A cognitive perspective on trust in science

(PhD thesis)

Jan Pfänder

2025-05-26

Table of contents

Acknowledgements

There are many wonderful people in my life who made writing this thesis possible.

First and foremost, I have to thank my supervisor, Hugo Mercier. Without Hugo, I would never have been able to write this thesis—nor would I have started any thesis in the first place. When he agreed to supervise me as an intern during my Master's, I was completely new to cognitive science and completely clueless. What followed were one and a half years of failed experiments on a topic we never returned to. And yet, somehow, Hugo made this experience fun and intellectually stimulating. More than that, he gave me the feeling that my thoughts mattered. Hugo always made time for me, whether in weekly meetings or within-the-minute replies to my emails. I would never have imagined myself as a researcher if it weren't for Hugo's kindness, generosity, and encouragement. I still have a lot to learn from him, and I hope we'll have many opportunities to keep working together in the future.

I thank Sacha Altay, who is not only a friend but has also been an academic mentor to me. Sacha's down-to-earth attitude while producing great research is a constant inspiration. He also taught me to be more pragmatic (though I still have a long way to go).

I've been fortunate to be part of a team of brilliant and kind people at the Institut Jean Nicod. I couldn't have hoped for a more intellectually stimulating place to grow up, academically speaking. I'm grateful for the fun conversations, the valuable feedback, and the delicious pizza after lab meetings. In particular, I'd like to thank my co-authors Benoît, Sophie, and Lou for their excellent collaboration. I also have to single out Coralie Chevalier, who's class on psychology and public policy first inspired me to apply for a Master's in Cognitive Science. I'm also grateful for her trust in giving me the opportunity to design and teach my own course.

Thanks to Lilas, for being a great friend, faithful lunch companion, and making sure I didn't miss countless administrative requirements and deadlines.

I was incredibly lucky to have made many friends in Paris who helped take my mind off research when needed, in particular Amelie, Elie, Edgar, Simon and Fiona. I also want to thank all the amazing people I was lucky enough to play music with–Adrian, Daniel, Enora, Luis, Bella, Chippy, Kenta, Malin, Michael, Octave, Rémy, Pablo, Marina, Reynaldo. These sessions and concerts always lifted me up, no matter how frustrating the day had been.

The two people who made Paris feel like home are my flatmates, Enara and Adrian. I owe much of my sanity and my sense of belonging in Paris to them. Lazy TV nights with ice cream-induced stomachaches or concerts in squads–with them, any plan is a good plan. Over the last couple of years, I've spent more time with Adrian than with anyone else, and I've cherished every single moment. This is what I will most miss about Paris.

There are other friends who, despite the distance, are always present in my life. Leon, Pauli, Jonas, Luis, Jakob, Chris, Eva, and Basti have been around for as long as I can remember, and they mean the world to me. There are also my dear friends Joseph and Sylvie who I wish I would see more often.

I would like to thank my parents, who love me and always trust me to do the right thing. Neither of them got a university degree, yet they unconditionally supported me through all those years of education. I thank my dad, Thorsten, who always inspired me to think critically. I look up to him for his unshakable optimism and stamina. I thank my mum, Katja, for putting me on the right track in so many ways: for making me learn an instrument, for the countless hours she spent reading to me as a child, for helping me study vocabulary, for running a household with four kids while working. I'm incredibly grateful to have my siblings Luca, Jona, and Franca in my life, and for the moments of laughter and empathy we share whenever we're all home together.

Thanks to my grandma, Oma Brunhilde, for being living proof that people who occasionally distrust science can be the loveliest people. Thanks to Opa Erwin, for always taking care of us. And to Oma Renate and to Opa Horst, who never got tired of sharing his passion for history with me and would probably have been the single most proud person to see me finish a dissertation.

I was fortunate enough to have been adopted by a second family in Spain. I thank Consu and Miguel for their endless generosity and for giving me a sense of calm I've never experienced anywhere else. Thanks to Maria and Dani for always supporting me.

I owe special thanks to Carmen, for never getting tired of showing me all that is beautiful in life. I may have found more insight in the novels that she has offered me over the years than I could ever find in research.

1 Summary

Recognizing and addressing some of the pressing challenges we face as human society, including global health and climate change, requires trust in science. Philosophers of science have argued that people should trust science for its epistemic qualities–its capacity to produce accurate knowledge. Under this premise, the literature on public understanding of science has long sought to explain people's trust in science by their knowledge of it–with sobering results: While people do tend to trust science, they do not tend to know much about it. If not grounded in knowledge, is public trust in science mostly irrational? In this thesis, I argue that no, not necessarily. From a cognitive perspective, this thesis aims to provide an explanation of the foundations of trust in science at the micro-level. I develop a 'rational impression' account of trust in science, according to which people do not need to understand or remember much about science to trust it. The account builds on two basic cognitive mechanisms of information evaluation: First, if someone finds out something that is hard-to-know, we tend to be impressed by it, if we believe it is true. This impression makes us infer that the person is competent, a crucial component of trustworthiness. Second, if something is highly consensual, we tend to infer that it is likely to be true, and that those who agree are competent. These inferences from consensus are particularly relevant in the context of science, where most people lack relevant background knowledge to evaluate claims for themselves. Scientists agree on hard-to-know findings such as the size of the milky way or the atomic structure of DNA. Although most people do not understand much of how the scientists came to make these findings, nor remember the details of the findings, the consensus provides good reasons to trust the scientists. This account underlines the critical role of education and science communication in fostering trust in science.

This thesis is structured as follows: Chapter 2 lays out the motivation for this thesis, summarizes the rational impression account and situates it in the literature on trust in science.

In Chapter 3 and Chapter 4, to explain public trust in science, we lay out the foundations of the rational impressions account. In Chapter 3, we show that exposure to impressive science increases people's trust in science, but that people tend to almost immediately forget much of the content that generated this impression. In Chapter 4, we show that in non-science related contexts, where participants were deprived of relevant background knowledge, they inferred that informants–individuals providing answers on some question–who agree more with each other had more accurate answers and were more competent. Using simulations and analytical arguments, we argue that these inferences from convergence–the extent to which informants agree on a piece of information, the most extreme form of which is consensus–are rational, under a wide range of parameters, given that informants are independent and

unbiased. Participants took this into account: when given cues that the informants might be biased, participants' inferences from convergence were weakened.

