 0.776]	< 0.001
			Bot score	0.911	[0.868, 0.957]	< 0.001
	False	$N = 3{,}873{,}614$	Unanimity	1.465	[1.452, 1.479]	< 0.001
			Bot score	0.747	[0.730, 0.765]	< 0.001
			Majority	1.484	[1.470, 1.497]	< 0.001
			Bot score	0.752	[0.734, 0.770]	< 0.001
Domain Factuality	High	$N = 791{,}477$	Unanimity	1.176	[1.151, 1.203]	< 0.001
			Bot score	0.702	[0.669, 0.737]	< 0.001
			Majority	1.180	[1.154, 1.206]	< 0.001
			Bot score	0.703	[0.670, 0.737]	< 0.001
	Low	$N = 1{,}213{,}228$	Unanimity	1.965	[1.938, 1.992]	< 0.001
			Bot score	0.795	[0.764, 0.827]	< 0.001
			Majority	1.943	[1.917, 1.969]	< 0.001
			Bot score	0.796	[0.765, 0.829]	< 0.001
Credibility Score	High	$N = 1{,}572{,}942$	Unanimity	1.301	[1.280, 1.323]	< 0.001
			Bot score	0.702	[0.679, 0.726]	< 0.001
			Majority	1.305	[1.284, 1.327]	< 0.001
			Bot score	0.702	[0.679, 0.727]	< 0.001
	Low	$N = 1{,}790{,}094$	Unanimity	1.620	[1.599, 1.642]	< 0.001
			Bot score	0.748	[0.724, 0.773]	< 0.001
			Majority	1.661	[1.639, 1.683]	< 0.001
			Bot score	0.761	[0.737, 0.786]	< 0.001
User Leaning	Left	$N = 1{,}551{,}200$	Unanimity	1.060	[1.039, 1.081]	< 0.001
			Bot score	0.654	[0.631, 0.679]	< 0.001
			Majority	1.052	[1.032, 1.073]	< 0.001
			Bot score	0.654	[0.630, 0.678]	< 0.001
	Right	$N = 2{,}383{,}478$	Unanimity	1.728	[1.708, 1.748]	< 0.001
			Bot score	0.680	[0.658, 0.702]	< 0.001
			Majority	1.721	[1.702, 1.740]	< 0.001
			Bot score	0.681	[0.660, 0.704]	< 0.001

Table A16: Negative Binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts on tweets with a *Neutral* rating from Condor (*Not rated* or *Not eligible*). Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Condition	Value	Subset Size	Parameter	IRR	CI	p
Source Dataset	Overall	N = 3,728,797	Unanimity	1.826	[1.796, 1.856]	< 0.001
			Majority	1.789	[1.761, 1.818]	< 0.001
			Any	1.750	[1.724, 1.775]	< 0.001
Domain Factuality	High	N = 509,137	Unanimity	1.103	[1.097, 1.164]	< 0.001
			Majority	1.119	[1.087, 1.152]	< 0.001
			Any	1.148	[1.117, 1.180]	< 0.001
	Mid	N = 784,872	Unanimity	2.169	[2.108, 2.232]	< 0.001
			Majority	2.101	[2.043, 2.160]	< 0.001
			Any	1.927	[1.877, 1.978]	< 0.001
	Low	N = 156,140	Unanimity	1.196	[1.126, 1.269]	< 0.001
			Majority	1.186	[1.119, 1.257]	< 0.001
			Any	1.282	[1.213, 1.355]	< 0.001
Credibility Score	High	N = 924,409	Unanimity	1.252	[1.219, 1.285]	< 0.001
			Majority	1.238	[1.206, 1.271]	< 0.001
			Any	1.231	[1.201, 1.261]	< 0.001
	Mid	N = 634,576	Unanimity	2.492	[2.411, 2.576]	< 0.001
			Majority	2.417	[2.340, 2.497]	< 0.001
			Any	2.140	[2.076, 2.206]	< 0.001
	Low	N = 196,997	Unanimity	1.135	[1.095, 1.176]	< 0.001
			Majority	1.115	[1.078, 1.154]	< 0.001
			Any	1.139	[1.103, 1.176]	< 0.001
User Leaning	Left	N = 931,317	Unanimity	1.001	[0.972, 1.030]	0.959
			Majority	0.993	[0.965, 1.021]	0.610
			Any	0.950	[0.926, 0.974]	< 0.001
	Right	N = 688,317	Unanimity	1.680	[1.632, 1.729]	< 0.001
			Majority	1.665	[1.619, 1.713]	< 0.001
			Any	1.695	[1.652, 1.740]	< 0.001

Table A17: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by poster *political self-identification* (Left, Right) and veracity measures. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Veracity Measure	Label	Leaning	Subset Size	Agreement	IRR	CI	p
None		Left	$N = 326{,}132$	Unanimity	1.349	[1.303, 1.397]	< 0.001
				Majority	1.359	[1.314, 1.407]	< 0.001
				Any	1.270	[1.230, 1.311]	< 0.001
		Right	$N = 716{,}867$	Unanimity	1.617	[1.586, 1.649]	< 0.001
				Majority	1.599	[1.568, 1.630]	< 0.001
				Any	1.574	[1.545, 1.604]	< 0.001
Fact-check Rating	True	Left	$N = 34{,}436$	Unanimity	0.769	[0.684, 0.864]	< 0.001
				Majority	0.760	[0.676, 0.854]	< 0.001
				Any	0.729	[0.652, 0.815]	< 0.001
		Right	$N = 25,690$	Unanimity	1.445	[1.210, 1.726]	< 0.001
				Majority	1.319	[1.114, 1.562]	0.001
				Any	1.523	[1.306, 1.776]	< 0.001
	False	Left	$N = 48{,}111$	Unanimity	1.353	[1.282, 1.428]	< 0.001
				Majority	1.337	[1.267, 1.410]	< 0.001
				Any	1.263	[1.199, 1.330]	< 0.001
		Right	$N = 139{,}756$	Unanimity	1.475	[1.443, 1.509]	< 0.001
				Majority	1.462	[1.430, 1.495]	< 0.001
				Any	1.433	[1.402, 1.465]	< 0.001
Factuality Level	High	Left	$N = 75{,}779$	Unanimity	0.956	[0.887, 1.030]	0.239
				Majority	0.957	[0.889, 1.031]	0.246
				Any	0.909	[0.848, 0.976]	0.008
		Right	$N = 32{,}863$	Unanimity	1.992	[1.823, 2.176]	< 0.001
				Majority	1.957	[1.792, 2.137]	< 0.001
				Any	1.864	[1.714, 2.027]	< 0.001
	Low	Left	$N = 22{,}406$	Unanimity	1.579	[1.438, 1.733]	< 0.001
				Majority	1.566	[1.427, 1.718]	< 0.001
				Any	1.406	[1.283, 1.542]	< 0.001
		Right	$N = 220{,}081$	Unanimity	2.080	[2.02, 2.141]	< 0.001
				Majority	2.044	[1.986, 2.104]	< 0.001
				Any	2.05	[1.992, 2.109]	< 0.001
Credibility Score	High	Left	$N = 150{,}255$	Unanimity	0.893	[0.846, 0.942]	< 0.001
				Majority	0.901	[0.855, 0.950]	< 0.001
				Any	0.855	[0.814, 0.898]	< 0.001
		Right	$N = 67{,}199$	Unanimity	1.488	[1.398, 1.584]	< 0.001
				Majority	1.467	[1.379, 1.562]	< 0.001
				Any	1.406	[1.324, 1.493]	< 0.001
	Low	Left	$N = 37{,}646$	Unanimity	1.310	[1.199, 1.431]	< 0.001
				Majority	1.329	[1.219, 1.448]	< 0.001
				Any	1.368	[1.259, 1.487]	< 0.001
		Right	$N = 249{,}677$	Unanimity	1.300	[1.262, 1.339]	< 0.001
				Majority	1.283	[1.246, 1.321]	< 0.001
				Any	1.342	[1.304, 1.380]	< 0.001

Table A18: Neg. bin. regression estimates of the effect of Scientific Authority Cues (SACs) on retweet count, disaggregated by shared URL's Media Bias/Fact Check *source slant* (Left, Right) and veracity measures. Only *cued* indicator estimates are reported. Parameter estimates are not reported for Label–Bias combinations lacking enough cases.

Veracity Measure	Label	Bias	Subset Size	Agreement	IRR	CI	p
None		Left	$N = 128{,}929$	Unanimity	1.003	[0.977, 1.029]	0.848
				Majority	1.003	[0.978, 1.029]	0.805
				Any	0.953	[0.930, 0.975]	< 0.001
		Right	$N = 402{,}816$	Unanimity	1.298	[1.261, 1.337]	< 0.001
				Majority	1.29	[1.254, 1.328]	< 0.001
				Any	1.28	[1.244, 1.317]	< 0.001
Fact-check Rating	True	Left	$N = 128{,}929$	Unanimity	1.177	[1.101, 1.260]	< 0.001
				Majority	1.177	[1.101, 1.260]	< 0.001
				Any	1.092	[1.028, 1.161]	0.005
		Right	$N = 26{,}205$	Unanimity	0.231	[0.183, 0.291]	< 0.001
				Majority	0.231	[0.183, 0.291]	< 0.001
				Any	0.251	[0.201, 0.314]	< 0.001
	False	Left	$N = 48{,}111$	Unanimity	1.198	[1.118, 1.284]	< 0.001
				Majority	1.198	[1.118, 1.284]	< 0.001
				Any	1.201	[1.121, 1.286]	< 0.001
		Right	$N = 139{,}756$	Unanimity	1.195	[1.147, 1.246]	< 0.001
				Majority	1.190	[1.142, 1.240]	< 0.001
				Any	1.177	[1.13, 1.226]	< 0.001
Factuality Level	High	Left	$N = 348{,}651$	Unanimity	0.837	[0.802, 0.873]	< 0.001
				Majority	0.833	[0.799, 0.869]	< 0.001
				Any	0.835	[0.802, 0.869]	< 0.001
		Right	$N = 37{,}277$	Unanimity	1.497	[1.343, 1.669]	< 0.001
				Majority	1.497	[1.343, 1.669]	< 0.001
				Any	1.457	[1.312, 1.619]	< 0.001
	Low	Left	$N = 3{,}633$	Unanimity			
				Majority			
				Any	0.569	[0.284, 1.139]	0.111
		Right	$N = 14{,}439$	Unanimity	2.441	[2.143, 2.780]	< 0.001
				Majority	2.441	[2.143, 2.780]	< 0.001
				Any	2.429	[2.134, 2.765]	< 0.001
Credibility Score	High	Left	$N = 605{,}800$	Unanimity	1.075	[1.045, 1.106]	< 0.001
				Majority	1.074	[1.044, 1.105]	< 0.001
				Any	1.013	[0.987, 1.040]	0.323
		Right	$N = 123{,}462$	Unanimity	1.076	[1.027, 1.128]	0.002
				Majority	1.076	[1.026, 1.127]	0.002
				Any	1.073	[1.024, 1.125]	0.003
	Low	Left	$N = 3{,}633$	Unanimity			
				Majority			
				Any	0.569	[0.284, 1.139]	0.111
		Right	$N = 15{,}879$	Unanimity	2.576	[2.275, 2.918]	< 0.001
				Majority	2.576	[2.275, 2.918]	< 0.001
				Any	2.551	[2.254, 2.888]	< 0.001

Table A19: Negative Binomial regression estimates of the effect of Scientific Authority Attribution (SAA) on retweet counts. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Condition	Value	Subset Size	Parameter	IRR	CI	p
Source Dataset	Combined	$N = 8{,}729{,}332$	Unanimity	1.485	[1.473, 1.498]	< 0.001
			Majority	1.490	[1.477, 1.502]	< 0.001
			Any	1.420	[1.409, 1.431]	< 0.001
	Condor	$N = 8{,}142{,}088$	Unanimity	1.462	[1.449, 1.476]	< 0.001
			Majority	1.466	[1.453, 1.479]	< 0.001
			Any	1.391	[1.380, 1.403]	< 0.001
	Science Feedback	$N = 1{,}903{,}546$	Unanimity	1.055	[1.041, 1.070]	< 0.001
			Majority	1.051	[1.037, 1.066]	< 0.001
			Any	1.032	[1.019, 1.046]	< 0.001
URL Fact-check	True	$N = 703{,}732$	Unanimity	0.680	[0.657, 0.704]	< 0.001
			Majority	0.737	[0.713, 0.763]	< 0.001
			Any	0.932	[0.903, 0.962]	< 0.001
	False	$N = 3{,}873{,}614$	Unanimity	1.216	[1.203, 1.228]	< 0.001
			Majority	1.220	[1.208, 1.233]	< 0.001
			Any	1.125	[1.114, 1.136]	< 0.001
Domain Factuality	High	$N = 791{,}477$	Unanimity	1.249	[1.218, 1.280]	< 0.001
			Majority	1.255	[1.225, 1.287]	< 0.001
			Any	1.230	[1.200, 1.260]	< 0.001
	Low	$N = 1{,}213{,}228$	Unanimity	1.097	[1.080, 1.115]	< 0.001
			Majority	1.086	[1.069, 1.104]	< 0.001
			Any	1.077	[1.06, 1.094]	< 0.001
Credibility Score	High	$N = 1{,}572{,}942$	Unanimity	1.390	[1.365, 1.416]	< 0.001
			Majority	1.401	[1.375, 1.427]	< 0.001
			Any	1.336	[1.313, 1.360]	< 0.001
	Low	$N = 1{,}790{,}094$	Unanimity	1.637	[1.614, 1.661]	< 0.001
			Majority	1.622	[1.600, 1.645]	< 0.001
			Any	1.664	[1.641, 1.687]	< 0.001
User Leaning	Left	$N = 1{,}551{,}200$	Unanimity	0.905	[0.885, 0.926]	< 0.001
			Majority	0.904	[0.884, 0.924]	< 0.001
			Any	0.888	[0.870, 0.906]	< 0.001
	Right	$N = 2{,}383{,}478$	Unanimity	1.205	[1.189, 1.221]	< 0.001
			Majority	1.205	[1.189, 1.221]	< 0.001
			Any	1.161	[1.146, 1.176]	< 0.001