In Chapter 5 and Chapter 6, we test two predictions that follow from the rational impressions account. In Chapter 5, we find that in a representative sample of the French population, trust in science–within and between different disciplines–was associated with perceptions of consensus and precision: the more precise and consensual people perceived science to be, the more they tended to trust it. According to the rational impression account, people who have received science education should have had the opportunity to form impressions of trustworthiness of science. This should have built a solid baseline of trust in science. In line with this prediction, in Chapter 6, we show that, in the US, almost everyone–even people who said they don't trust science in general or who held specific beliefs blatantly violating scientific knowledge (e.g. that the earth is flat)–trusted most basic science knowledge (e.g. that electrons are smaller than atoms). This finding suggests some conclusions about distrust in science: The fact that trust in basic science knowledge is nearly at ceiling for almost everyone suggests that those who nevertheless report not trusting science do so because they are driven by specific, partial rejections of science (e.g. climate change denial). Given the overwhelming trust in basic science, these rejections are likely to stem from motivations exogenous to science. This lends support to motivated reasoning accounts of science rejection.

The rational impression account for trust in science is built on the hypothesis that people tend to be good at evaluating information, by relying on mechanisms of epistemic vigilance. In Chapter 7, I explore another consequence of this hypothesis, which is that people should be good at judging the veracity of news. In a meta-analysis, we found that this was largely the case: people around the world were generally able to distinguish true from false news. When they erred, people were slightly more skeptical towards true news than they were gullible towards false news. We do not conclude from these results that all misinformation is harmless, but that people don't simply believe all misinformation they encounter–if anything, they tend to have the opposite tendency to not believe information, even if accurate. Based on this, we argue that if we are concerned about an informed public, we should not only focus on fighting against misinformation, but also on fighting for accurate information.

In Chapter 8, I discuss limitations of this thesis. I argue that the broader picture that emerges from the evidence presented in this thesis is that people come to trust science, and information more generally, based on mechanisms of information evaluation that are, on average, sound, and do, in most cases, work.

2 Introduction

2.1 Trust in science matters

Numerous studies have demonstrated that individuals with higher levels of trust in science are more likely to engage in pro-environmental behavior, support climate policies, and accept the scientific consensus on global warming (Cologna and Siegrist 2020; Hornsey et al. 2016; Bogert et al. 2024).

Trust in science has also been shown to be positively associated with willingness to get vaccinated, both in a Covid-19 context (Lindholt et al. 2021) and beyond (Sturgis, Brunton-Smith, and Jackson 2021).

During the pandemic, a panel study in 12 countries found that trust in scientists was the strongest predictor of whether people followed public health guidelines, such as mask-wearing or social distancing (Algan et al. 2021). Similar results have been found by other studies, e.g. positive effects of trust in science on acceptance of social distancing in the US (Koetke, Schumann, and Porter 2021).

In short, trust in science matters, because it is related to several more specific and desirable attitudes, beliefs and behaviors.

2.2 What is trust in science?

What exactly is trust in science? Answering this question is challenging, because trust in science involves many subtleties than can be grouped into two broad categories: subjective representation and domain-specificity. First, any general definition of trust in science necessarily glosses over variation in the public's representation of science: The term science can mean different things to different individuals or groups. It can be seen, for example, as a body of literature, an institution, a method, or certain individual scientists (Gauchat 2011). This diversity of possible representations can have consequences on people's reported trust in science. For instance, it has been shown that people tend to trust scientific methods more than scientific institutions—particularly among less-educated segments of the population (Achterberg, De Koster, and Van Der Waal 2017).

Second, trust is, at least to some extent, domain-specific. Domains can be, for example, scientific disciplines. In the US, people trust some disciplines, such as biology or physics, considerably more than others, such as economics or sociology (Altenmüller, Wingen, and Schulte 2024; Gligorić, Kleef, and Rutjens 2024; Gauchat and Andrews 2018).

Domains can also be particular pieces of knowledge. For certain contentious science topics, general trust can be a poor indicator of acceptance of scientific knowledge. For example, in 2014, 41.9% of Americans had a great deal of confidence in the scientific community, according to the US General Social Survey (GSS)[1]. A Pew study run in the same year found that 88% of a representative sample of scientists connected to the American Association for the Advancement of Science (AAAS) believed that it was safe to eat genetically modified foods–but only 37% of Americans shared that belief (Rainie and Funk 2015).

Domains can also be character traits of scientists. In social psychology, a popular model suggests that people evaluate others along two fundamental dimensions: competence and warmth (Cuddy, Fiske, and Glick 2008). Similarly, for trust in scientists, researchers have distinguished between an epistemological and an ethical dimension (Wilholt 2013; Intemann 2023). Sometimes, researchers make more fine-grained distinctions: For example, Hendriks, Kienhues, and Bromme (2015) have argued for three dimensions: expertise/competence, integrity, and benevolence. Besley, Lee, and Pressgrove (2021a) has suggested openness as an additional fourth dimension. Across these dimensions, competence is typically the one on which scientists score highest in the perception of the public. A recent Pew survey found that 89% of Americans viewed research scientists as intelligent, but only 65% viewed them as honest, and only 45% described research scientists as good communicators (Kennedy and Brian 2024; see also Fiske and Dupree 2014). Beyond the US, a recent study confirmed this tendency on a global scale (Cologna et al. 2025): People perceived scientists as highly competent, with 78% tending to believe that scientists are qualified to conduct high-impact research. By contrast, people held scientists in lower esteem with regards to their integrity and benevolence: Only 57% of people tended to believe that most scientists are honest, and only 56% tended to believe that most scientists are concerned about people's well-being.

Finally, domains can also be areas of legitimacy. The fact that people tend to evaluate scientists in particular as competent already suggests that they see scientists as legitimate in their role of knowledge producers. However, this legitimacy does not necessarily extent to other roles, in particular as policy advocates: According to the just before mentioned Pew survey, only 43% of Americans think scientists are usually better than other people at making good policy decisions on scientific issues (Kennedy and Brian 2024). However, the evidence on the public opinion about scientists in the role of policy advocates is not conclusive: A recent global study found that people tend to agree that scientists should engage with society and be involved in policymaking (Cologna et al. 2025). In the context of climate science, Cologna et al. (2021) found that in the US, and Germany, people support policy advocacy by climate researchers and expect greater political engagement of the scientists.