Table A20: Negative Binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, including commercial entities. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Condition	Value	Subset Size	Parameter	IRR	CI	p
Source Dataset	Combined	$N = 8{,}729{,}332$	Unanimity	1.973	[1.960, 1.987]	< 0.001
			Majority	1.975	[1.962, 1.989]	< 0.001
			Any	1.884	[1.871, 1.896]	< 0.001
	Condor	$N = 8{,}142{,}088$	Unanimity	2.022	[2.007, 2.036]	< 0.001
			Majority	2.025	[2.010, 2.040]	< 0.001
			Any	1.923	[1.910, 1.936]	< 0.001
	Science Feedback	$N = 1{,}903{,}546$	Unanimity	2.102	[2.078, 2.126]	< 0.001
			Majority	2.086	[2.062, 2.110]	< 0.001
			Any	2.049	[2.026, 2.073]	< 0.001
URL Fact-check	True	$N = 703{,}732$	Unanimity	0.701	[0.680, 0.722]	< 0.001
			Majority	0.746	[0.724, 0.768]	< 0.001
			Any	0.873	[0.849, 0.897]	< 0.001
	False	$N = 3{,}873{,}614$	Unanimity	1.796	[1.780, 1.812]	< 0.001
			Majority	1.811	[1.795, 1.827]	< 0.001
			Any	1.689	[1.674, 1.703]	< 0.001
Domain Factuality	High	$N = 791{,}477$	Unanimity	1.181	[1.155, 1.207]	< 0.001
			Majority	1.185	[1.159, 1.211]	< 0.001
			Any	1.166	[1.142, 1.190]	< 0.001
	Low	$N = 1{,}213{,}228$	Unanimity	2.068	[2.041, 2.096]	< 0.001
			Majority	2.051	[2.023, 2.078]	< 0.001
			Any	2.042	[2.015, 2.07]	< 0.001
Credibility Score	High	$N = 1{,}572{,}942$	Unanimity	1.293	[1.273, 1.314]	< 0.001
			Majority	1.298	[1.277, 1.319]	< 0.001
			Any	1.234	[1.216, 1.253]	< 0.001
	Low	$N = 1{,}790{,}094$	Unanimity	1.845	[1.821, 1.869]	< 0.001
			Majority	1.880	[1.856, 1.904]	< 0.001
			Any	1.942	[1.918, 1.967]	< 0.001
User Leaning	Left	$N = 1{,}551{,}200$	Unanimity	1.092	[1.071, 1.113]	< 0.001
			Majority	1.084	[1.064, 1.104]	< 0.001
			Any	1.045	[1.027, 1.064]	< 0.001
	Right	$N = 2{,}383{,}478$	Unanimity	2.232	[2.208, 2.256]	< 0.001
			Majority	2.219	[2.195, 2.243]	< 0.001
			Any	2.142	[2.120, 2.165]	< 0.001

Table A21: Logistic regression estimates of the effect of Scientific Authority Cues (SACs) on retweet likelihood for *Condor*, *Science Feedback*, and the combined dataset. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Database	**Agreement**	**OR**	**CI**	***p***
Combined	Unanimity	1.457	[1.449, 1.465]	< 0.001
	Majority	1.463	[1.456, 1.471]	< 0.001
	Any	1.361	[1.354, 1.368]	< 0.001
Condor	Unanimity	1.468	[1.459, 1.476]	< 0.001
	Majority	1.473	[1.464, 1.481]	< 0.001
	Any	1.360	[1.353, 1.367]	< 0.001
Science Feedback	Unanimity	1.435	[1.422, 1.448]	< 0.001
	Majority	1.439	[1.426, 1.451]	< 0.001
	Any	1.435	[1.423, 1.448]	< 0.001

Table A22: Logistic regression estimates of the effect of Scientific Authority Cues (SACs) on retweet likelihood, disaggregated by veracity measures: URL Fact-check (True, False), Domain Factuality Level (High, Low), and Domain Credibility Score (High, Low). Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Veracity Measure	**Label**	**Subset Size**	**Parameter**	**OR**	**CI**	***p***
URL Fact-check	True	$N = 703{,}732$	Unanimity	0.942	[0.918, 0.967]	< 0.001
			Majority	0.962	[0.939, 0.987]	0.002
			Any	0.950	[0.928, 0.972]	< 0.001
	False	$N = 3{,}873{,}614$	Unanimity	1.390	[1.381, 1.399]	< 0.001
			Majority	1.399	[1.390, 1.408]	< 0.001
			Any	1.282	[1.274, 1.290]	< 0.001
Factuality Level	High	$N = 791{,}477$	Unanimity	1.276	[1.254, 1.299]	< 0.001
			Majority	1.282	[1.259, 1.305]	< 0.001
			Any	1.272	[1.251, 1.294]	< 0.001
	Low	$N = 1{,}213{,}228$	Unanimity	1.356	[1.342, 1.371]	< 0.001
			Majority	1.351	[1.337, 1.365]	< 0.001
			Any	1.341	[1.327, 1.355]	< 0.001
Credibility Score	High	$N = 1{,}572{,}942$	Unanimity	1.282	[1.266, 1.298]	< 0.001
			Majority	1.291	[1.275, 1.307]	< 0.001
			Any	1.280	[1.265, 1.295]	< 0.001
	Low	$N = 1{,}790{,}094$	Unanimity	1.508	[1.493, 1.523]	< 0.001
			Majority	1.510	[1.495, 1.525]	< 0.001
			Any	1.521	[1.506, 1.535]	< 0.001

Table A23: Logistic regression estimates of the effect of Scientific Authority Cues (SACs) on retweet likelihood, disaggregated by inferred poster partisanship (Left, Right) and veracity measures. Follower count and intercept omitted for clarity, only *Cued* indicator estimates are reported.

Veracity Measure	Label	Leaning	Subset Size	Agreement	OR	CI	p
None		Left	$N = 1{,}551{,}200$	Unanimity	1.163	[1.146, 1.182]	< 0.001
				Majority	1.170	[1.152, 1.188]	< 0.001
				Any	1.209	[1.192, 1.226]	< 0.001
		Right	$N = 2{,}383{,}478$	Unanimity	1.412	[1.400, 1.424]	< 0.001
				Majority	1.412	[1.400, 1.424]	< 0.001
				Any	1.355	[1.344, 1.366]	< 0.001
Fact-check Rating	True	Left	$N = 157{,}200$	Unanimity	1.173	[1.12, 1.229]	< 0.001
				Majority	1.175	[1.122, 1.23]	< 0.001
				Any	1.119	[1.07, 1.169]	< 0.001
		Right	$N = 64{,}349$	Unanimity	0.967	[0.881, 1.061]	0.483
				Majority	0.986	[0.904, 1.075]	0.746
				Any	1.086	[1.006, 1.173]	0.035
	False	Left	$N = 402{,}798$	Unanimity	1.155	[1.126, 1.184]	< 0.001
				Majority	1.167	[1.139, 1.196]	< 0.001
				Any	1.288	[1.259, 1.317]	< 0.001
		Right	$N = 1{,}584{,}988$	Unanimity	1.289	[1.276, 1.301]	< 0.001
				Majority	1.290	[1.278, 1.302]	< 0.001
				Any	1.225	[1.213, 1.236]	< 0.001
Factuality Level	High	Left	$N = 320{,}921$	Unanimity	1.124	[1.088, 1.162]	< 0.001
				Majority	1.131	[1.094, 1.168]	< 0.001
				Any	1.131	[1.097, 1.167]	< 0.001
		Right	$N = 97{,}100$	Unanimity	1.565	[1.504, 1.630]	< 0.001
				Majority	1.554	[1.492, 1.617]	< 0.001
				Any	1.583	[1.522, 1.647]	< 0.001
	Low	Left	$N = 69{,}269$	Unanimity	1.206	[1.149, 1.267]	< 0.001
				Majority	1.204	[1.147, 1.264]	< 0.001
				Any	1.173	[1.118, 1.231]	< 0.001
		Right	$N = 720{,}597$	Unanimity	1.352	[1.334, 1.371]	< 0.001
				Majority	1.345	[1.327, 1.363]	< 0.001
				Any	1.333	[1.316, 1.351]	< 0.001
Credibility Score	High	Left	$N = 603{,}312$	Unanimity	1.207	[1.18, 1.234]	< 0.001
				Majority	1.212	[1.185, 1.239]	< 0.001
				Any	1.184	[1.160, 1.209]	< 0.001
		Right	$N = 212{,}938$	Unanimity	1.680	[1.633, 1.727]	< 0.001
				Majority	1.675	[1.629, 1.723]	< 0.001
				Any	1.691	[1.645, 1.738]	< 0.001
	Low	Left	$N = 198{,}868$	Unanimity	1.473	[1.409, 1.54]	< 0.001
				Majority	1.522	[1.459, 1.589]	< 0.001
				Any	1.531	[1.469, 1.596]	< 0.001
		Right	$N = 802{,}983$	Unanimity	1.268	[1.250, 1.286]	< 0.001
				Majority	1.265	[1.248, 1.283]	< 0.001
				Any	1.267	[1.250, 1.285]	< 0.001

Table A24: Negative Binomial regression estimates of the effect of Scientific Authority Cues (SACs) on overall spread for *Condor*, *Science Feedback*, and the combined dataset.

Database	**Agreement**	**IRR**	**CI**	***p***
Combined	Unanimity	1.072	[1.014, 1.133]	0.014
	Majority	1.063	[1.006, 1.123]	0.028
	Any	1.036	[0.983, 1.092]	0.186
Condor	Unanimity	1.238	[1.163, 1.318]	< 0.001
	Majority	1.219	[1.146, 1.296]	< 0.001
	Any	1.172	[1.105, 1.243]	< 0.001
Science Feedback	Unanimity	0.986	[0.884, 1.098]	0.793
	Majority	0.971	[0.872, 1.082]	0.592
	Any	0.929	[0.835, 1.034]	0.178

Table A25: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on overall spread, disaggregated by veracity measures: URL Fact-check (True, False), Domain Factuality Level (High, Low), and Domain Credibility Score (High, Low).

Veracity Measure	**Label**	**Subset Size**	**Parameter**	**IRR**	**CI**	***p***
URL Fact-check	True	$N = 6{,}113$	Unanimity	0.786	[0.641, 0.964]	0.021
			Majority	0.811	[0.664, 0.991]	0.040
			Any	0.848	[0.703, 1.024]	0.086
	False	$N = 31{,}865$	Unanimity	1.339	[1.247, 1.438]	< 0.001
			Majority	1.334	[1.244, 1.432]	< 0.001
			Any	1.271	[1.187, 1.361]	< 0.001
Domain Factuality	High	$N = 3{,}723$	Unanimity	1.173	[0.961, 1.432]	0.118
			Majority	1.156	[0.950, 1.406]	0.149
			Any	1.104	[0.917, 1.329]	0.295
	Low	$N = 7{,}969$	Unanimity	1.532	[1.371, 1.712]	< 0.001
			Majority	1.497	[1.341, 1.672]	< 0.001
			Any	1.430	[1.283, 1.594]	< 0.001
Credibility Score	High	$N = 10{,}099$	Unanimity	1.173	[1.034, 1.331]	0.0131
			Majority	1.167	[1.031, 1.321]	0.0144
			Any	1.115	[0.991, 1.253]	0.0696
	Low	$N = 10{,}225$	Unanimity	0.955	[0.861, 1.059]	0.383
			Majority	0.958	[0.865, 1.062]	0.416
			Any	1.017	[0.920, 1.125]	0.742

Table A26: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on overall spread, disaggregated by average poster partisanship (Left, Right) and veracity measures.

Veracity Measure	Label	Leaning	Subset Size	Agreement	IRR	CI	p
None		Left	$N = 7{,}098$	Unanimity	0.611	[0.510, 0.732]	< 0.001
				Majority	0.618	[0.517, 0.739]	< 0.001
				Any	0.665	[0.562, 0.786]	< 0.001
		Right	$N = 2{,}383{,}478$	Unanimity	1.709	[1.559, 1.874]	< 0.001
				Majority	1.674	[1.529, 1.832]	< 0.001
				Any	1.586	[1.454, 1.730]	< 0.001
Fact-check Rating	True	Left	$N = 767$	Unanimity	0.896	[0.564, 1.422]	0.641
				Majority	0.927	[0.586, 1.469]	0.748
				Any	0.925	[0.599, 1.427]	0.723
		Right	$N = 498$	Unanimity	0.518	[0.245, 1.097]	0.086
				Majority	0.528	[0.254, 1.095]	0.086
				Any	0.565	[0.292, 1.093]	0.090
	False	Left	$N = 402{,}798$	Unanimity	0.396	[0.264, 0.595]	< 0.001
				Majority	0.432	[0.288, 0.646]	< 0.001
				Any	0.435	[0.295, 0.643]	< 0.001
		Right	$N = 10{,}758$	Unanimity	1.865	[1.677, 2.074]	< 0.001
				Majority	1.840	[1.657, 2.044]	< 0.001
				Any	1.758	[1.588, 1.947]	< 0.001
Factuality Level	High	Left	$N = 1{,}349$	Unanimity	0.598	[0.434, 0.825]	0.002
				Majority	0.590	[0.430, 0.810]	0.001
				Any	0.592	[0.438, 0.799]	< 0.001
		Right	$N = 416$	Unanimity	3.913	[2.180, 7.022]	< 0.001
				Majority	3.825	[2.143, 6.828]	< 0.001
				Any	3.785	[2.170, 6.602]	< 0.001
	Low	Left	$N = 69{,}269$	Unanimity	0.271	[0.137, 0.538]	< 0.001
				Majority	0.256	[0.131, 0.501]	< 0.001
				Any	0.230	[0.122, 0.434]	< 0.001
		Right	$N = 4482$	Unanimity	1.798	[1.580, 2.047]	< 0.001
				Majority	1.751	[1.540, 1.990]	< 0.001
				Any	1.656	[1.460, 1.877]	< 0.001
Credibility Score	High	Left	$N = 3{,}055$	Unanimity	0.807	[0.646, 1.008]	0.059
				Majority	0.792	[0.636, 0.986]	0.034
				Any	0.778	[0.634, 0.955]	0.017
		Right	$N = 1{,}216$	Unanimity	2.626	[1.880, 3.669]	< 0.001
				Majority	2.561	[1.841, 3.561]	< 0.001
				Any	2.386	[1.736, 3.280]	< 0.001
	Low	Left	$N = 411$	Unanimity	0.569	[0.299, 1.083]	0.086
				Majority	1.491	[0.795, 2.796]	0.213
				Any	1.377	[0.747, 2.538]	0.306
		Right	$N = 4{,}081$	Unanimity	0.985	[0.854, 1.137]	0.838
				Majority	0.972	[0.844, 1.120]	0.697
				Any	1.048	[0.913, 1.202]	0.507

Table A27: Negative Binomial estimates of the effect of Scientific Authority Cues (SACs) on retweet counts for *Condor*, *Science Feedback*, and the combined dataset. Standard errors are clustered at the user level to account for the non-independence of tweets from the same poster. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Database	**Agreement**	**IRR**	**CI**	***p***
Combined	Unanimity	1.745	[1.586, 1.920]	< 0.001
	Majority	1.756	[1.599, 1.928]	< 0.001
	Any	1.670	[1.526, 1.828]	< 0.001
Condor	Unanimity	1.550	[1.431, 1.679]	< 0.001
	Majority	1.537	[1.418, 1.665]	< 0.001
	Any	1.508	[1.392, 1.634]	< 0.001
Science Feedback	Unanimity	1.714	[1.572, 1.869]	< 0.001
	Majority	1.723	[1.583, 1.875]	< 0.001
	Any	1.647	[1.517, 1.788]	< 0.001

Table A28: Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by veracity measures. Standard errors are clustered at the user level to account for the non-independence of tweets from the same poster. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Veracity Measure	Label	Subset Size	Parameter	IRR	CI	p
URL Fact-check	True	$N = 703{,}732$	Unanimity	0.707	[0.599, 0.834]	< 0.001
			Majority	0.754	[0.643, 0.885]	< 0.001
			Any	0.886	[0.719, 1.091]	0.254
	False	$N = 3{,}873{,}614$	Unanimity	1.468	[1.344, 1.604]	< 0.001
			Majority	1.487	[1.361, 1.624]	< 0.001
			Any	1.385	[1.264, 1.518]	< 0.001
Domain Factuality	High	$N = 791{,}477$	Unanimity	1.183	[1.045, 1.339]	0.008
			Majority	1.187	[1.048, 1.344]	0.007
			Any	1.166	[1.029, 1.320]	0.016
	Low	$N = 1{,}213{,}228$	Unanimity	1.971	[1.807, 2.148]	< 0.001
			Majority	1.948	[1.787, 2.124]	< 0.001
			Any	1.935	[1.776, 2.109]	< 0.001
Credibility Score	High	$N = 1{,}572{,}942$	Unanimity	1.304	[0.906, 1.879]	0.153
			Majority	1.309	[0.918, 1.866]	0.137
			Any	1.240	[0.909, 1.692]	0.175
	Low	$N = 1{,}790{,}094$	Unanimity	1.637	[1.485, 1.805]	< 0.001
			Majority	1.679	[1.522, 1.852]	< 0.001
			Any	1.739	[1.580, 1.914]	< 0.001