[1] Number based on own calculations using publicly available of the US General Social Survey (GSS). Data provided by NORC and available here: https://gss.norc.org/us/en/gss/get-the-data.html

Although trust in science can be domain-specific and depends on one's subjective representation of science, empirically, general trust in science appears to be a meaningful concept. First, while index measures of trustworthiness with different dimensions (e.g., competence, benevolence etc.), are preferable to direct, single-item measures of general trust (Besley and Tiffany 2023), such as the Pew research centers' question "how much confidence, if any, do you have in scientists to act in the best interests of the public?", in their recent global study on trust in science, Cologna et al. (2025) report that the Pew's single item question was highly correlated with an index measure based on several trustworthiness dimensions. Second, as reviewed earlier, general trust in science, whether based on a single-item or an index measure, appears to be a relevant predictor of other more tangible outcomes, from vaccine intentions to beliefs about climate change.

A widely agreed-upon definition of trust is the willingness to be vulnerable to another party–whether an individual, a group, or an institution (Mayer, Davis, and Schoorman 1995; Rousseau et al. 1998). Here, I adopt a definition that builds on this idea, by defining trust in science as "one's willingness to rely on science and scientists (as representatives of the system) despite having a bounded understanding of science" (Wintterlin et al. 2022, 2).

To which extent do people rely on science and scientists? And why? In the following section, I show that despite a research focus on distrust, most people do tend to trust science. However, I will argue that the classic normative reason on why people should trust science–namely understanding and knowledge of science–is not likely to be the main explanation.

2.3 The puzzle of trust in science

2.3.1 People tend to trust science

On the whole, people across the world tend to trust science: A recent large survey in 68 countries found trust in scientists to be "moderately high" across all countries (mean = 3.62; sd= 0.70; Scale: 1 = very low, 2 = somewhat low, 3 = neither high nor low, 4 = somewhat high, 5 = very high), with not a single country below midpoint trust (Cologna et al. 2025).

In the US, where long term data is available from the US General Social Survey (GSS), this public trust appears to be both remarkably stable and elevated relative to other institutions (Funk and Kennedy 2020; Funk et al. 2020; T. W. Smith and Son 2013): From the early 1970s to 2022, the currently latest year available in the GSS, on average 40% of Americans say they have a great deal of confidence in the scientific community (see Figure 2.1). This is the second highest score (just behind the military) among 13 institutions listed in the GSS, including, e.g., the government, press, organized religion or medicine. Note, however, that very recently, polls suggest a drop in trust in science in the US (Lupia et al. 2024).

Global data on trust in science across time is sparse. Yet, the available data suggests, if anything, a recent increase of trust in science: In 2018, the Wellcome Global Monitor (WGM)

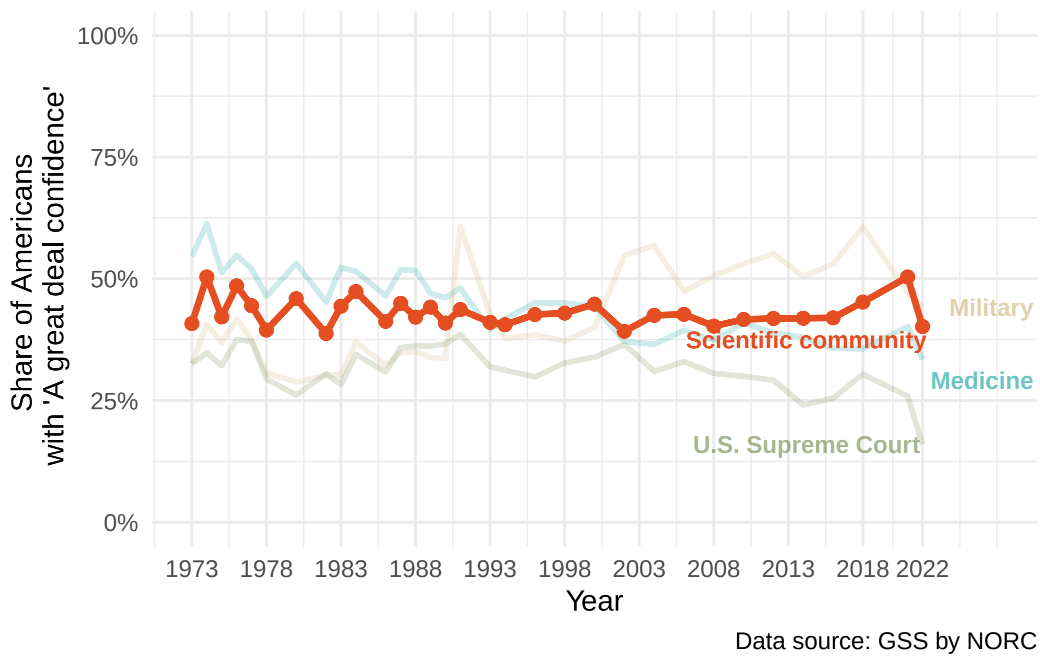

Figure 2.1: *Top four trusted institutions in the US.* The plot shows the variation of confidence in the top four trusted institutions in the US over time. The data for this plot is from the cumulative file of the US General Social Survey (GSS).

surveyed of over 140000 people in over 140 countries on trust in science (Wellcome Global Monitor 2018). In 2020, during the first year of the Covid pandemic and before vaccines were widely available, a follow-up survey was run in 113 countries, involving 119000 participants (Wellcome Global Monitor 2020). Between these two surveys, on average, trust in science has risen (Wellcome Global Monitor 2021): In 2020, 41% (32% in 2018) of respondents said they trust science a lot, 39% (45% in 2018) said they trust science to some extent, 13% (also 13% in 2018) said they trust science "not much or not at all", and 7% (10% in 2018) answered "don't know" (Figure 2.2).