Table A29: Negative Binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet count, disaggregated by inferred poster partisanship (Left, Right) and veracity measures. Standard errors are clustered at the user level. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Veracity Measure	Label	Leaning	Subset Size	Agreement	IRR	CI	p
None		Left	$N = 1{,}551{,}200$	Unanimity	1.068	[0.945, 1.207]	0.295
				Majority	1.060	[0.939, 1.195]	0.346
				Any	1.023	[0.913, 1.147]	0.694
		Right	$N = 2{,}383{,}478$	Unanimity	1.731	[1.607, 1.864]	< 0.001
				Majority	1.724	[1.602, 1.855]	< 0.001
				Any	1.665	[1.547, 1.791]	< 0.001
Fact-check Rating	True	Left	$N = 157{,}200$	Unanimity	0.870	[0.616, 1.230]	0.431
				Majority	0.877	[0.623, 1.236]	0.454
				Any	0.850	[0.614, 1.176]	0.326
		Right	$N = 64{,}349$	Unanimity	0.967	[0.881, 1.061]	0.483
				Majority	0.986	[0.904, 1.075]	0.746
				Any	1.086	[1.006, 1.173]	0.035
	False	Left	$N = 402{,}798$	Unanimity	1.099	[0.890, 1.358]	0.381
				Majority	1.097	[0.889, 1.353]	0.388
				Any	1.050	[0.851, 1.296]	0.647
		Right	$N = 1{,}584{,}988$	Unanimity	1.517	[1.394, 1.650]	< 0.001
				Majority	1.510	[1.388, 1.641]	< 0.001
				Any	1.454	[1.337, 1.581]	< 0.001
Factuality Level	High	Left	$N = 320{,}921$	Unanimity	0.870	[0.741, 1.021]	0.088
				Majority	0.865	[0.738, 1.015]	0.075
				Any	0.857	[0.734, 1.000]	0.050
		Right	$N = 97{,}100$	Unanimity	1.767	[1.481, 2.108]	< 0.001
				Majority	1.744	[1.462, 2.080]	< 0.001
				Any	1.758	[1.467, 2.108]	< 0.001
	Low	Left	$N = 69{,}269$	Unanimity	2.067	[1.554, 2.749]	< 0.001
				Majority	2.023	[1.525, 2.683]	< 0.001
				Any	1.973	[1.496, 2.602]	< 0.001
		Right	$N = 720{,}597$	Unanimity	2.179	[1.945, 2.441]	< 0.001
				Majority	2.149	[1.919, 2.406]	< 0.001
				Any	2.089	[1.867, 2.336]	< 0.001
Credibility Score	High	Left	$N = 603{,}312$	Unanimity	0.931	[0.819, 1.058]	0.273
				Majority	0.926	[0.815, 1.051]	0.234
				Any	0.880	[0.780, 0.993]	0.039
		Right	$N = 212{,}938$	Unanimity	1.668	[1.372, 2.027]	< 0.001
				Majority	1.667	[1.370, 2.029]	< 0.001
				Any	1.659	[1.355, 2.033]	< 0.001
	Low	Left	$N = 198{,}868$	Unanimity	1.424	[0.975, 2.079]	0.067
				Majority	1.426	[0.978, 2.080]	0.065
				Any	1.445	[0.989, 2.112]	0.057
		Right	$N = 802{,}983$	Unanimity	1.358	[1.237, 1.492]	< 0.001
				Majority	1.352	[1.232, 1.484]	< 0.001
				Any	1.376	[1.257, 1.505]	< 0.001

Table A30: Mixed Effects Negative Binomial model estimates of the effect of Scientific Authority Cues (SACs) on Retweet Counts for *Condor*, *Science Feedback*, and the combined dataset. Models include user-level random intercepts to account for baseline differences in engagement across posters. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Database	**Agreement**	**IRR**	**CI**	***p***
Condor	Unanimity	1.255	[1.246, 1.264]	< 0.001
	Majority	1.254	[1.245, 1.263]	< 0.001
	Any	1.250	[1.242, 1.259]	< 0.001
Science Feedback	Unanimity	1.304	[1.290, 1.319]	< 0.001
	Majority	1.304	[1.290, 1.318]	< 0.001
	Any	1.296	[1.282, 1.310]	< 0.001
Combined	Unanimity	1.250	[1.241, 1.258]	< 0.001
	Majority	1.251	[1.242, 1.259]	< 0.001
	Any	1.247	[1.239, 1.255]	< 0.001

Table A31: Mixed Effects Negative binomial regression estimates of the effect of Scientific Authority Cues (SACs) on Retweet Counts, disaggregated by veracity measures. Models include user-level random intercepts to account for baseline differences in engagement across posters. Follower count and intercept omitted for clarity, only *cued* indicator estimates are reported.

Veracity Measure	Label	Subset Size	Parameter	IRR	CI	p
URL Fact-check	True	$N = 703{,}732$	Unanimity	0.918	[0.891, 0.947]	< 0.001
			Majority	0.940	[0.912, 0.968]	< 0.001
			Any	0.925	[0.899, 0.951]	< 0.001
	False	$N = 3{,}873{,}614$	Unanimity	1.217	[1.207, 1.227]	< 0.001
			Majority	1.222	[1.212, 1.232]	< 0.001
			Any	1.223	[1.213, 1.232]	< 0.001
Domain Factuality	High	$N = 791{,}477$	Unanimity	1.102	[1.078, 1.128]	< 0.001
			Majority	1.104	[1.079, 1.129]	< 0.001
			Any	1.095	[1.071, 1.119]	< 0.001
	Low	$N = 1{,}213{,}228$	Unanimity	1.292	[1.276, 1.309]	< 0.001
			Majority	1.278	[1.262, 1.295]	< 0.001
			Any	1.271	[1.255, 1.288]	< 0.001
Credibility Score	High	$N = 1{,}572{,}942$	Unanimity	1.109	[1.092, 1.126]	< 0.001
			Majority	1.112	[1.096, 1.129]	< 0.001
			Any	1.097	[1.081, 1.113]	< 0.001
	Low	$N = 1{,}790{,}094$	Unanimity	1.315	[1.299, 1.332]	< 0.001
			Majority	1.312	[1.296, 1.328]	< 0.001
			Any	1.320	[1.305, 1.336]	< 0.001

Table A32: Mixed Effects Negative Binomial regression estimates of the effect of Scientific Authority Cues (SACs) on retweet counts, disaggregated by inferred poster partisanship (Left, Right) and veracity measures. Models include user-level random intercepts to account for baseline differences in engagement across posters.

Veracity Measure	Label	Leaning	Subset Size	Agreement	OR	CI	p
None		Left	$N = 1{,}551{,}200$	Unanimity	1.099	[1.079, 1.120]	< 0.001
				Majority	1.102	[1.082, 1.122]	< 0.001
				Any	1.105	[1.087, 1.124]	< 0.001
		Right	$N = 2{,}383{,}478$	Unanimity	1.272	[1.259, 1.286]	< 0.001
				Majority	1.266	[1.253, 1.279]	< 0.001
				Any	1.258	[1.245, 1.271]	< 0.001
Fact-check Rating	True	Left	$N = 157{,}200$	Unanimity	1.203	[1.139, 1.270]	< 0.001
				Majority	1.209	[1.145, 1.276]	< 0.001
				Any	1.150	[1.092, 1.211]	< 0.001
		Right	$N = 64{,}349$	Unanimity	0.922	[0.820, 1.035]	0.169
				Majority	0.909	[0.815, 1.015]	0.090
				Any	0.995	[0.901, 1.098]	0.915
	False	Left	$N = 402{,}798$	Unanimity	1.100	[1.065, 1.136]	< 0.001
				Majority	1.106	[1.071, 1.142]	< 0.001
				Any	1.186	[1.150, 1.224]	< 0.001
		Right	$N = 1{,}584{,}988$	Unanimity	1.202	[1.187, 1.216]	< 0.001
				Majority	1.198	[1.184, 1.212]	< 0.001
				Any	1.190	[1.176, 1.204]	< 0.001
Factuality Level	High	Left	$N = 320{,}921$	Unanimity	1.022	[0.984, 1.061]	0.270
				Majority	1.026	[0.989, 1.065]	0.175
				Any	1.020	[0.984, 1.057]	0.271
		Right	$N = 97{,}100$	Unanimity	1.544	[1.460, 1.633]	< 0.001
				Majority	1.519	[1.437, 1.606]	< 0.001
				Any	1.567	[1.483, 1.656]	< 0.001
	Low	Left	$N = 69{,}269$	Unanimity	1.114	[1.047, 1.185]	< 0.001
				Majority	1.101	[1.035, 1.172]	0.002
				Any	1.091	[1.026, 1.161]	0.006
		Right	$N = 720{,}597$	Unanimity	1.333	[1.311, 1.355]	< 0.001
				Majority	1.317	[1.295, 1.339]	< 0.001
				Any	1.300	[1.278, 1.321]	< 0.001
Credibility Score	High	Left	$N = 603{,}312$	Unanimity	1.129	[1.100, 1.158]	< 0.001
				Majority	1.129	[1.101, 1.158]	< 0.001
				Any	1.093	[1.067, 1.120]	< 0.001
		Right	$N = 212{,}938$	Unanimity	1.433	[1.381, 1.487]	< 0.001
				Majority	1.421	[1.369, 1.475]	< 0.001
				Any	1.434	[1.383, 1.487]	< 0.001
	Low	Left	$N = 198{,}868$	Unanimity	1.136	[1.070, 1.206]	< 0.001
				Majority	1.152	[1.086, 1.222]	< 0.001
				Any	1.190	[1.123, 1.260]	< 0.001
		Right	$N = 802{,}983$	Unanimity	1.191	[1.172, 1.212]	< 0.001
				Majority	1.183	[1.163, 1.202]	< 0.001
				Any	1.178	[1.159, 1.197]	< 0.001

Table A33: Distribution of URLs and SAC presence by veracity measures.

Veracity Measure	Label	URLs N.	URLs %	SAC (unanimity)	SAC (majority)	SAC (any)
Fact-Check Rating	True	6,113	10.8%	9.2%	9.7%	11.1%
	False	31,865	56.4%	25.4%	26.3%	28.8%
	Neutral	15,474	27.3%	9.1%	9.8%	11.7%
	Other	3,056	5.4%	8.2%	8.4%	10.2%
Factuality level	High	3,723	19.8%	12.3%	12.7%	14.6%
	Mid	7,090	37.7%	17.7%	18.4%	20.0%
	Low	7,969	42.4%	33.5%	34.6%	37.2%
Credibility Score	High	10,099	32.9%	12.1%	12.6%	14.4%
	Mid	10,320	33.6%	21.8%	22.8%	24.7%
	Low	10,225	33.3%	29.1%	30.0%	32.6%

Table A34: Percentage of Scientific Authority Attribution (over all SACs) overall and per veracity category and annotation agreement level.

Category	Label	Ratio of Scientific Authority Attribution		
		Unanimity	Majority	Any
Overall		75,8%	75.5%	73.5%
Fact-Check Rating	True	71.5%	71.8%	70.7%
	False	77.5%	77.4%	75.6%
	Neutral	67.6%	66.6%	63.9%
	Other	74.9%	73.7%	72.4%
Political Slant	Left	75.3%	74.3%	72.3%
	Right	75.7%	75.7%	73.7%

Table A35: Average Bot Score for retweeters of content in the different subsets of the observational study. Bot percentage is the proportion of accounts in the subset with a Bot Score > 0.5.

Political Leaning	Veracity	Mean Bot Score		Estimated bot percentage	
		Cued	Uncued	Cued	Uncued
Left	**True**	0.186	0.183	2.900%	2.500%
	False	0.190	0.188	3.500%	3.100%
	Neutral	0.201	0.199	3.400%	3.000%
	Other	0.182	0.169	2.300%	1.500%
Right	**True**	0.171	0.184	2.300%	3.000%
	False	0.200	0.201	2.000%	2.900%
	Neutral	0.189	0.193	2.500%	2.700%
	Other	0.181	0.181	2.600%	2.700%
Overall		0.188	0.187	3.000%	2.950%

Appendix 4.14: Experimental study tables

Table A36: Complete list of *claims*. Participants were randomly assigned to one of the seven possible conditions per claim. The manipulation is bolded here for the reader.

No Mention	Named			Unnamed		
	Individual	**Institution**	**Source**	**Individual**	**Institution**	**Source**
Cannabis use disorder may be linked to growing number of heart attacks in younger adults.	**Dr. Dan Stock says** cannabis use disorder may be linked to growing number of heart attacks in younger adults.	Cannabis use disorder may be linked to growing number of heart attacks in younger adults**, according to the World Health Organization (WHO).**	**The journal Lancet says** cannabis use disorder may be linked to growing number of heart attacks in younger adults.	**Health experts say** cannabis use disorder may be linked to growing number of heart attacks in younger adults.	**Health authorities say** cannabis use disorder may be linked to growing number of heart attacks in younger adults.	**According to medical trials,** cannabis use disorder may be linked to growing number of heart attacks in younger adults.
McDonald's fries contain silicone.	McDonald's fries contain silicone**, says Dr. Dan Stock.**	**According to the World Health Organization (WHO),** McDonald's fries contain silicone.	McDonald's fries contain silicone**, says the American Journal of Medicine.**	**Doctors say** McDonald's fries contain silicone.	McDonald's fries contain silicone, **according to research laboratory.**	**Research says** McDonald's fries contain silicone.
Vaccines don't cause autism.	**Dr. Anthony Fauci says** vaccines don't cause autism.	Vaccines don't cause autism**, says Harvard University.**	**According to The Lancet,** vaccines don't cause autism.	Vaccines don't cause autism**, say health experts.**	Vaccines don't cause autism**, according to health authorities.**	Vaccines don't cause autism**, according to research.**
Vaccines are linked to sterilization of women.	Vaccines are linked to sterilization of women**, says Dr. Anthony Fauci.**	**According to Harvard,** vaccines are linked to sterilization of women.	Vaccines are linked to sterilization of women**, according to the American Journal of Medicine.**	**Doctors say** vaccines are linked to sterilization of women.	**Research laboratory says** vaccines are linked to sterilization of women.	Vaccines are linked to sterilization of women**, according to medical trials.**

Continued on the next page

Table A36 – continued from previous page

No Mention	Named			Unnamed		
	Individual	**Institution**	**Source**	**Individual**	**Institution**	**Source**
Carbon dioxide has reached levels not seen in 3 million years.	**Dr. Patrick Moore says** carbon dioxide has reached levels not seen in 3 million years.	Carbon dioxide has reached levels not seen in 3 million years, **according to the IPCC (Intergovernmental Panel on Climate Change)**.	**According to the journal Nature,** carbon dioxide has reached levels not seen in 3 million years.	Carbon dioxide has reached levels not seen in 3 million years, **says leading climate scientist.**	Carbon dioxide has reached levels not seen in 3 million years, **according to a research center.**	**Study says** carbon dioxide has reached levels not seen in 3 million years.
Climate change is caused by shifts in Earth's solar orbit, not SUVs and fossil fuels.	Climate change is caused by shifts in Earth's solar orbit, not SUVs and fossil fuels **says Dr. Peter Ridd.**	**According to the Massachusetts Institute of Technology (MIT),** Climate change is caused by shifts in Earth's solar orbit, not SUVs and fossil fuels.	Climate change is caused by shifts in Earth's solar orbit, not SUVs and fossil fuels**, according to the European Scientific Journal.**	**Researchers say** climate change is caused by shifts in Earth's solar orbit, not SUVs and fossil fuels.	**According to a research center,** climate change is caused by shifts in Earth's solar orbit, not SUVs and fossil fuels.	Climate change is caused by shifts in Earth's solar orbit, not SUVs and fossil fuels**, says a climate report.**
Seas are rising at the fastest rate in the last 28 centuries.	Seas are rising at the fastest rate in the last 28 centuries**, says Dr. Patrick Moore.**	**According to the IPCC (Intergovernmental Panel on Climate Change),** seas are rising at the fastest rate in the last 28 centuries.	Seas are rising at the fastest rate in the last 28 centuries**, according to the journal Nature.**	**Leading climate scientist says** seas are rising at the fastest rate in the last 28 centuries.	**According to a climate research center,** seas are rising at the fastest rate in the last 28 centuries.	**Study says** seas are rising at the fastest rate in the last 28 centuries.

Continued on the next page

Table A36 – continued from previous page

No Mention	Named			Unnamed		
	Individual	**Institution**	**Source**	**Individual**	**Institution**	**Source**
Electric cars are dirtier for the planet than diesel vehicles.	**Dr. Peter Ridd says** electric cars are dirtier for the planet than diesel vehicles.	Electric cars are dirtier for the planet than diesel vehicles**, according to the Massachusetts Institute of Technology (MIT).**	**According to the European Scientific Journal,** electric cars are dirtier for the planet than diesel vehicles.	Electric cars are dirtier for the planet than diesel vehicles**, researchers say.**	Electric cars are dirtier for the planet than diesel vehicles**, says a research center.**	**Climate report says** electric cars are dirtier for the planet than diesel vehicles.
Single people are healthier and more satisfied compared to those in couples.	**Dr. Peter McCullough says** single people are healthier and more satisfied compared to those in couples.	Single people are healthier and more satisfied compared to those in couples**, according to the Office for National Statistics.**	**According to the journal Scientific American,** single people are healthier and more satisfied compared to those in couples.	**Scientists say** single people are healthier and more satisfied compared to those in couples.	Single people are healthier and more satisfied compared to those in couples**, says a research center.**	**According to statistics,** single people are healthier and more satisfied compared to those in couples.
Women are actually attracted to men with big bellies.	Women are actually attracted to men with big bellies**, says Dr. Robert Malone.**	**According to the University of Oxford,** women are actually attracted to men with big bellies.	**According to the International Conference on Science and Society,** women are actually attracted to men with big bellies.	**Psychologists say** women are actually attracted to men with big bellies.	**A top university says** women are actually attracted to men with big bellies.	**According to peer reviewed study,** women are actually attracted to men with big bellies.
People who swear a lot make good friends.	People who swear a lot make good friends**, says Dr. Peter McCullough.**	**According to the Office for National Statistics,** people who swear a lot make good friends.	People who swear a lot make good friends**, according to the journal Scientific American.**	People who swear a lot make good friends**, scientists say.**	**A research center says** people who swear a lot make good friends.	People who swear a lot make good friends**, according to statistics.**

Continued on the next page

Table A36 – continued from previous page

No Mention	Named			Unnamed		
	Individual	**Institution**	**Source**	**Individual**	**Institution**	**Source**
Intelligence is passed from mothers, not from fathers.	**Dr. Robert Malone says** intelligence is passed from mothers, not from fathers.	Intelligence is passed from mothers, not from fathers, **according to the University of Oxford.**	Intelligence is passed from mothers, not from fathers, **according to the International Conference on Science and Society.**	Intelligence is passed from mothers, not from fathers, **psychologists say.**	Intelligence is passed from mothers, not from fathers, **according to a top university.**	Intelligence is passed from mothers, not from fathers, **according to peer-reviewed study.**
Migration linked to 82% of growth in the population.	Migration linked to 82% of growth in the population, **says Dr. Dolores Cahill.**	Migration linked to 82% of growth in the population, **according to Stanford University.**	According to the journal Science, migration linked to 82% of growth in the population.	Migration linked to 82% of growth in the population, **researchers say.**	**A research institute says** migration linked to 82% of growth in the population.	Migration linked to 82% of growth in the population, **according to scientific evidence.**
The first contraceptive injection for men could arrive in the coming months.	**Dr. Sherri Tenpenny says** the first contraceptive injection for men could arrive in the coming months.	The first contraceptive injection for men could arrive in the coming months, **according to Johns Hopkins University.**	The first contraceptive injection for men could arrive in the coming months, **according to the journal Sage.**	The first contraceptive injection for men could arrive in the coming months, **experts say.**	The first contraceptive injection for men could arrive in the coming months, **says a research center.**	The first contraceptive injection for men could arrive in the coming months, **according to new study.**
Chocolate is on track to go extinct within 40 years.	**Dr. Dolores Cahill says** chocolate is on track to go extinct within 40 years.	**According to Johns Hopkins University,** chocolate is on track to go extinct within 40 years.	Chocolate is on track to go extinct within 40 years, **according to the journal Science.**	**Researchers say** chocolate is on track to go extinct within 40 years.	Chocolate is on track to go extinct within 40 years, **says a research institute.**	**Scientific evidence says** chocolate is on track to go extinct within 40 years.

Continued on the next page

Table A36 – continued from previous page

No Mention	Named			Unnamed		
	Individual	**Institution**	**Source**	**Individual**	**Institution**	**Source**
Living standards won't improve until all the baby boomers are dead.	Living standards won't improve until all the baby boomers are dead**, says Dr. Sherri Tenpenny.**	**According to Stanford University,** living standards won't improve until all the baby boomers are dead.	**According to the journal Sage,** living standards won't improve until all the baby boomers are dead.	**Experts say** living standards won't improve until all the baby boomers are dead.	**A research center says** living standards won't improve until all the baby boomers are dead.	**New study says** living standards won't improve until all the baby boomers are dead.

Table A37: Complete list of *news*. These headlines were not manipulated.

Headline
A doctor was suspended for two years for treating cancer with black radish.
Big Pharma, Gates, Fauci, UK officials accused of crimes against humanity in complaint to International Court.
91-year-old doctor refuses to retire and performs 4 surgeries daily.
CDC under fire from vaccine autism lawsuit.
Volkswagen apologizes for forcing monkeys to breathe exhaust gases.
Iceland prohibits teaching religion to minors under 21 years old.
Mexico is now the ninth richest country in the whole world.
Fines over 700€ for those who throw trash on the ground in Portugal.

Table A38: Contrasts testing hypothesized differences in *sharing intention*. Group estimates from robust linear regression models and pairwise contrasts. Sharing intention was measured on a scale from 0 ("Not at all likely to share") to 4 ("Extremely likely to share"). Abbreviations: SAA (Scientific Authority Attribution), NAA (Named Authority Attribution), UAA (Unnamed Authority Attribution). Estimated means with standard errors and mean differences with confidence intervals in brackets. ** $p < 0.01$, * $p < 0.05$.

Hypothesis	Condition	Mean [SE]	Contrast	Difference [CI]	*t*, *p*
H1	SAA	1.240 [0.027]	No Attribution	0.057* [0.006, 0.109]	$t = 2.174$, $p = 0.030$
H2	NAA	1.197 [0.038]	UAA	−0.020 [−0.094, 0.054]	$t = 0.532$, $p = 0.595$
H3	SAA + False Claim	1.094 [0.032]	No Attribution + False Claim	0.114** [0.043, 0.184]	$t = 3.165$, $p = 0.002$
H3	SAA + True Claim	1.386 [0.035]	No Attribution + True Claim	−0.001 [−0.074, 0.073]	t = -0.017, $p = 0.987$
H4	SAA + Republican	1.322 [0.039]	No Attribution + Republican	0.061 [−0.010, 0.131]	$t = 1.693$, $p = 0.091$
H4	SAA + Democrat	1.156 [0.039]	No Attribution + Democrat	0.050 [−0.027, 0.125]	$t = 1.286$, $p = 0.199$
H5	SAA + False Claim + Democrat	0.899 [0.042]	No Attribution + False Claim + Democrat	0.079 [−0.015, 0.172]	$t = 1.654$, $p = 0.098$
H5	SAA + True Claim + Democrat	1.410 [0.049]	No Attribution + True Claim + Democrat	−0.003 [−0.111, 0.104]	t = -0.062, $p = 0.950$
H5	SAA + False Claim + Republican	1.283 [0.046]	No Attribution + False Claim + Republican	0.124* [0.021, 0.226]	$t = 2.365$, $p = 0.018$
H5	SAA + True Claim + Republican	1.362 [0.049]	No Attribution + True Claim + Republican	0.002 [−0.098, 0.102]	$t = 0.033$, $p = 0.974$

Table A39: Contrasts exploring trust and familiarity differences in *sharing intention*. Group estimates from robust linear regression models and pairwise contrasts. Sharing intention was measured on a scale from 0 ("Not at all likely to share") to 4 ("Extremely likely to share"). Estimated means with standard errors and mean differences with confidence intervals in brackets. *** $p < 0.001$, * $p < 0.05$.

Analysis	Condition	Mean [SE]	Contrast	Difference [CI]	*t*, *p*
Trust × Fact-checked Veracity	Trusted + False	1.168 [0.049]	Not trusted + False	0.127* [0.015, 0.239]	$t = 2.214$, $p = 0.027$
Trust × Fact-checked Veracity	Trusted + True	1.681 [0.053]	Not trusted + True	0.506*** [0.383, 0.628]	$t = 8.087$, $p < 0.001$
Trust × Political Affiliation	Republican + Trusted	1.643 [0.063]	Republican + Not trusted	0.471*** [0.337, 0.605]	$t = 6.877$, $p < 0.001$
Trust × Political Affiliation	Democrat + Trusted	1.285 [0.052]	Democrat + Not trusted	0.272*** [0.149, 0.395]	$t = 4.348$, $p < 0.001$

Table A40: Contrasts exploring trust and familiarity differences in *perceived accuracy*. Group estimates from robust linear regression models and pairwise contrasts. Perceived accuracy was measured on a scale from 0 ("Not at all accurate") to 4 ("Extremely accurate"). Estimated means with standard errors and mean differences with confidence intervals in brackets. ** $p < 0.01$, * $p < 0.05$.

Analysis	**Condition**	**Mean [SE]**	**Contrast**	**Difference [CI]**	***t*, *p***
Trust × Fact-checked Veracity	Trusted + False	1.254 [0.043]	Not trusted + False	-0.035 $[-0.140, 0.069]$	$t = -0.663$, $p = 0.507$
Trust × Fact-checked Veracity	Trusted + True	2.431 [0.041]	Not trusted + True	0.450^{***} $[0.344, 0.555]$	$t = 8.360$, $p < 0.001$
Trust × Political Affiliation	Republican + Trusted	1.832 [0.055]	Republican + Not trusted	0.216^{***} $[0.091, 0.341]$	$t = 3.389$, $p < 0.001$
Trust × Political Affiliation	Democrat + Trusted	1.775 [0.044]	Democrat + Not trusted	0.219^{***} $[0.106, 0.332]$	$t = 3.801$, $p < 0.001$

Table A41: Tests of H1–H5. Models are estimated with robust standard errors, with confidence intervals in brackets. F-differences are computed between nested OLS models with conventional standard errors and degrees of freedom based on parameter differences. N = 1,187, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	H1	H2	H3		H4		H5	
Intercept	1.183^{***} $[1.130, 1.236]$	1.250^{***} $[1.187, 1.314]$	1.358^{***} $[1.296, 1.420]$	1.387^{***} $[1.319, 1.454]$	1.264^{***} $[1.192, 1.336]$	1.261^{***} $[1.185, 1.337]$	1.440^{***} $[1.361, 1.519]$	1.361^{***} $[1.264, 1.457]$
Attribution (vs. No Attribution)	0.057^{*} $[0.006, 0.109]$	–	0.056^{*} $[0.005, 0.108]$	-0.001 $[-0.074, 0.073]$	0.055^{*} $[0.004, 0.107]$	0.061 $[-0.010, 0.131]$	0.054^{*} $[0.003, 0.105]$	0.002 $[-0.099, 0.102]$
Named (vs. Unnamed)	–	-0.020 $[-0.094, 0.054]$	–	–	–	–	–	–
No Attribution (vs. Unnamed)	–	-0.067^{*} $[-0.129, -0.005]$	–	–	–	–	–	–
False (vs. True)	–	–	-0.349^{***} $[-0.403, -0.296]$	-0.407^{***} $[-0.478, -0.335]$	–	–	-0.350^{***} $[-0.403, -0.296]$	-0.202^{***} $[-0.306, -0.098]$
Democrat (vs. Republican)	–	–	–	–	-0.161^{***} $[-0.254, -0.068]$	-0.155^{**} $[-0.260, -0.050]$	-0.162^{***} $[-0.254, -0.069]$	0.053 $[-0.083, 0.188]$
Attribution × False	–	–	–	0.114^{*} $[0.014, 0.215]$	–	–	–	0.122 $[-0.024, 0.268]$
Attribution × Democrat	–	–	–	–	–	-0.011 $[-0.114, 0.092]$	–	-0.005 $[-0.152, 0.142]$
False × Democrat	–	–	–	–	–	–	–	-0.402^{***} $[-0.544, -0.260]$
Attribution × False × Democrat	–	–	–	–	–	–	–	-0.039 $[-0.239, 0.160]$
Adjusted R^2	0.001	0.001	0.019	0.019	0.004	0.004	0.023	0.030
F-Difference (df)	–	–	–	19.390^{*} (1)	–	0.043 (1)	–	17.764^{***} (4)

Table A42: Tests of H6 and exploratory interaction model of perceived accuracy. Models are estimated with robust standard errors, with confidence intervals in brackets. F-differences are computed between nested OLS models with conventional standard errors and degrees of freedom based on parameter differences. N = 1,187, *** $p < 0.001$, ** $p < 0.01$.

	H6	**Exploratory Interaction**	
Intercept	1.541^{***}	1.954^{***}	1.963^{***}
	$[1.500, 1.583]$	$[1.904, 2.004]$	$[1.907, 2.019]$
Attribution (vs. No Attribution)	0.136^{***}	0.134^{***}	0.117^{**}
	$[0.084, 0.189]$	$[0.084, 0.184]$	$[0.045, 0.188]$
False (vs. True)	–	-0.824^{***}	-0.841^{***}
		$[-0.876, -0.772]$	$[-0.911, -0.772]$
Attribution × False	–	–	0.035
			$[-0.062, 0.132]$
Adjusted R^2	0.003	0.107	0.107
F-Difference (df)	–	–	0.511 (1)

Table A43: Test of H7 (top, N = 1,187) and the roles of trust in and familiarity with an attributed entity (bottom, N = 1,182). Unstandardized model parameters estimated using SEMs with robust standard errors. F-differences computed between nested OLS models with conventional standard errors and degrees of freedom based on parameter differences. Fit indices indicate a perfect fit with the data because the models are just-identified.