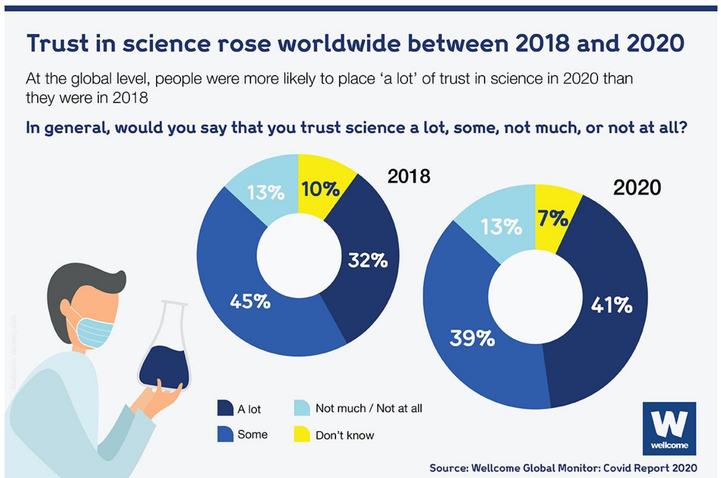

Figure 2.2: Graphic from Wellcome Global Monitor (2020)

2.3.2 Why should we trust science?

According to some classical views in epistemology, trust should in fact play no role in science: In a way, the whole point of science is to "know the truth instead of just trusting what you are told" (Hendriks, Kienhues, and Bromme 2016). Some early scholars, e.g. John Locke, saw the "uncritical acceptance of the claims of others [...] as a failure to meet rationality requirements imposed on genuine knowledge" (Sperber et al. 2010, 361). This perspective is best understood historically, as an argument against defiance to authority (Sperber et al. 2010). However, this ideal of the individual as a radically skeptical, self-reliant reasoner is at odds with reality. Most objects of scientific study are impossible for the public to observe and judge for themselves (e.g., atoms, genes). And even scientists fundamentally need to trust each other: Modern science is deeply collaborative and specialized, to the extent that no individual scientist independently verifies all the claims they rely on. This is not only true for empirical science–even in mathematics, there are proofs that most mathematicians could not verify for themselves (Hardwig 1991).

Because explicit verification of everything is impossible for any individual, there must be some features to science that make it trustworthy. What are these features? Philosophers and sociologists have long argued over this question (for an overview, see Oreskes 2019). Famously, Karl Popper emphasized the role of the scientific method–the hallmark of which, for him, was falsifiability. Scientific knowledge, in Popper's view, is trustworthy because it constantly puts itself to test, through bold conjectures and rigorous attempts at refutation (Popper 2002).

Popper's conception of the scientific method, and more generally the very idea of any singly universal method for trustworthy science, has been widely criticized (Oreskes 2019). Moving beyond a purely method-based view, Robert Merton for example rooted science's trustworthiness in its institutional norms (Merton 1979). Merton layed out four defining norms of science, while acknowledging that they were often more ideals than empirical descriptions: communism, i.e. findings are shared publicly rather than kept secret or privatized; universalism, i.e. claims are evaluated based on impersonal criteria, not the identity (gender, race, nationality, or status) of the person making them; disinterestedness, i.e. scientists are (ideally) motivated by the pursuit of knowledge, not personal gain; and organized skepticism, i.e. claims are constantly subject to critical scrutiny and testing.

Both Popper and Merton idealized, to some extent, the nature of scientists: Although Popper highlighted the (impersonal) benefits of the (singular, in his view) scientific method, executing this method required heroic individual scientists who make bold predictions and put their own claims to the greatest scrutiny. Merton focuses on a shared ethos of the scientific community, but this implies that individual scientists' have internalized the common value ideals and are committed to them. By contrast, Strevens (2020) assumes that scientists can be, just like other humans, biased and ideological. However, in science, they ultimately need to convince other scientists by presenting evidence. This, "iron rule of explanation", in Strevens' view, is what makes science trustworthy.

This very short and incomplete review of normative perspectives on trust in science is meant to illustrate that even very different accounts on what should make science trustworthy have something in common: Popper, Merton and Strevens highlight different epistemic qualities of science, but they share the underlying assumption that, to (rationally) trust science, one needs to know about these qualities.

2.3.3 People do not know much about science

This premise, that knowledge and understanding of science should lead to trust in–and more generally positive attitudes towards–science, has dominated research on public understanding of science (Bauer, Allum, and Miller 2007). Studies testing the relationship between science knowledge and attitudes have mostly relied on versions of what is known as the "Oxford scale" (Table 2.1[2]) to measure science knowledge–a set of specific true/false or multiple-choice

[2]The original scale proposed by Durant, Evans, and Thomas (1989) comprised 20-items. Several reduced versions of this have been used in different survey projects and studies since then. For example, Miller (1998) reports that a Eurobarometer in 1992 included only 9, and a Science and Engineering Indicator survey in 1995 only 10 of these items. Gauchat (2011) reports relying on a 14-item scale based on several Science and Engineering Indicators surveys included in the GSS (but does not provide an overview of the

questions about basic science facts.[3] As the name suggests, this measure was developed by scholars associated with the University of Oxford in the late 1980s, in collaboration with Jon D. Miller, a pioneer in research on science literacy in the US (Durant, Evans, and Thomas 1989). A version of these items has been used in several large-scale survey projects, including several Eurobarometer surveys, as well as in the National Science Foundation's (NSF) "Science and Engineering Indicators" surveys in the US (National Academies of Sciences, Engineering, and Medicine 2016). Survey results from the Oxford scale have been taken to showcase a science knowledge deficit among the public. For example, from the first surveys in which they were used, Durant, Evans, and Thomas (1989) (p.11) report that only "34% of Britons and 46% of Americans appeared to know that the Earth goes round the Sun once a year, and just 28% of Britons and 25% of Americans knew that antibiotics are ineffective against viruses". According to the National Academies of Sciences, Engineering, and Medicine (2016), performance on the Oxford scale items in the US has been "fairly stable across 2 decades" (National Academies of Sciences, Engineering, and Medicine 2016, 51)[4].

The results of this research program have been overall sobering-not only because Oxford scale knowledge was rather low in the population, but also because it did not seem to explain much variation in people's attitudes towards science: In a seminal meta-analysis Allum et al. (2008) found that Oxford scale type science knowledge was only weakly associated with attitudes towards science. Moreover, this weak association only held for general attitudes–for attitudes towards specific contentious science topics (e.g., climate change), the study did not find an association.

Table 2.1: An 11-item version of the Oxford-scale, as reported in a comprehensive review of the literature on scientific literacy (National Academies of Sciences, Engineering, and Medicine 2016).

1	The center of the Earth is very hot. (True)
2	The continents on which we live have been moving their locations for millions of years and will continue to move in the future. (True)
3	Does the Earth go around the Sun, or does the Sun go around the Earth? (Earth around Sun)
4	How long does it take for the Earth to go around the Sun? (One year)*
5	All radioactivity is man-made. (False)
6	It is the father's gene that decides whether the baby is a boy or a girl. (True)
7	Antibiotics kill viruses as well as bacteria. (False)
8	Electrons are smaller than atoms. (True)

included items).

[3] The meta-analysis included two types of studies: one that measured general scientific knowledge, and one that measured knowledge specific to biology and genetics

[4] Note that for this trend scale, the National Science Board who publishes results from the Science and Engineering Indicators selected a subset of 9 of the items shown in Table 2.1.

9	Lasers work by focusing sound waves. (False)
10	Human beings, as we know them today, developed from earlier species of animals. (True)
11	The universe began with a huge explosion. (True)

*Only asked if previous question was answered correctly.

The Oxford scale has received various forms of criticism, ranging from minor concerns for being–ironically, in light of the bad performance–too easy to answer (Kahan 2015), to the fundamental critique that it only captures factual recall (Pardo and Calvo 2004; Bauer, Allum, and Miller 2007). In line with the normative literature reviewed here earlier, what should actually matter for trust is a different kind of knowledge, namely an institutional and methodological understanding of how science works.

The literature on science literacy has sought to measure how well people understand science. Even the creators of the Oxford scale were, to some extent, aware of the limitations of using factual knowledge alone(National Academies of Sciences, Engineering, and Medicine 2016). They never intended the Oxford scale items to serve as a comprehensive measure of science literacy. In fact, Durant, Evans, and Thomas (1989) also developed a scale of "understanding of processes of scientific inquiry"–several multiple-choice about the scientific method and basic concepts of probability. Miller (2004) viewed a scientifically literate citizen as someone who has both a "(1) a basic vocabulary of scientific terms and constructs; and (2) a general understanding of the nature of scientific inquiry." His measure of science literacy included open-ended questions, for example on what people understand as the meaning of scientific study (Miller 1998).

This institutional knowledge and understanding of science has been tested less extensively for its association with attitudes towards science (but see e.g., Weisberg et al. 2021), as was the case for Oxford scale knowledge, likely because understanding via, for instance, open-ended questions is less obviously quantifiable. However, in isolation, results of measures intended to capture an understanding of science have hardly drawn a more positive image of the public's knowledge of science. Using an index of various measures, Miller (2004) (p. 288) concluded that "approximately 10 percent of US adults qualified as civic scientifically literate in the late 1980s and early 1990s, but this proportion increased to 17 percent in 1999". Miller described that according to his measure, someone qualifies as scientifically literate if they possess "the level of skill required to read most of the articles in the Tuesday science section of The New York Times, watch and understand most episodes of Nova, or read and understand many of the popular science books sold in bookstores today" (Miller 2004, 288). These low scientific literacy rates stand in sharp contrast to the rather elevated and stable levels of trust in science in the US (Figure 2.1).

More recent data suggest that science literacy in the US may have improved slightly since Miller's assessment during the early 2000s, but it still remains rather low. Based on results from the 2018 US Science & Engineering Indicators, Scheufele and Krause (2019) (p. 7663) report that "one in three Americans (36%) misunderstood the concept of probability; half of the population (49%) was unable to provide a correct description of a scientific experiment; and three in four (77%) were unable to describe the idea of a scientific study."

The 2024 US Science & Engineering Indicators, based on data from the Pew Research Center's American Trends Panel (ATP) from 2020, report that "60% of U.S. adults could correctly note that a control group can be useful in making sense of study results" and that "only half of U.S. adults (50%) could correctly identify a scientific hypothesis" (National Science Board, National Science Foundation 2024, 24). The report also suggests a positive association between people's knowledge on these two questions and their trust in scientists to act in the best interest of the public[5]. However, this association appears rather small, given how fundamental these questions are: Among those who accurately responded that assigning a control group is useful, 44% also expressed a great deal of confidence, compared to 32% of those who did not think a control gorup was useful.

The here reviewed evidence on the relatively weak link between knowledge of and attitudes towards science comes with three limitations: First, direct evidence is based mostly on narrow, factual knowledge as measured by the Oxford scale, and not on a broader, institutional understanding of science. Second, the relevant studies have typically assessed attitudes towards science more broadly, rather than trust in science in particular. Third, the evidence is–to a large extent–based on data from the US and, to some extent, Europe. A recent global study addresses these limitations–in particular the second and third–and points towards a similar conclusion: Cologna et al. (2025) tested the relationship between national science literacy scores, based on the Program for International Student Assessment (PISA), and national average trust in scientists for the 68 countries included in their study. They found no statistically significant association.

This is the puzzle of trust in science that I take as a starting point: Why do people, to a large extent, trust science, despite not knowing or understanding much about it?

2.3.4 Existing explanations focus on why people distrust in science

The idea that knowledge about science causes positive science attitudes is best known under the term "deficit model", because much of the literature attested the public "depressingly low levels of scientific knowledge" that were assumed to be the principle cause of negative attitudes

[5]The report does not test for whether this association is statistically significant. The exact wording from the report is: 'In those data, accurate understanding of the scientific process is positively associated with respondents' expression of "a great deal' of confidence in scientists to act in the public's best interests. For example, among those who accurately responded that assigning a control group to not receive medication would be a useful way to test whether a medication works, 44% also expressed a great deal of confidence in scientists to act in the best interests of the public. By comparison, a lower percentage (32%) of those who did not correctly identify the value of a control group—i.e., those who did not demonstrate understanding of experimental logic—expressed such confidence. In addition, approximately 47% of respondents who knew what a hypothesis is expressed a great deal of confidence in scientists. By contrast, only 31% of those who did not demonstrate knowledge of what a hypothesis is expressed a great deal of confidence in scientists 'to act in the best interests of the public' "

towards science (Sturgis and Allum 2004, 56). The deficit model has been widely criticized for idealizing science: for implying that "to know science is to love it" (Bauer, Allum, and Miller 2007) and for portraying science knowledge as "superior to whatever 'nonscientific' or 'local' knowledge the public may (also) possess" (Gauchat 2011). The literature has since moved beyond the idea of science knowledge as the sole driver of trust in science.