Path	**Estimate**	**CI**	***p***	**Contribution**
Attribution (vs. No Attribution) → Perceived Accuracy → Sharing Intention	0.067	$[0.041, 0.093]$	< 0.001	117.54%
Attribution (vs. No Attribution) → Sharing Intention	-0.010	$[-0.055, 0.035]$	0.672	−17.54%
Total Effect	0.057	$[0.006, 0.109]$	0.006	100%
F-Difference (df)	2983.700 (1)	–	< 0.001	–
Trust → Perceived Accuracy → Sharing Intention	0.055	$[0.033, 0.077]$	< 0.001	25.58%
Trust → Sharing Intention	0.085	$[0.041, 0.130]$	< 0.001	39.53%
Familiarity → Perceived Accuracy → Sharing Intention	0.004	$[-0.014, 0.021]$	0.692	1.86%
Familiarity → Sharing Intention	0.071	$[0.035, 0.108]$	< 0.001	33.02%
Total Effect	0.215	$[0.170, 0.260]$	< 0.001	100%
F-Difference (df)	1648.700 (1)	–	< 0.001	–

Table A44: Complementary analyses of the interaction between attribution to a trusted entity and fact-checked veracity. Models are estimated with robust standard errors, with confidence intervals in brackets. F-differences are computed between nested OLS models with conventional standard errors and degrees of freedom based on parameter differences. N = 1,182, *** $p < 0.001$, ** $p < 0.01$.

	Sharing Intention		**Perceived Accuracy**	
Intercept	1.254^{***}	1.175^{***}	1.993^{***}	1.891^{***}
	$[1.183, 1.326]$	$[1.098, 1.253]$	$[1.930, 2.055]$	$[1.821, 1.961]$
Trusted (vs. Not Trusted)	0.317^{***}	0.506^{***}	0.208^{***}	0.450^{***}
	$[0.227, 0.406]$	$[0.383, 0.629]$	$[0.130, 0.286]$	$[0.344, 0.555]$
False (vs. True)	-0.294^{***}	-0.134^{**}	-0.805^{***}	-0.601^{***}
	$[-0.369, -0.219]$	$[-0.228, -0.041]$	$[-0.878, -0.733]$	$[-0.694, -0.508]$
Trusted × False	–	-0.379^{***}	–	-0.485^{***}
		$[-0.532, -0.226]$		$[-0.627, -0.343]$
Adjusted R^2	0.028	0.033	0.105	0.114
F-difference (df)	–	26.671^{***} (1)	–	47.280^{***} (1)

Table A45: Complementary analyses of the interaction between attribution to a trusted entity and party affiliation. Models are estimated with robust standard errors, with confidence intervals in brackets. F-differences are computed between nested OLS models with conventional standard errors and degrees of freedom based on parameter differences. N = 1,182, *** $p < 0.001$, ** $p < 0.01$.

	Sharing Intention		**Perceived Accuracy**	
Intercept	1.206^{***}	1.172^{***}	1.616^{***}	1.617^{***}
	$[1.129, 1.282]$	$[1.092, 1.252]$	$[1.552, 1.680]$	$[1.548, 1.686]$
Trusted (vs. Not Trusted)	0.367^{***}	0.471^{***}	0.217^{***}	0.216^{***}
	$[0.276, 0.457]$	$[0.336, 0.605]$	$[0.134, 0.301]$	$[0.091, 0.341]$
Democrat (vs. Republican)	-0.242^{***}	-0.159^{**}	-0.059	-0.061
	$[-0.348, -0.136]$	$[-0.279, -0.038]$	$[-0.145, 0.026]$	$[-0.165, 0.043]$
Trusted × Democrat	–	-0.199^{**}	–	0.003
		$[-0.381, -0.017]$		$[-0.165, 0.172]$
Adjusted R^2	0.023	0.024	0.006	0.006
F-difference (df)	–	6.944 ** (1)	–	0.003 (1)

Table A46: Control models for the tests of H1–H5. Models are estimated with robust standard errors, with confidence intervals in brackets. N = 1,187, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	H1	**H2**	**H3**	**H4**	**H5**
Intercept	1.435*** [1.263, 1.607]	1.498*** [1.325, 1.671]	1.644*** [1.467, 1.820]	1.433*** [1.260, 1.605]	1.533*** [1.351, 1.714]
Attribution (vs. No Attribution)	0.055* [0.003, 0.106]	–	−0.002 [−0.076, 0.071]	0.059 [−0.011, 0.130]	−0.002 [−0.102, 0.098]
Named (vs. Unnamed)	–	−0.017 [−0.091, 0.057]	–	–	–
No Attribution (vs. Unnamed)	–	−0.063* [−0.125, −0.001]	–	–	–
False (vs. True)	–	–	−0.407*** [−0.478, −0.335]	–	−0.207*** [−0.311, −0.103]
Democrat (vs. Republican)	−0.161*** [−0.253, −0.068]	−0.160*** [−0.253, −0.067]	−0.161*** [−0.253, −0.068]	−0.156** [−0.261, −0.050]	0.049 [−0.086, 0.184]
Attribution × False	–	–	0.112* [0.013, 0.212]	–	0.127 [−0.019, 0.273]
Attribution × Democrat	–	–	–	−0.010 [−0.113, 0.094]	−0.003 [−0.150, 0.144]
False × Democrat	–	–	–	–	−0.396*** [−0.538, −0.254]
Attribution × False × Democrat	–	–	–	–	−0.042 [−0.241, 0.158]
Age	−0.004* [−0.007, −0.001]	−0.004* [−0.007, −0.001]	−0.004** [−0.007, −0.001]	−0.004* [−0.007, −0.001]	−0.004* [−0.007, −0.001]
Male (vs. Female)	−0.054 [−0.147, 0.040]	−0.054 [−0.147, 0.040]	−0.056 [−0.149, 0.038]	−0.054 [−0.147, 0.040]	−0.054 [−0.147, 0.039]
College (vs. No College)	0.080 [−0.014, 0.174]	0.080 [−0.014, 0.174]	0.080 [−0.014, 0.173]	0.080 [−0.014, 0.174]	0.079 [−0.015, 0.173]
Adjusted R^2	0.007	0.007	0.026	0.007	0.032

Table A47: Control models for the tests of H6 and the exploratory interaction. Models are estimated with robust standard errors, with confidence intervals in brackets. N = 1,187, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	H6	**Exploratory Interaction**
Intercept	1.811*** [1.689, 1.932]	2.244*** [2.118, 2.369]
Attribution (vs. No Attribution)	0.136*** [0.083, 0.188]	0.114** [0.042, 0.185]
False (vs. True)	–	−0.845*** [−0.914, −0.775]
Democrat (vs. Republican)	−0.004 [−0.069, 0.062]	−0.006 [−0.070, 0.058]
Attribution × False	–	0.040 [−0.057, 0.137]
Age	−0.005*** [−0.008, −0.003]	−0.006*** [−0.008, −0.003]
Male (vs. Female)	−0.081* [−0.147, −0.016]	−0.086* [−0.150, −0.021]
College (vs. No College)	0.024 [−0.042, 0.090]	0.022 [−0.043, 0.087]
Adjusted R^2	0.006	0.111

Table A48: Control models for the complementary analyses of the interactions between attribution to a trusted entity, fact-checked veracity, and party affiliation. Models are estimated with robust standard errors, with confidence intervals in brackets. N = 1,182. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Sharing Intention		**Perceived Accuracy**	
Intercept	1.351***	1.349***	2.087***	1.826***
	[1.158, 1.544]	[1.155, 1.543]	[1.676, 1.977]	[1.676, 1.977]
Trusted (vs. Not Trusted)	0.489***	0.444***	0.445***	0.202***
	[0.366, 0.611]	[0.310, 0.578]	[0.338, 0.552]	[0.076, 0.329]
False (vs. True)	−0.134**	–	−0.601***	–
	[−0.227, −0.040]		[−0.695, −0.508]	
Democrat (vs. Republican)	–	−0.163**	–	−0.065
		[−0.283, −0.043]		[−0.168, 0.039]
Trusted × False	−0.376***	–	−0.483***	–
	[−0.529, −0.223]		[−0.625, −0.341]	
Trusted × Democrat	–	−0.178	–	0.019
		[−0.359, 0.003]		[−0.150, 0.187]
Age	−0.004*	−0.004*	−0.004**	−0.004**
	[−0.008, −0.001]	[−0.008, −0.001]	[−0.007, −0.001]	[−0.007, −0.001]
Male (vs. Female)	−0.084	−0.077	−0.091*	−0.091*
	[−0.190, 0.021]	[−0.182, 0.027]	[−0.172, −0.009]	[−0.175, −0.006]
College (vs. No College)	0.127*	0.120*	0.029	0.037
	[0.020, 0.234]	[0.014, 0.226]	[−0.054, 0.112]	[−0.050, 0.123]
Adjusted R^2	0.037	0.028	0.116	0.008

Table A49: Ordinal logistic regression models for the tests of H1–H5 and H7. Model parameter estimates are in log odds, with confidence intervals in brackets. N = 1,187, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	H1	**H2**	**H3**	**H4**	**H5**	**H7**
Not at all | Not very	-0.661^{***} $[-0.804, -0.519]$	-0.779^{***} $[-0.930, -0.629]$	-0.980^{***} $[-1.133, -0.827]$	-0.661^{*} $[-0.808, -0.514]$	-0.815^{***} $[-0.981, -0.650]$	0.704^{***} $[0.547, 0.861]$
Not very | Somewhat	0.146^{*} $[0.004, 0.288]$	0.028 $[-0.121, 0.178]$	-0.160^{*} $[-0.312, -0.008]$	0.146 $[-0.000, 0.293]$	0.010 $[-0.155, 0.176]$	1.706^{***} $[1.546, 1.866]$
Somewhat | Very	1.154^{***} $[1.109, 1.299]$	1.036^{***} $[0.884, 1.188]$	0.860^{***} $[0.706, 1.014]$	1.154^{*} $[1.005, 1.304]$	1.035^{***} $[0.867, 1.202]$	2.944^{***} $[2.776, 3.113]$
Very | Extremely	2.300^{***} $[2.143, 2.457]$	2.182^{***} $[2.019, 2.346]$	2.013^{***} $[1.849, 2.178]$	2.300^{*} $[2.139, 2.461]$	2.190^{***} $[2.012, 2.367]$	4.259^{***} $[4.075, 4.443]$
Attribution (vs. No Attribution)	0.101^{**} $[0.028, 0.174]$	–	0.014 $[-0.088, 0.116]$	0.102 $[-0.001, 0.204]$	0.022 $[-0.123, 0.166]$	0.001 $[-0.075, 0.077]$
Named (vs. Unnamed)	–	-0.034 $[-0.137, 0.068]$	–	–	–	–
No Attribution (vs. Unnamed)	–	-0.118^{***} $[-0.207, -0.029]$	–	–	–	–
False (vs. True)	–	–	-0.603^{***} $[-0.708, -0.498]$	–	-0.276^{***} $[-0.423, -0.128]$	–
Democrat (vs. Republican)	-0.245^{***} $[-0.318, -0.172]$	-0.244^{*} $[-0.318, -0.171]$	-0.250^{***} $[-0.324, -0.176]$	-0.244^{***} $[-0.349, -0.140]$	0.077 $[-0.069, 0.222]$	-0.293^{***} $[-0.369, -0.217]$
Attribution × False	–	–	0.178^{*} $[0.031, 0.325]$	–	0.166 $[-0.039, 0.371]$	–
Attribution × Democrat	–	–	–	-0.001 $[-0.147, 0.145]$	-0.014 $[-0.219, 0.190]$	–
False × Democrat	–	–	–	–	-0.668^{***} $[-0.879, -0.457]$	–
Attribution × False × Democrat	–	–	–	–	0.009 $[-0.285, 0.304]$	–
Perceived Accuracy	–	–	–	–	0.851^{***} $[0.816, 0.885]$	–
Age	-0.007^{***} $[-0.009, -0.004]$	-0.007^{***} $[-0.009, -0.004]$	-0.007^{***} $[-0.009, -0.004]$	-0.007^{***} $[-0.009, -0.004]$	-0.007^{***} $[-0.009, -0.004]$	-0.003^{*} $[-0.006, -0.000]$
Male (vs. Female)	-0.057 $[-0.131, 0.017]$	-0.057 $[-0.131, 0.017]$	-0.062 $[-0.136, 0.012]$	-0.057 $[-0.131, 0.017]$	-0.060 $[-0.134, 0.014]$	0.005 $[-0.072, 0.082]$
College (vs. No College)	0.093^{*} $[0.019, 0.167]$	0.093^{*} $[0.019, 0.167]$	0.097^{*} $[0.022, 0.171]$	0.093^{*} $[0.019, 0.167]$	0.096^{*} $[0.022, 0.170]$	0.098^{*} $[0.022, 0.175]$
AIC	27299.618	27297.090	27107.233	27297.525	27035.263	24649.554
BIC	27356.905	27368.699	27186.002	27369.133	27135.515	24721.163
Log Likelihood	–13641.809	–13638.545	–13542.616	–13638.762	–13503.631	–12314.777

Table A50: Ordinal regression models for the tests of of H6 and the exploratory interaction. Model parameter estimates are in log odds, with confidence intervals in brackets. N = 1,187, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	H6	**Interaction**
Not at all | Not very	-2.092^{***}	-2.107^{***}
	$[-2.242, -1.942]$	$[-2.262, -1.953]$
Not very | Somewhat	-1.016^{***}	-1.032^{***}
	$[-1.162, -0.871]$	$[-1.181, -0.882]$
Somewhat | Very	0.226^{**}	0.210^{***}
	$[0.082, 0.370]$	$[0.062, 0.359]$
Very | Extremely	1.525^{***}	1.509^{***}
	$[1.373, 1.677]$	$[1.352, 1.666]$
Attribution (vs. No Attribution)	0.202^{***}	0.172^{***}
	$[0.129, 0.274]$	$[0.070, 0.273]$
False (vs. True)	-1.219^{***}	-1.249^{***}
	$[-1.294, -1.143]$	$[-1.354, -1.145]$
Attribution × False	–	0.061
		$[-0.083, 0.205]$
Age	-0.008^{***}	-0.008^{***}
	$[-0.011, -0.006]$	$[-0.011, -0.006]$
Male (vs. Female)	-0.124^{***}	-0.125^{***}
	$[-0.197, -0.052]$	$[-0.198, -0.052]$
College (vs. No College)	0.030	0.031
	$[-0.042, 0.103]$	$[-0.042, 0.103]$
AIC	28429.227	28430.537
BIC	28500.835	28509.307
Log Likelihood	–14204.613	–14204.269