Researchers have increasingly studied how people's values, world views, and identities shape their attitudes towards science (Hornsey and Fielding 2017; Lewandowsky and Oberauer 2021). The psychological literature has focused on explaining negative attitudes towards science with motivated reasoning–selecting and interpreting information to match one's existing beliefs or behaviors (Lewandowsky and Oberauer 2016; Hornsey 2020; Lewandowsky, Gignac, and Oberauer 2013). This research mostly suggests that certain psychological traits, such as a social dominance orientation, or a tendency of engaging in conspiracy thinking, lead people to reject science. Arguments on a general conspiratory thinking style as one of the root causes of science rejection shift the debate, to some extent, from a knowledge deficit to a broader reasoning deficit (Hornsey and Fielding 2017; Rutjens and Većkalov 2022)[6].

Motivated reasoning accounts have been popular in light of accumulating evidence in the US for a widening partisan gap regarding trust in science, with Republicans trusting less and Democrats trusting more (Gauchat 2012; Krause et al. 2019; Lee 2021). Contrary to what the deficit model would suggest, some influential work on motivated reasoning has shown that partisan-driven rejection of science does not appear to be the result of a lack of cognitive sophistication: Kahan et al. (2012) have shown that greater science literacy was associated with more polarized beliefs on climate change. Drummond and Fischhoff (2017) have extended these findings to other controversial science topics, namely stem cell research and evolution: they show that both greater science literacy and education are associated with more polarized beliefs on these topics. However, this phenomenon that "people with high reasoning capacity will use that capacity selectively to process information in a manner that protects their own valued beliefs" (Persson et al. 2021, 1), known under the term 'motivated numeracy', has largely failed to replicate (Persson et al. 2021; Hutmacher, Reichardt, and Appel 2024; Stagnaro, Tappin, and Rand 2023). Other studies, focusing on the case of climate change, have argued that partisan divides might simply be the result of bayesian information updating, rather than motivated reasoning (Bayes and Druckman 2021; Druckman and McGrath 2019).

Since the Covid-19 pandemic, the role of misinformation in fostering distrust in science has been increasingly studied (National Academies of Sciences 2024; Scheufele and Krause 2019; Druckman 2022). For this literature, by contrast with the deficit model, the problem of trust in science is less a lack of information, and more the abundance of harmful information.

The premise of the deficit model was that people would trust science because of their knowledge

[6]Note that in these studies, conspiracy thinking is conceived of as a general psychological trait, i.e. a general a way of thinking. For a distinction between general conspiracist worldviews and conspiracy beliefs about science specifically, see Rutjens and Većkalov (2022).

and understanding of science. Perhaps in response to this, recent literature in science communication has shifted the focus on source-based evaluations of trust: from science knowledge to perceptions of scientists. This work has established that people evaluate scientists along different dimensions (Intemann 2023), including competence, but also also integrity, benevolence or openness (Hendriks, Kienhues, and Bromme 2015; Besley, Lee, and Pressgrove 2021a). This literature suggests that, for enhancing trust in science, the latter, warmth-related (i.e. other than competence) dimensions could be particularly relevant (Fiske and Dupree 2014). The idea is that people already perceive scientists as very competent, but not as very warm, thus offering a greater margin for improvement.

On the whole, the literature on public trust in science focuses not on explaining why people do trust science, but on why they do not, or not enough, and how to foster trust. As reviews of science-society research have noted, the literature has traditionally operated (Bauer, Allum, and Miller 2007), and according to some still has a tendency to operate (Scheufele 2022), in "deficit" paradigms—at times focusing on the public's lack of scientific knowledge, and at other times on its lack of trust in science.

2.4 The rational impression account of trust in science

The previous section has established that public trust in science is relatively high, but that, contrary to what normative accounts would predict, knowledge and understanding of science do not seem to be very powerful determinants of this trust. Here, I try to make sense of this by taking a cognitive perspective: I develop a 'rational impression' account of trust in science, according to which people trust science because they have been impressed by it. This impression of trust persists even after knowledge of the specific content has vanished. The account builds on two basic mechanisms of information evaluation: First, if someone finds out something that is hard-to-know, we tend to be impressed by it, if we deem it true. This impression makes us infer that the person is competent, a crucial component of trustworthiness. However, it is hard or even impossible to recall exactly how we formed this impression. Second, if something is highly consensual, it is likely to be true. This is particularly relevant when people lack relevant background knowledge to evaluate claims for themselves, as is often the case in science. I begin by reviewing existing evidence for this account. Then I present the evidence contributed as part of this thesis.

2.4.1 Existing evidence for the rational impression account

2.4.1.1 The role of competence

As noted earlier, it has been suggested that targeting the perceived warmth of scientists might be particularly fruitful for science communication to enhance public trust, as people already perceive scientists as very competent, but not so much as warm (Fiske and Dupree 2014).

This might indeed be an effective strategy for science communication–at the same time, the fact that many people already trust science even without perceiving scientists as particularly warm highlights the central importance of perceived competence in explaining existing public trust in science.

How could people come to view scientists as competent? It seems plausible that most people do not have much first-hand evidence to judge scientists' character. As for personal contact, there are relatively very few scientists in the world, and most people probably do not know any personally. As for contact via the media, news on science, by contrast, for example, with news on politicians, mostly concern the science, not the scientists. As people only consume very little news in general (Newman et al. 2023), the bulk of exposure to science can be assumed to happen during education.

Taken together, it seems plausible that people make inferences about the trustworthiness of science based on scientific content and that these inferences primarily relate to scientists' competence. But if this process is supposed to happen mostly during the years of education, how is this compatible with the observation that people, even those who received a science education, do not seem to know much about science?

2.4.1.2 The role of education

The fact that knowledge and understanding of science are only weakly correlated with trust in science is puzzling not least because of the positive effects of education on trust in science: Education, and in particular science education, has been consistently identified as one of the strongest correlates of trust in science (Noy and O'Brien 2019; Wellcome Global Monitor 2018, 2020; but see Cologna et al. 2025 who only find a small positive relationship between tertiary education and trust in science). How to square the strong association with education, and the weak association with knowledge? One interpretation is that, if we assume that education has some causal effect on trust in science, this effect might be driven by something else than a pure transmission of knowledge and understanding (for a similar argument, see Bak 2001). The candidate mechanism I develop here is impression generation: Education exposes students to scientific content. Students might not understand much of it, and potentially recall even less later on; but they might have been impressed by it, to the point that they come to perceive scientists as competent, and thus, everything else equal, as trustworthy. This impression might persist even when specific knowledge vanishes.