Table A51: Ordinal logistic regression models for the complementary analyses of the interactions between attribution to a trusted entity, fact-checked veracity, and party affiliation. Model parameter estimates are in log odds, with confidence intervals in brackets. N = 1,182, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Sharing Intention		**Perceived Accuracy**	
Not at all | Not very	-0.615^{***}	-0.626^{***}	-1.398^{***}	-1.902^{***}
	$[-0.816, -0.414]$	$[-0.829, -0.423]$	$[-1.597, -1.200]$	$[-2.108, -1.695]$
Not very | Somewhat	0.217^{*}	0.213^{*}	-0.410^{***}	-0.829^{***}
	$[0.017, 0.417]$	$[0.011, 0.415]$	$[-0.604, -0.216]$	$[-1.029, -0.629]$
Somewhat | Very	1.305^{***}	1.307^{***}	0.730^{***}	0.420^{***}
	$[1.100, 1.510]$	$[1.100, 1.515]$	$[0.535, 0.925]$	$[0.221, 0.619]$
Very | Extremely	2.487^{***}	2.493^{***}	2.035^{***}	1.796^{***}
	$[2.264, 2.710]$	$[2.267, 2.718]$	$[1.826, 2.243]$	$[1.584, 2.008]$
Trusted (vs. Not Trusted)	0.610^{***}	0.670^{***}	0.267^{***}	0.644^{***}
	$[0.455, 0.766]$	$[0.523, 0.817]$	$[0.113, 0.420]$	$[0.497, 0.790]$
False (vs. True)	–	-0.204^{**}	–	-0.893^{***}
		$[-0.340, -0.069]$		$[-1.027, -0.759]$
Democrat (vs. Republican)	-0.243^{***}	–	-0.103	–
	$[-0.382, -0.105]$		$[-0.237, 0.031]$	
Trusted × False	–	-0.544^{***}	–	-0.740^{***}
		$[-0.754, -0.334]$		$[-0.948, -0.533]$
Trusted × Democrat	-0.256^{*}	–	0.035	–
	$[-0.470, -0.041]$		$[-0.176, 0.246]$	
Age	-0.007^{***}	-0.007^{***}	-0.006^{**}	-0.006^{**}
	$[-0.011, -0.003]$	$[-0.011, -0.003]$	$[-0.010, -0.002]$	$[-0.010, -0.002]$
Male (vs. Female)	-0.102	-0.109^{*}	-0.136^{**}	-0.134^{*}
	$[-0.207, 0.003]$	$[-0.214, -0.004]$	$[-0.238, -0.034]$	$[-0.237, -0.032]$
College (vs. No College)	0.146^{**}	0.157^{**}	0.057	0.049
	$[0.041, 0.251]$	$[0.052, 0.262]$	$[-0.046, 0.159]$	$[-0.054, 0.152]$
AIC	13677.158	13632.286	14826.752	14278.624
BIC	13741.836	13696.964	14891.429	14343.302
Log Likelihood	–6828.579	–6806.143	–7403.376	–7129.312

Table A52: Complementary analysis of scientific authority mentions (SAMs). Models are estimated with robust standard errors, with confidence intervals in brackets. N = 1,185, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Sharing Intention	
Intercept	1.052^{***}	1.413^{***}
	$[0.989, 1.114]$	$[1.205, 1.621]$
Mention (vs. No Mention)	0.151^{***}	0.146^{***}
	$[0.084, 0.218]$	$[0.080, 0.212]$
Democrat (vs. Republican)	–	-0.308^{***}
		$[-0.419, -0.197]$
Age	–	-0.006^{**}
		$[-0.010, -0.002]$
Male	–	-0.041
		$[-0.153, 0.070]$
College	–	0.133^{*}
		$[0.020, 0.246]$
Adjusted R^2	0.003	0.023

Retying the Knot

Quatre articles sur la science et l'expertise dans un monde polarisé

Résumé

Résumé

Introduction

Au début du millénaire, George W. Bush retire les États-Unis du Protocole de Kyoto, en contradiction avec le consensus scientifique sur la nécessité de contrer le réchauffement climatique. Pendant la pandémie de COVID-19, les gouvernements qui ont imposé des mesures sanitaires souvent préconisées par des épidémiologistes et des médecins se sont heurtés à la résistance du public et à une contestation ouverte dans de nombreux pays. Ces deux exemples ont un point commun : l'expertise scientifique est de plus en plus contestée.

Pourtant, les sociétés démocratiques contemporaines s'appuient fortement sur l'expertise scientifique pour éclairer leurs politiques, fixer les réglementations économiques et gérer leurs risques. Au cours des dernières décennies, les désaccords publics concernant la science se sont intensifiés, en particulier lorsque les faits viennent appuyer la décision publique. Ces désaccords ne signalent pas une dispute sur la factualité, mais des tensions plus larges concernant l'exercice du pouvoir, l'utilisation de la science pour l'action publique et l'autorité de l'expertise, qui se déploient dans le contexte d'une société interconnectée de plus en plus polarisée.

Quatre évolutions majeures survenues au cours des soixante-dix dernières années permettent d'expliquer ce contexte. Premièrement, l'essor de la technocratie a conféré à la science un rôle central dans la prise de décision des gouvernants, suscitant une résistance parmi les groupes qui perçoivent l'autorité des experts comme une remise en cause des processus démocratiques. Deuxièmement, la démocratisation de l'enseignement supérieur a acculturé de larges segments de la population à la science, permettant aux citoyens

non seulement de s'appuyer sur l'autorité des experts, mais aussi de la remettre en question. Troisièmement, l'essor d'Internet et des réseaux sociaux a facilité la production et la diffusion d'informations destinées à des publics spécifiques. Enfin, dans de nombreux contextes nationaux, les opinions des citoyens se sont de plus en plus polarisées. Ensemble, ces évolutions ont favorisé un contexte où l'opposition à un mode de gouvernance technocratique peut être exprimée par des citoyens familiarisés avec le fonctionnement de la science, tandis que des ressources alternatives sont mises à disposition en ligne pour venir contester l'autorité des experts, souvent au service d'objectifs politiques.

La science dans la modernité

La science est généralement comprise comme une méthode organisée et cumulative permettant de connaître le monde, produisant des connaissances via l'expérimentation directe et, idéalement, exempte de tout parti pris (Dooremalen et al., 2015). Le statut privilégié de la science (Cologna et al., 2025) repose sur sa capacité à générer des énoncés factuellement vrais qui semblent parler d'eux-mêmes (Latour, 1993), ce qui la positionne comme notre outil le plus puissant pour produire des connaissances (Shapin, 2008). La méthode scientifique confère à la science un statut d'arbitre rationnel et « objectif de la « Vérité » » (Lewis, 2010, p. 11).

La séparation entre les faits et les valeurs est ce que Latour décrit comme le « nœud gordien » de la modernité : « à gauche la connaissance des choses, à droite l'intérêt, le pouvoir et la politique des hommes » (Latour, 1993, p. 3). Pourtant, dans la pratique, la production de connaissances est une entreprise sociale indissociable de son contexte social (Shapin, 1994). La science est ancrée dans tous les éléments constitutifs de ce que nous appelons le « social » (Jasanoff, 2004a, p. 3), et est sensible aux intérêts politiques et économiques (Foucault, 2008; Jasanoff, 2004b).

Comme le soutient Shapin (2008), « l'autorité de la science repose de plus en plus non pas sur ce que les scientifiques savent, mais sur ce qu'ils peuvent contribuer à faire advenir » (p. 443). L'autorité scientifique n'est donc pas seulement épistémique, mais également instrumentale. Cette imbrication n'est pas propre à la science contemporaine : les pratiques scientifiques anciennes étaient déjà liées à des intérêts privés et étatiques (Biagioli, 1990a, 1990b; Habinek, 2010; Schiebinger, 2004; Shapin, 1994), et après la

Seconde Guerre mondiale, la recherche scientifique aux États-Unis a été de plus en plus présentée comme un moyen de favoriser l'innovation industrielle, les progrès médicaux et d'asseoir une domination militaire (Forman, 2007; Gauchat, 2023; Schnaiberg, 1980). La répartition inégale de l'autorité scientifique contribue à expliquer les désaccords sur ce qui rend un domaine de recherche scientifique : il n'existe pas de méthode scientifique unique, mais plutôt différentes sciences pour différents *usages* (Eyal, 2019), et les domaines de recherche dans lesquels la séparation entre les faits et les valeurs n'est pas facilement maintenue peuvent voir leur légitimité scientifique contestée.

Expertise

Un expert se définit comme une personne qui détient des connaissances approfondies et spécialisées dans un domaine spécifique, et qui participe à leur production (Eyal, 2019). Après la Seconde Guerre mondiale,les experts ont été de plus en plus associés aux processus de décision en matière de politiques publiques (Gauchat, 2023), donnant ainsi naissance à une *science réglementaire*: « le type de science sur lequel s'appuient les agences de régulation pour chercher à protéger la santé humaine, la sécurité ou l'environnement » (Dear & Jasanoff, 2010, p. 773).

Le sentiment anti-intellectuel n'est pas né de l'ignorance, mais s'est exprimé à travers une rhétorique populiste qui présente les experts comme utilisant les connaissances scientifiques pour contrôler « le peuple » (van der Meer & Hameleers, 2025). Les données montrent que, si la confiance générale dans la science reste élevée, des désaccords apparaissent autour de questions *spécifiques* où des affirmations scientifiques factuelles sont utilisées pour justifier une réglementation (Blank & Shaw, 2015; Funk et al., 2019; Pfänder et al., 2025). Ce qui est contesté, alors, n'est peut-être pas la connaissance scientifique en soi, mais la manière dont les affirmations sont liées à l'autorité et leur utilisation dans l'exercice du pouvoir. Pendant de nombreuses années, la solution proposée s'est inspirée du modèle du déficit de connaissances (Bak, 2001; Simis et al., 2016; Sturgis & Allum, 2004), qui expliquait la méfiance envers les experts par un déficit de connaissances scientifiques. Pourtant, les données empiriques suggèrent que la relation entre les connaissances et la confiance envers les experts est plus complexe et, parfois, paradoxale.

Éducation

Au cours de la seconde moitié du XXe siècle, l'accès à l'enseignement supérieur s'est rapidement développé (Breen, 2010; Morrisson & Murtin, 2009; Schofer & Meyer, 2005), familiarisant de larges segments du public avec la science qui développent des attitudes positives à son égard (Allum et al., 2008; Brossard & Nisbet, 2007; Noy & O'Brien, 2019). Cette expansion de l'éducation a doté le public instruit d'un *capital scientifique*, permettant aux « individus d'interpréter les informations scientifiques qu'ils rencontrent dans leur vie quotidienne » (Archer et al., 2015, p. 926).

Pourtant, le capital scientifique ne se traduit pas automatiquement par une déférence inconditionnelle de la science. Comme le suggère Eyal (2019), le fait d'apprendre comment les experts produisent des connaissances peut, paradoxalement, rendre les individus *plus* enclins à remettre en question l'autorité des experts. Des niveaux d'éducation plus élevés ont été associés à un désaccord accru avec les experts sur des questions politisées (Drummond & Fischhoff, 2017; Joslyn & Sylvester, 2019; Kahan et al., 2012), tout en prédisposant généralement les citoyens à apprécier les vertus de la science (Alper et al., 2023). Une telle vigilance n'implique pas un rejet de la science en général, mais plutôt un rejet des institutions scientifiques traditionnelles en tant qu'arbitres incontestés de la vérité (Achterberg et al., 2017; van der Meer & Hameleers, 2025).

Société en réseau

Les sociétés contemporaines sont entrées dans un nouvel âge de l'information défini par un paradigme centré « autour des technologies de l'information et de la communication basées sur la microélectronique » (Castells, 2000, p. 5). Dans cette *société en réseau*, l'autorité épistémique est de plus en plus diffuse, les médias d'information traditionnels ont perdu leur monopole sur la diffusion de l'information (Van Aelst et al., 2025), et l'autorité des institutions scientifiques sur la production des faits est de plus en plus contestée (Gauchat, 2023).

Les plateformes en ligne ont tendance à privilégier les contenus chargés en émotion et sensationnels (Brady et al., 2020; Leach et al., 2026; Puryear et al., 2024), ce qui peut fausser la perception qu'ont les utilisateurs des croyances d'autrui, conduisant à des

croyances biaisées insensibles à des tentatives de corrections (Friggeri et al., 2014; Walter & Tukachinsky, 2020). Dans le contexte des connaissances spécialisées, ces conditions peuvent favoriser la diffusion d'affirmations qui s'écartent du consensus scientifique mais s'alignent sur les croyances préexistantes d'un individu (Kata, 2010; Lewandowsky et al., 2013).

Polarisation politique

La polarisation est un processus d'alignement et de tri dans lequel les attitudes deviennent de plus en plus corrélées à des identités partisanes exclusives (Baldassarri & Gelman, 2008). La littérature distingue la polarisation *idéologique* — divergence des attitudes sur les enjeux entre les lignes partisanes — et la polarisation *affective* — aversion émotionnelle et méfiance envers les membres des autres partis (Iyengar et al., 2019). La polarisation réduit la politique à une opposition entre « nous » et « eux » (McCoy et al., 2018), avec des conséquences négatives sur le monde du travail, la situation économique et les relations interpersonnelles (Berntzen et al., 2024; Frye, 2002; McConnell et al., 2018).

Aux États-Unis, plutôt qu'une radicalisation généralisée, les Américains ont connu un processus de « tri partisan » (Fiorina & Abrams, 2008) : un alignement accru entre l'idéologie, les positions sur les enjeux et l'affiliation partisane (Gauchat, 2023; Levendusky, 2009), souvent attribué aux élites politiques qui ont adopté des positions plus extrêmes depuis les années 1970 (Baldassarri & Gelman, 2008). Cela a créé un état de *fausse* polarisation (Druckman et al., 2022) où les citoyens surestiment le fossé idéologique entre eux-mêmes et « l'autre » (Ahler, 2014; Levendusky & Malhotra, 2016; Westfall et al., 2015). La polarisation ne se limite pas aux États-Unis : elle a été observée en Europe, en Asie, en Afrique et en Amérique latine (Cheong & Haggard, 2023; Huang & Kuo, 2022; Kibet & Ward, 2018; McCoy et al., 2018; Rekker, 2021; Sarsfield et al., 2024).

Questions de recherche et méthodologie

Même si le contexte actuel est parfois décrit comme une crise de la légitimité de la science (Gauchat, 2023; Saltelli & Funtowicz, 2017), la science elle-même reste très respectée (Cologna et al., 2025; van der Meer & Hameleers, 2025), tandis que certaines affirmations

spécifiques et l'autorité de ceux qui les avancent sont contestées (McCright et al., 2013; Pfänder et al., 2025; Rekker, 2025). La crise ne concerne pas la science en tant que méthode de production de connaissances, mais l'autorité des experts et ses applications.

S'appuyant sur des travaux antérieurs qui s'éloignent du modèle de déficit de connaissances (Gauchat, 2010; Kahan, 2012, 2013; McCright & Dunlap, 2010), la présente thèse remet en cause la caractérisation courante des libéraux comme « pro-science » et des conservateurs comme « anti-science » (Drummond & Fischhoff, 2017; Gauchat, 2012; Mann & Schleifer, 2020; Oreskes & Conway, 2022), en se concentrant plutôt sur les préférences pour différentes formes d'expertise. Elle s'appuie sur la distinction théorique entre la *science de l'impact* — recherche qui traite des conséquences de l'industrie et des politiques publiques, utilisée pour soutenir la régulation des marchés — et la *science de la production* — recherche alignée sur les intérêts industriels et commerciaux (McCright & Dunlap, 2010; Schnaiberg, 1980). Cette distinction met en évidence la manière dont différentes formes d'expertise sont imbriquées avec des intérêts politiques et économiques, et souligne l'un des domaines les plus polarisants de la politique contemporaine : la redistribution économique et la taille de l'État (Webster & Abramowitz, 2017). Dans les environnements en ligne contemporains, où les individus s'appuient sur des heuristiques pour traiter l'information (Chaiken, 1980; Metzger & Flanagin, 2013), les experts et les sources scientifiques peuvent servir d'indicateurs de crédibilité façonnant la perception de l'exactitude et la volonté de partager l'information, avec des effets potentiellement différents selon les groupes politiques (Rekker, 2021).