2.4.1.3 People's impressions can be detached from knowledge

We commonly form impressions of the people around us while forgetting the details of how we formed these impressions: If a colleague fixes our computer, we might forget exactly how they fixed it, yet remember that they are good at fixing computers. As an extreme example, patients with severe amnesia can continue to experience emotions linked to events they could not recall (Feinstein, Duff, and Tranel 2010). In the context of science, Liquin and Lombrozo (2022) have shown that while people find some science-related explanations more satisfying than others, this did not predict how well they could recall the explanations shortly after, suggesting that impressions and knowledge formation can be quite detached.

What makes scientific findings impressive? For an information to be impressive, at least two criteria should be met: (i) it is perceived as hard to uncover and (ii) there is reason to believe it is true. By (i), I mean information that most people would have no idea how to find out themselves. In the case of science, this can be mostly taken for granted. For example, most people would probably only be mildly impressed by someone telling them that a given tree has exactly 110201 leaves. Even though obtaining this information implies an exhausting counting effort, everyone in principle knows how to do it. By contrast, finding out that it takes light approximately 100,000 years to travel from one end of the Milky Way to the other is probably impressive to most people, as they would not know how such a distance can be measured. Regarding (ii), I argue that the main relevant cue for the case of science is consensus.

2.4.1.4 People infer accuracy from consensus

In order to make the best of communicated information, animals need to be able to evaluate it, i.e. being able to distinguish inaccurate and harmful from accurate and beneficial information (Maynard-Smith and Harper 2003). It has been argued that humans have evolved a suite of cognitive mechanisms to serve this function (Sperber et al. 2010; Mercier 2020). In particular, we rely on cues of an informant's trustworthiness, and check the plausibility of an information against our background knowledge.

In the case of science, reliable cues and background information are scarce: As I have argued above, people generally have little first-hand information to evaluate individual scientists' trustworthiness. People also largely lack relevant background knowledge to evaluate the plausibility of scientific findings. Sometimes, to a certain extent, people might be able to judge the accuracy of scientific findings for themselves, for example when they are exposed to accessible and convincing explanations in school (Read and Marcus-Newhall 1993; Lombrozo 2007; for a review, see Lombrozo 2006). But for most scientific research, people cannot possibly evaluate the quality of the information for themselves, let alone make their own observations (e.g. quantum mechanics, genes).

An additional way to evaluate whether something is true or not is by aggregating opinions. It has been shown that, when no better information is available, people rely on majority heuristics: the more others agree on something, the more likely we are to believe them to be right (Mercier and Morin 2019). Literature on the wisdom of crowds has shown that this inference is often appropriate (see e.g., Hastie and Kameda 2005). In the absence of other reliable cues and background knowledge, inferences from consensus are likely to be given greater weight in a cognitive system of epistemic vigilance. In the case of science, this extra weight should play in favor of science's perceived trustworthiness: It has been argued that, by contrast with other intellectual enterprises, consensus is the defining trait of science (Collins 2002). Not only do scientists agree on things, but they agree on impressive things–things that would be impossible for any individual to ever uncover for themselves.

There is some suggestive evidence that people infer accuracy from consensus in the case of science. Research has demonstrated 'consensus gaps' for many science topics: gaps between the scientific consensus and public opinion (Rainie and Funk 2015). Some of the most worrying consensus gaps relate to climate change, as substantial segments of the population disagree with scientists on what is happening and what to do about it (e.g., Egan and Mullin 2017). According to a popular psychological model, the "gateway model", informing people about the scientific consensus on specific issues acts as a gateway to change their beliefs on these issues (Linden 2021). Studies have demonstrated the effectiveness of consensus messaging in changing people's beliefs on contentious science topics such as climate change (Većkalov et al. 2024) or vaccination (Salmon et al. 2015; for an overview of results on vaccination, climate change, and genetically modified food, see Van Stekelenburg et al. 2022). This evidence is only suggestive, because because the fact that consensus changes people's beliefs or attitudes does not necessarily require an inference of accuracy. An alternative explanation, for example, is normative conformity-that is, when people follow the majority because of social pressure rather than a belief that the majority is correct (Mercier and Morin 2019).[7] The literature on consensus messaging also does not provide evidence on whether learning about scientific consensus changes not only beliefs on specific science topics, but also increases trust in science more generally.

2.4.2 Novel evidence for the rational impression account

2.4.2.1 People (in the UK) trust but forget impressive science

Existing evidence suggests that impressions can be detached from remembering specific content. However, to the best of my knowledge, it has not yet been tested whether exposure to impressive science increases trustworthiness, and whether this process requires or not an ability of being able to recall the content that generated these impressions. In Chapter 3, we provide such a test. We found that reading about impressive scientific findings in the disciplines of archaeology and entomology increased participants' perceptions of both the scientists' competence and the trustworthiness of their discipline. At the same time, participants forgot almost immediately about the specific content that generated these impressions.

In these experiments, participants had reason to trust the information they saw for two reasons. First, while we did not provide explicit cues on consensus, the scientific findings were implicitly formulated as consensual (e.g. "Archaeologists are able to tell…"). This probably matches people's priors on science, as the bulk of people's exposure to science is likely to happen

[7] However, an accuracy inference seems to be the more plausible mechanism here: Studies on consensus messaging do not seem to be settings of high social pressure that we might expect to produce instances of normative conformity, compared to, for instance, the famous Asch experiments (Asch 1956).

during education, and schoolbook science knowledge is typically highly consensual. Second, as illustrated above, people already believe scientists to be competent, and this general prior is likely to extend to archaeologists and entomologists.

2.4.2.2 People (in the UK) infer both accuracy and competence from consensus

To test whether people infer accuracy and competence purely from consensus, in Chapter 4, we used non-science related experimental scenarios.

Participants were presented with answers by fictional individuals–the informants–on problems they knew nothing about other than the broad context (e.g., results of a game, or stock market predictions). While this created a slightly artificial experimental setting, it allowed us to ensure that participants did not have any priors, by contrast with previous studies, in which participants often had cues on the informants' competence. Participants were asked to judge the accuracy and competence of informants. We manipulated convergence, i.e. the extent to which different informants agreed on something, the most extreme form of which was consensus, where all informants agreed on exactly the same answer. Our experiments showed that, in the absence of any priors, participants indeed inferred accuracy from convergence; more than that, participants even inferred competence.