Au-delà des préférences politiques, cette thèse examine également si le raisonnement motivé les sous-tend. Des recherches antérieures ont opérationnalisé le raisonnement motivé à travers les interactions entre l'idéologie politique et des mesures du savoir (Drummond & Fischhoff, 2017; Fischer et al., 2022; Savadori et al., 2022), avec des résultats incohérents. Partant de l'idée que l'engagement vis-à-vis des affirmations d'experts relève moins de la *connaissance* que de la familiarité avec les processus de production de connaissances (Eyal, 2019), le raisonnement motivé est plutôt opérationnalisé à travers les interactions entre l'identification politique, l'engagement partisan et l'éducation, capturant à la fois la motivation à défendre les croyances du groupe et la capacité à traiter l'information d'une manière qui s'aligne sur ces croyances (Kunda, 1990). Étant donné que l'imbrication entre science et politique est plus marquée dans certains domaines de recherche que dans

d'autres, le raisonnement motivé peut influencer les préférences concernant des types *spécifiques* de science, et est donc évalué à travers la science générale, la science d'impact et la science de la production.

Cette thèse examine également le rôle des hypothèses individuelles dans le partage d'informations. S'appuyant sur la théorie de la pertinence (Wilson & Sperber, 1995, 1998) et la théorie de l'information (Shannon, 1948), elle établit une distinction entre les *croyances* — propositions considérées comme vraies (Egan, 1986) — et les *attentes* — des prédictions sur ce qui est susceptible de se produire (Barto et al., 2013) — en les conceptualisant comme interagissant avec l'information par le biais de différents canaux : contradictions et renforcements des croyances, et violations et validations des attentes.

Une approche quantitative et multiméthodologique est adoptée tout au long de l'étude. Les données d'enquêtes nationales examinent les tendances au niveau de la population concernant les perceptions de l'expertise. Des expériences d'enquête en ligne isolent les effets causaux des sources expertes et des hypothèses individuelles (Gaines et al., 2007). Les données issues des réseaux sociaux capturent les comportements de partage à grande échelle en ligne (Edelmann et al., 2020). Bien qu'elles soient bien adaptées pour tester les attentes théoriques, ces méthodes réduisent des concepts tels que la confiance, les perceptions de l'exactitude, les croyances et les attentes à des indicateurs numériques qui ne peuvent pas saisir pleinement les significations culturelles et les expériences subjectives impliquées dans la manière dont les individus interagissent avec l'expertise.

Résumés des chapitres

Chapitre 1

Les conservateurs américains sont souvent perçus comme antiscientifiques, et les libéraux comme pro-scientifiques (Kennedy et al., 2022; Mann & Schleifer, 2020; McCright & Dunlap, 2011). Ce point de vue néglige les travaux récents montrant que les groupes politiques ne rejettent ni ne soutiennent la science dans son ensemble (Gligorić et al., 2025; Rekker, 2025; Shi et al., 2017), leurs positions dépendent du domaine de recherche ou de l'affirmation en question. Le chapitre 1 examine comment l'alignement entre le domaine

d'expertise d'un expert et l'idéologie politique d'un individu influence la perception de ces affirmations, à l'aide d'une expérience en ligne pré-enregistrée menée auprès d'adultes américains, dans laquelle les participants ont fait part de leur confiance envers des experts issus des domaines de l'impact et de la production, et ont évalué l'exactitude d'affirmations attribuées de manière aléatoire à ces experts ou restant non sourcées.

Les résultats montrent que les libéraux font davantage confiance aux experts que les conservateurs, bien que l'écart soit moins important pour les experts en production. L'attribution à un expert a accru la précision perçue de certaines affirmations, quelle que soit l'idéologie politique du lecteur. Ces résultats suggèrent *certaines* préférences idéologiques, mais celles-ci ne correspondent pas clairement à la distinction entre impact et production telle qu'initialement théorisée (McCright & Dunlap, 2010; Schnaiberg, 1980).

Chapitre 2

Le chapitre 2 évalue le raisonnement motivé en tant que facteur de polarisation scientifique en Europe. L'identification du raisonnement motivé nécessite d'observer la combinaison de l'engagement partisan et de la capacité cognitive (Kunda, 1990; Leeper & Slothuus, 2014) : le premier fournissant la motivation pour défendre les positions de son groupe politique, et la seconde la capacité à écarter par le raisonnement les informations scientifiques discordantes. À partir de données d'enquête provenant de 11 pays européens, ces éléments ont été opérationnalisés par le biais d'interactions entre l'identification gauche-droite, la proximité avec un parti et le niveau d'éducation, testées en tant que prédicteurs de la confiance dans la science dans son ensemble, la science de l'impact et la science de la production à l'aide de modèles d'équations structurelles.

Les résultats montrent que les Européens alignent leurs préférences sur la distinction entre science d'impact et science de production (Rekker, 2025), et que la proximité avec un parti et le niveau d'éducation amplifient les différences gauche-droite *uniquement* pour la science d'impact. Cela indique que le raisonnement motivé oriente les différences idéologiques spécifiquement vers la recherche qui remet en cause les intérêts industriels et soutient l'intervention réglementaire, ce qui est compatible avec l'idée selon laquelle le raisonnement motivé est plus probable lorsque l'imbrication entre expertise et politique

est plus marquée.

Chapitre 3

Le chapitre 3 examine l'interaction entre l'information et les hypothèses de l'observateur. S'appuyant sur la théorie de la pertinence (Wilson & Sperber, 1995, 1998) et la théorie de l'information (Shannon, 1948), il établit une distinction entre les croyances et les attentes, en appliquant ce cadre pour étudier les intentions de partage dans le cadre d'une expérience en ligne pré-enregistrée menée auprès de participants américains. Les participants ont été exposés de manière aléatoire à des affirmations d'experts qui contredisaient ou confirmaient leurs croyances, violaient ou validaient leurs attentes, ou combinaient ces manipulations.

Les contradictions avec les croyances augmentent davantage l'intérêt perçu que les violations des attentes, mais le renforcement des croyances n'augmente pas les intentions de partage, contrairement à la validation des attentes. La diffusion d'informations d'experts sur les réseaux sociaux pourrait donc dépendre moins de ce que les utilisateurs considèrent comme vrai que de la question de savoir si cela correspond à ce qu'ils *anticipent*.

Chapitre 4

Le chapitre 4 examine comment les *indices d'autorité scientifique* (SAC), définis comme des mentions d'entités signalant une expertise ou des découvertes scientifiques, influencent la diffusion d'informations d'experts vérifiées sur les réseaux sociaux. Il combine une analyse observationnelle de millions de tweets partagés entre 2010 et 2023 — où les SAC ont été identifiés à l'aide d'un modèle linguistique de grande envergure et où les utilisateurs ont été classés comme républicains ou démocrates — avec une expérience en ligne pré-enregistrée dans laquelle les participants ont évalué leur volonté de partager des titres attribués soit à des entités scientifiques, soit à aucune source.

Les résultats montrent que le fait d'attribuer des affirmations à des entités scientifiques augmente la volonté de partager, au prorata de la précision perçue. Cependant, l'effet sur l'intention de partager est spécifique aux fausses affirmations parmi les républicains, ce qui suggère que l'autorité scientifique est détournée pour légitimer la désinformation qui

séduit les utilisateurs en ligne de droite.

Conclusion

Cette thèse a examiné plusieurs facteurs qui déterminent la manière dont les informations fournies par les experts sont perçues et diffusées. Les résultats montrent que les experts sont généralement considérés comme des sources d'information crédibles, ce qui renforce la perception de la fiabilité des affirmations factuelles ; pourtant, la contestation ne vise pas la science dans son ensemble, mais plutôt la manière dont l'autorité des experts et certaines affirmations spécifiques sont utilisées dans la pratique. Cette distinction remet en question l'image couramment véhiculée selon laquelle les libéraux seraient pro-science et les conservateurs anti-science. Aux États-Unis, les différences dans l'évaluation de l'expertise suivent davantage les lignes partisanes, tandis qu'en Europe, les préférences dépendent davantage du type de science évalué, la contestation semblant motivée et concentrée autour de l'expertise liée à la réglementation. Ces tendances mettent en évidence une tension : l'expertise scientifique est largement valorisée, mais elle fait l'objet d'une contestation politique lorsqu'il s'agit d'affirmations et d'experts spécifiques.

Plutôt qu'une crise de la science, cette thèse suggère un contexte plus paradoxal dans lequel l'expertise scientifique est hautement valorisée, mais où l'autorité des experts est contestée. Cette contestation peut refléter la prise de conscience du public que la production de connaissances est un processus social et politique, et que les affirmations peuvent servir des intérêts privés ou être utilisées pour justifier des décisions politiques. Le scepticisme n'est pas toujours un rejet total, mais une résistance face à une autorité incontestée sur les faits, qui reflète un engagement substantiel, parfois critique, vis-à-vis de l'expertise. Cela dit, toutes les contestations ne reposent pas sur des bases équivalentes : le scepticisme à l'égard des affirmations d'experts concernant les risques sociétaux peut avoir des conséquences collectives, et les contre-arguments sur lesquels il s'appuie méritent tout autant d'être examinés de façon critique.

Retying the Knot

Cuatro artículos sobre ciencia y pericia en un mundo polarizado

Resumen

Resumen

Introducción

A principios del milenio, George W. Bush retiró a Estados Unidos del Protocolo de Kioto, en contra del consenso científico sobre la necesidad de combatir el calentamiento global. Durante la pandemia de COVID-19, los gobiernos que impusieron medidas sanitarias —a menudo recomendadas por epidemiólogos y médicos— se enfrentaron a la resistencia del público y a protestas abiertas en numerosos países. Estos dos ejemplos tienen un punto en común: la autoridad científica es cada vez más cuestionada.

Sin embargo, las sociedades democráticas contemporáneas se basan en gran medida en la especialización científica para orientar sus políticas, establecer regulaciones económicas y gestionar riesgos. En las últimas décadas, los desacuerdos públicos en torno a la ciencia se han intensificado, especialmente cuando los hechos científicos respaldan decisiones públicas. Estos desacuerdos no indican una disputa sobre la veracidad de los hechos, sino tensiones más amplias relacionadas con el ejercicio del poder, el uso de la ciencia para la acción pública y la autoridad de la pericia, que se desarrollan en el contexto de una sociedad interconectada y cada vez más polarizada.

Cuatro cambios importantes ocurridos en los últimos setenta años permiten explicar este contexto. En primer lugar, el auge de la tecnocracia ha otorgado a la ciencia un papel central en la toma de decisiones gubernamentales, lo que ha suscitado resistencia entre grupos que perciben la autoridad experta como una amenaza a los procesos democráticos. En segundo lugar, la democratización de la educación superior ha familiarizado a amplios segmentos de la población con la ciencia, lo que permite a los ciudadanos no solo confiar en la autoridad experta, sino también cuestionarla. En tercer lugar, el auge del

Internet y las redes sociales ha facilitado la producción y difusión de información dirigida a públicos específicos. Por último, en numerosos contextos nacionales, las opiniones de los ciudadanos se han polarizado cada vez más. En conjunto, estas evoluciones han propiciado un contexto en el que la oposición a un modo de gobernanza tecnocrático puede ser expresada por ciudadanos familiarizados con el funcionamiento de la ciencia, mientras que ponen a disposición recursos alternativos en línea para cuestionar la autoridad experta, a menudo al servicio de objetivos políticos.

La ciencia en la modernidad

La ciencia se entiende generalmente como un método organizado y acumulativo que permite conocer el mundo, generando conocimiento a través de la experimentación directa y, en teoría, libre de cualquier sesgo (Dooremalen et al., 2015). El estatus privilegiado de la ciencia (Cologna et al., 2025) se basa en su capacidad para generar enunciados verídicos que parecen hablar por sí mismos (Latour, 1993), lo que la posiciona como nuestra herramienta más poderosa para producir conocimiento (Shapin, 2008). El método científico confiere a la ciencia un estatus de árbitro racional y «objetivo de la "Verdad"» (Lewis, 2010, p. 11).

La separación entre hechos y valores es lo que Latour describe como el corte del «nudo gordiano» en la modernidad: «a la izquierda, el conocimiento de las cosas; a la derecha, el interés, el poder y la política de los hombres» (Latour, 1993, p. 3). Sin embargo, en la práctica, la producción de conocimiento es un emprendimiento inseparable de su contexto social (Shapin, 1994). La ciencia está arraigada en todos los elementos constitutivos de lo que llamamos «lo social» (Jasanoff, 2004a, p. 3), y es sensible a intereses políticos y económicos (Foucault, 2008; Jasanoff, 2004b).

Como sostiene Shapin (2008), «la autoridad de la ciencia se basa cada vez más no en lo que los científicos saben, sino en lo que pueden contribuir a hacer realidad» (p. 443). La autoridad científica no es, por lo tanto, solo epistémica, sino también instrumental. Esta imbricación no es exclusiva de la ciencia contemporánea: las prácticas científicas antiguas ya estaban vinculadas a intereses privados y estatales (Biagioli, 1990a, 1990b; Habinek, 2010; Schiebinger, 2004; Shapin, 1994), y tras la Segunda Guerra Mundial, la investigación científica en Estados Unidos se presentó cada vez más como un medio para fomentar la

innovación industrial, los avances médicos y afianzar el dominio militar (Forman, 2007; Gauchat, 2023; Schnaiberg, 1980). La distribución desigual de la autoridad científica contribuye a explicar los desacuerdos sobre qué constituye un campo de investigación científico: no existe un único método científico, sino más bien diferentes ciencias para diferentes *usos* (Eyal, 2019), y los campos de investigación en los que no es fácil mantener la separación entre hechos y valores pueden ver cuestionada su legitimidad científica.

Pericia

Un experto se define como una persona que posee conocimientos profundos y especializados en un ámbito específico, y que participa en su producción (Eyal, 2019). Tras la Segunda Guerra Mundial, los expertos se asociaron cada vez más a procesos de toma de decisiones en materia de políticas públicas (Gauchat, 2023), dando así origen a una *ciencia regulatoria*: «el tipo de ciencia en la que se basan las agencias reguladoras para tratar de proteger la salud humana, la seguridad o el medio ambiente» (Dear & Jasanoff, 2010, p. 773).

El sentimiento antiintelectual no surge de la ignorancia, sino que se ha expresado a través de una retórica populista que presenta a los expertos como personas que utilizan el conocimiento científico para controlar «al pueblo» (van der Meer & Hameleers, 2025). Los datos muestran que, si bien la confianza general en la ciencia sigue siendo alta, hay desacuerdos en torno a cuestiones *específicas* en las que se utilizan afirmaciones científicas basadas en hechos para justificar acciones regulatorias (Blank & Shaw, 2015; Funk et al., 2019; Pfänder et al., 2025).

Lo que se cuestiona, entonces, tal vez no sea el conocimiento científico en sí mismo, sino la forma en la que afirmaciones verídicas se vinculan con autoridad experta y su uso en el ejercicio del poder. Durante muchos años, la solución propuesta se inspiró en el modelo del déficit de conocimiento (Bak, 2001; Simis et al., 2016; Sturgis & Allum, 2004), que explicaba la desconfianza hacia expertos por una falta de conocimientos científicos. Sin embargo, los datos empíricos sugieren que la relación entre el conocimiento y la confianza en expertos es más compleja y, en ocasiones, paradójica.

Educación

Durante la segunda mitad del siglo XX, el acceso a la educación superior se expandió rápidamente (Breen, 2010; Morrisson & Murtin, 2009; Schofer & Meyer, 2005), familiarizando a amplios segmentos del público con la ciencia, lo que les llevó a desarrollar actitudes positivas hacia ella (Allum et al., 2008; Brossard & Nisbet, 2007; Noy & O'Brien, 2019). Esta expansión de la educación ha dotado al público instruido de un *capital científico*, lo que permite a «los individuos interpretar la información científica con la que se encuentran en su vida cotidiana» (Archer et al., 2015, p. 926).

Sin embargo, el capital científico no se traduce automáticamente en deferencia incondicional hacia la ciencia. Como sugiere Eyal (2019), el hecho de aprender cómo los expertos producen conocimiento puede, paradójicamente, hacer que las personas *sean más* propensas a cuestionar su autoridad. Se ha asociado un mayor nivel educativo con un mayor desacuerdo con expertos en cuestiones politizadas (Drummond & Fischhoff, 2017; Joslyn & Sylvester, 2019; Kahan et al., 2012), al mismo tiempo que, en general, predispone a los ciudadanos a apreciar las virtudes de la ciencia (Alper et al., 2023). Tal cautela no implica un rechazo de la ciencia en general, sino más bien un rechazo de las instituciones científicas tradicionales como árbitros indiscutibles de la verdad (Achterberg et al., 2017; van der Meer & Hameleers, 2025).

Sociedad en red

Las sociedades contemporáneas han entrado en una nueva era de la información definida por un paradigma centrado «en torno a las tecnologías de la información y la comunicación basadas en la microelectrónica» (Castells, 2000, p. 5). En esta *sociedad en red*, la autoridad epistémica es cada vez más difusa, los medios de comunicación tradicionales han perdido su monopolio sobre la difusión de la información (Van Aelst et al., 2025), y la autoridad de las instituciones científicas sobre la producción de hechos es cada vez más cuestionada (Gauchat, 2023).

Las plataformas en línea tienden a dar prioridad a los contenidos cargados de emoción y sensacionalistas (Brady et al., 2020; Leach et al., 2026; Puryear et al., 2024), lo que puede distorsionar la percepción que tienen los usuarios del consenso social y político.

Esto conduce a creencias sesgadas insensibles a los intentos de corrección (Friggeri et al., 2014; Walter & Tukachinsky, 2020). En el contexto del conocimiento especializado, estas condiciones pueden favorecer la difusión de afirmaciones que se alejan del consenso científico, pero que se alinean con las creencias preexistentes de un individuo (Kata, 2010; Lewandowsky et al., 2013).

Polarización política

La polarización es un proceso de alineación y selección en el que las actitudes se vuelven cada vez más correlacionadas con identidades partidistas exclusivas (Baldassarri & Gelman, 2008). La literatura distingue entre la polarización *ideológica* —divergencia de actitudes sobre los temas entre las líneas partidistas— y la polarización *afectiva* —aversión emocional y desconfianza hacia los miembros de otros partidos (Iyengar et al., 2019)—. La polarización reduce la política a una oposición entre «nosotros» contra «ellos» (McCoy et al., 2018), con consecuencias negativas para el mundo laboral, la situación económica y las relaciones interpersonales (Berntzen et al., 2024; Frye, 2002; McConnell et al., 2018).

En Estados Unidos, en lugar de una radicalización generalizada, el público ha experimentado un proceso de «selección partidista» (Fiorina & Abrams, 2008): una mayor alineación entre la ideología, las posturas sobre distintos temas y la afiliación partidista (Gauchat, 2023; Levendusky, 2009), a menudo atribuida a las élites políticas que han adoptado posturas más extremas desde la década de 1970 (Baldassarri & Gelman, 2008). Esto ha creado un estado de *falsa* polarización (Druckman et al., 2022) en el que los ciudadanos sobreestiman la brecha ideológica entre ellos mismos y «el otro» (Ahler, 2014; Levendusky & Malhotra, 2016; Westfall et al., 2015). La polarización no se limita a los Estados Unidos: se ha observado en Europa, Asia, África y América Latina (Cheong & Haggard, 2023; Huang & Kuo, 2022; Kibet & Ward, 2018; McCoy et al., 2018; Rekker, 2021; Sarsfield et al., 2024).

Preguntas de investigación y metodología

Aunque el contexto actual se describe a veces como una crisis de legitimidad de la ciencia (Gauchat, 2023; Saltelli & Funtowicz, 2017), la ciencia en sí misma sigue gozando de

gran respeto (Cologna et al., 2025; van der Meer & Hameleers, 2025), mientras que se cuestionan ciertas afirmaciones de expertos específicas y la autoridad de quienes las plantean (McCright et al., 2013; Pfänder et al., 2025; Rekker, 2025). La crisis no afecta a la ciencia como método de producción de conocimiento, sino a la autoridad experta y sus aplicaciones.

Basándose en trabajos anteriores que se alejan del modelo del déficit de conocimiento (Gauchat, 2010; Kahan, 2012, 2013; McCright & Dunlap, 2010), la presente tesis cuestiona la caracterización habitual de los liberales como «procientíficos» y de los conservadores como «anticientíficos » (Drummond & Fischhoff, 2017; Gauchat, 2012; Mann & Schleifer, 2020; Oreskes & Conway, 2022), centrándose más bien en preferencias por diferentes formas de pericia. Se basa en la distinción teórica entre la *ciencia de impacto* —investigación que aborda las consecuencias de la industria y políticas públicas, utilizada para respaldar la regulación de mercados— y la *ciencia de producción* —investigación alineada con intereses industriales y comerciales (McCright & Dunlap, 2010; Schnaiberg, 1980)—. Esta distinción pone de manifiesto la forma en que las diferentes especialidades se entrelazan con intereses políticos y económicos, y destaca uno de los ámbitos más polarizantes de la política contemporánea: la redistribución económica y el tamaño del Estado (Webster & Abramowitz, 2017). En los entornos en línea contemporáneos, donde las personas se basan en heurísticas para procesar información (Chaiken, 1980; Metzger & Flanagin, 2013), los expertos y las fuentes científicas pueden servir como indicadores de credibilidad que moldean la percepción de credibilidad y la disposición a compartir información, con efectos potencialmente diferentes según grupos políticos (Rekker, 2021).

Más allá de preferencias políticas, esta tesis examina también si el razonamiento motivado las subyace. Investigaciones anteriores han operacionalizado el razonamiento motivado a través de las interacciones entre ideología política y medidas de conocimiento (Drummond & Fischhoff, 2017; Fischer et al., 2022; Savadori et al., 2022), con resultados inconsistentes. Partiendo de la idea de que la percepción de afirmaciones de expertos depende menos del *conocimiento* que de la familiaridad con los procesos de producción de conocimiento (Eyal, 2019), el razonamiento motivado se operacionaliza más bien a través de las interacciones entre identificación política, compromiso partidista y educación, captando tanto la motivación para defender las creencias grupales como la capacidad de procesar información de una manera que se alinee con dichas creencias (Kunda, 1990).

Dado que la imbricación entre ciencia y política es más marcada en algunos campos de investigación que en otros, el razonamiento motivado puede influir en las preferencias respecto a *tipos específicos* de ciencia, y por lo tanto se evalúa a través de la ciencia general, la ciencia de impacto y la ciencia de la producción.

Esta tesis también examina el papel de suposiciones individuales en el intercambio de información. Basándose en la teoría de la relevancia (Wilson & Sperber, 1995, 1998) y la teoría de la información (Shannon, 1948), establece una distinción entre *creencias* —proposiciones consideradas verdaderas (Egan, 1986)— y *expectativas* —predicciones sobre lo que es probable que ocurra en el futuro (Barto et al., 2013)—, conceptualizándolas como elementos que interactúan con la información a través de diferentes canales: contradicciones y refuerzos de creencias, y violaciones y validaciones de expectativas.

A lo largo del estudio se adopta un enfoque cuantitativo y multimétodo. Los datos de encuestas nacionales examinan las tendencias a nivel de la población en cuanto a percepciones de la pericia. Los experimentos de encuestas en línea aíslan los efectos causales de fuentes expertas y suposiciones individuales (Gaines et al., 2007). Los datos de redes sociales capturan comportamientos de intercambio de información a gran escala en línea (Edelmann et al., 2020). Aunque son muy adecuados para poner a prueba las expectativas teóricas, estos métodos reducen conceptos como la confianza, las percepciones de credibilidad, las creencias y las expectativas a indicadores numéricos que no pueden captar plenamente los significados culturales y las experiencias subjetivas implicadas en la forma en que los individuos interactúan con la pericia.

Resúmenes de los capítulos

Capítulo 1

A menudo se percibe a los conservadores estadounidenses como anticientíficos, y a los liberales como pro-científicos (Kennedy et al., 2022; Mann & Schleifer, 2020; McCright & Dunlap, 2011). Esta perspectiva pasa por alto los trabajos recientes que demuestran que los grupos políticos no rechazan ni apoyan la ciencia en general (Gligorić et al., 2025; Rekker, 2025; Shi et al., 2017), sino que sus posiciones dependen del ámbito de investi-

gación o de la afirmación en cuestión. El capítulo 1 examina cómo la coincidencia entre el ámbito de especialización de un experto y la ideología política de un individuo influye en la percepción de dichas afirmaciones, mediante un experimento en línea preregistrado realizado con adultos estadounidenses. En él, los participantes expresaron su confianza en expertos de los ámbitos de impacto y producción, y evaluaron la exactitud de afirmaciones atribuidas aleatoriamente a dichos expertos o que carecían de fuente.

Los resultados muestran que los liberales confían más en expertos que los conservadores, aunque la diferencia es menor en el caso de los expertos de producción. La atribución a un experto aumentó la precisión percibida de ciertas afirmaciones, independientemente de la ideología política del lector. Estos resultados sugieren *ciertas* preferencias ideológicas, pero no corresponden claramente con la distinción entre impacto y producción tal como se teorizó inicialmente (McCright & Dunlap, 2010; Schnaiberg, 1980).

Capítulo 2

El capítulo 2 evalúa el razonamiento motivado como factor de polarización científica en Europa. La identificación del razonamiento motivado requiere observar la combinación entre compromiso partidista y capacidad cognitiva (Kunda, 1990; Leeper & Slothuus, 2014): el primero proporciona la motivación para defender las posiciones de un grupo político, y la segunda, la capacidad de descartar mediante el razonamiento información científica discordante. A partir de datos de encuestas procedentes de 11 países europeos, estos elementos se operacionalizaron mediante interacciones entre identificación izquierda-derecha, proximidad a un partido y nivel de educación. Estas variables se evaluaron como predictoras de confianza en la ciencia en general, la ciencia de impacto y la ciencia de producción utilizando modelos de ecuaciones estructurales.

Los resultados muestran que los europeos alinean sus preferencias con la distinción entre ciencia de impacto y ciencia de producción (Rekker, 2025), y que la afinidad con un partido y el nivel de educación amplifican las diferencias entre izquierda y derecha *únicamente* en el caso de la ciencia de impacto. Esto indica que el razonamiento motivado orienta las diferencias ideológicas específicamente hacia la investigación que cuestiona intereses industriales y respalda la intervención regulatoria, lo cual es compatible con la idea de que es más probable que el razonamiento motivado surja cuando la imbricación entre la

especialización científica y la política es más marcada.

Capítulo 3

El capítulo 3 examina la interacción entre información y suposiciones del observador. Basándose en la teoría de la relevancia (Wilson & Sperber, 1995, 1998) y la teoría de la información (Shannon, 1948), establece una distinción entre creencias y expectativas, y aplica este marco para estudiar intenciones de compartir en el contexto de un experimento en línea preregistrado realizado con participantes estadounidenses. Los participantes fueron expuestos de manera aleatoria a afirmaciones de expertos que contradecían o confirmaban sus creencias, violaban o validaban sus expectativas o combinaban estas manipulaciones.

Las contradicciones con creencias aumentan el interés percibido en mayor medida que las violaciones de expectativas, pero el refuerzo de creencias no aumenta las intenciones de compartir información, a diferencia de la validación de expectativas. Por lo tanto, la difusión de información de expertos en redes sociales podría depender menos de lo que los usuarios consideran verdadero que de si esto corresponde con lo que ellos *anticipan*.

Capítulo 4

El capítulo 4 examina cómo los *índices de autoridad científica* (SACs en inglés), definidos como menciones de entidades que indican especialización o descubrimientos científicos, influyen en la difusión de información verificada por expertos en las redes sociales. Combina un análisis observacional de millones de tuits compartidos entre 2010 y 2023 —en los que se identificaron los SAC mediante un modelo lingüístico a gran escala y se clasificaron a los usuarios como Republicanos o Demócratas— con un experimento en línea preregistrado en el que los participantes evaluaron su disposición a compartir titulares atribuidos a entidades científicas o a ninguna fuente.

Los resultados muestran que atribuir afirmaciones a entidades científicas aumenta la disposición a compartir, en proporción a la precisión percibida. Sin embargo, el efecto sobre la intención de compartir es específico de las afirmaciones falsas entre los Republicanos, lo que sugiere que la autoridad científica se utiliza indebidamente para legitimar

la desinformación que atrae a los usuarios de derecha en línea.

Conclusión

Esta tesis ha examinado varios factores que determinan la forma en la que se perciben y difunden informaciones atribuidas a expertos. Los resultados muestran que los expertos suelen considerarse fuentes de información creíbles, lo que refuerza la percepción de la exactitud de afirmaciones fácticas. La controversia no se dirige contra la ciencia en general, sino más bien contra la forma en la que se utiliza en la práctica la autoridad experta y ciertas afirmaciones específicas. Esta distinción pone en tela de juicio la imagen comúnmente difundida de que los liberales serían pro-ciencia y los conservadores anti-ciencia. En Estados Unidos, las diferencias en la evaluación de pericia siguen más líneas partidistas, mientras que en Europa las preferencias dependen en mayor medida del tipo de ciencia evaluada, pareciendo que la impugnación está motivada y se concentra en torno a la especialización relacionada con la regulación. Estas tendencias ponen de manifiesto una tensión: la investigación científica es ampliamente valorada, pero es objeto de impugnación política cuando se trata de afirmaciones y expertos específicos.

Más que una crisis de la ciencia, esta tesis sugiere un contexto más paradójico en el que la pericia científica es muy valorada, pero en el que se cuestiona la autoridad experta. Esta controversia puede reflejar la toma de conciencia del público de que la producción de conocimiento es un proceso social y político, y que las afirmaciones pueden servir a intereses privados o utilizarse para justificar decisiones políticas. El escepticismo no siempre es un rechazo total, sino una resistencia frente a una autoridad indiscutible sobre los hechos, lo que refleja un grado de interacción sustancial, a veces crítico, con respecto a la pericia. Dicho esto, no todas las impugnaciones se basan en fundamentos equivalentes: el escepticismo respecto a las afirmaciones de los expertos sobre los riesgos sociales puede tener consecuencias colectivas, y los contraargumentos en los que se basa merecen igualmente ser examinados de manera crítica.