Are these inferences sound? The literature on the wisdom of crowds does not provide normative grounds that would justify these inferences: First, part of the beauty of wisdom of crowds phenomena is that they do not require much individual competence. For example, in the case of averaging, what defines the wisdom of the crowd is that the average opinion of a set of individuals is, in a wide range of circumstances, more likely to be accurate than that of a single individual (e.g. Larrick and Soll 2006). Second, a minimum of competence is sometimes posited as a requirement for some wisdom of the crowd effects to emerge. That is the case, for example, in the Condorcet Jury Theorem, where jurors need to be at least slightly better than chance for the majority decision to be correct (De Condorcet 2014).

To fill this gap, in Chapter 4, we present analytical arguments and simulations which suggest that inferences from convergence to both accuracy and competence are in fact sound, under a wide range of circumstances–assuming that informants are independent and unbiased. If truly rational, these inferences should not be general behavioral rules; they should be sensitive to contextual information and only work well within these boundary conditions. In line with that, our experiments show that inferences from convergence were weakened when participants were given reason to believe that the informants were biased by a conflict of interest.

2.4.2.3 People (in France) tend to trust scientists more when they perceive their work as consensual and precise

The experimental work on inferences from convergence is rather abstract and quite detached from public perceptions of science. In Chapter 5, we provide correlational evidence that fits predictions from this work for the case of science: We asked a representative sample of the French population about their trust in scientists in general, as well as their trust in scientists of different disciplines (Biology, Physics, Climate science, Economics, Sociology). We also asked participants about their impressions of science in general and the different disciplines in terms of precision and consensus. We found that trust in scientists–within and between different disciplines, as well as for science in general–was associated with perceptions of consensus and precision: the more precise and consensual people perceived the science to be, the more they tended to trust the scientists.

2.4.2.4 People (in the US) trust almost all of basic science

According to the rational impression account, people who have received science education should have had the opportunity to form impressions of trustworthiness of science. This should have built a solid baseline of trust in science. People might deviate from this default and distrust science on certain specific science topics for other reasons, but they should trust most of science. In Chapter 6, we put this prediction to test. We found that, in the US, almost everyone–even people who say they don't trust science in general or who hold specific beliefs blatantly violating scientific knowledge (e.g. that the earth is flat)–trusts almost all of basic science knowledge (e.g. that electrons are smaller than atoms). This suggests that basic trust in science is even higher than what large-scale surveys measure when they assess general trust in science. Our findings also support motivated reasoning accounts of science rejection: The fact that trust in basic science knowledge is nearly at ceiling for almost everyone suggests that those who nevertheless report not trusting science do so because they are driven by specific, partial rejections of science (e.g. climate change denial). These rejections, in light of the overwhelming trust in basic science, are likely to stem from motivations exogenous to science.

2.5 Other research included in this thesis

The rational impression account for trust in science is built on the hypothesis that people are good at evaluating information, by relying on mechanisms of epistemic vigilance (Sperber et al. 2010; Mercier 2020). In Chapter 7, I explore another consequence of this hypothesis, which is that people should be good at judging the veracity of news. In a meta-analysis, we found that this was largely the case: people around the world were generally able to distinguish true from false news. When they erred, people were slightly more skeptical towards true news than they were gullible towards false news. We do not conclude from these results that all misinformation is harmless, but that people don't simply believe all misinformation they encounter–if anything, they tend to have the opposite tendency to not believe information, even if accurate. Based on this, we argue that if we are concerned about an informed public, we should not only focus on fighting against misinformation, but also on fighting for accurate information.

3 Trusting but forgetting impressive science

Trust in science is associated with significant outcomes such as intent to vaccinate and belief in climate change. Around the globe, most people trust science at least to some extent. However, the causes of this trust aren't well understood. Here we propose a 'rational impression' model of trust in science. In this model, people trust scientists because they are impressed by their findings, and this impression of trust persists even after knowledge of the specific content has vanished. We present evidence for this model in two experiments (total n = 696) with UK participants. In Experiment 1, impressive scientific findings lead participants to think of the scientists as more competent and their scientific discipline as more trustworthy. In Experiment 2, we show that participants have these impressions despite forgetting the content that generated them. The rational impression model can explain why people trust science without remembering much of it. It also stresses the relevance of science communication.

available as a preprint here:

Pfänder, J., Rouilhan, S. D., & Mercier, H. (2025). *Trusting but forgetting impressive science.* https://doi.org/10.31219/osf.io/argq5_v1
For supplementary materials, please refer to the preprint.

3.1 Introduction

Jon D. Miller, a pioneer in measuring science literacy, found that at the end of the 20th century, approximately 17% of Americans qualified as scientifically literate–a measure he described as the "level of skill required to read most of the articles in the Tuesday science section of The New York Times" (Miller 2004). His estimate for the early 1990s had been even lower, around 10%. By contrast, the level of trust in science in the US has been relatively high and remarkably stable (Funk and Kennedy 2020; Funk et al. 2020; T. W. Smith and Son 2013): Since the 1970s, on average 40% of Americans say they have a great deal of confidence in the scientific community, the second highest score among 13 institutions included in the US General Social Survey (GSS). How can we explain this gap between knowing science and trusting it? Here, we argue that people might trust science in large part because they have been impressed by it.

We begin by presenting global evidence that most people trust science at least to some extent. We then review existing explanations for why people trust science, as well as empirical observations that seem to violate their predictions. As an alternative explanation, we propose the 'rational impression model'. Two experiments provide evidence in favor of this model.

Trust in science matters. It is related to many desirable outcomes, from acceptance of anthropogenic climate change (Cologna and Siegrist 2020) or vaccination (Sturgis, Brunton-Smith, and Jackson 2021; Lindholt et al. 2021) to following recommendations during COVID (Algan et al. 2021).

Across the globe, most people report trusting science at least to some extent (Wellcome Global Monitor 2018, 2020). Recently, a large survey in 68 countries found trust in scientists to be "moderately high" across all countries (mean = 3.62; sd= 0.70; Scale: 1 = very low, 2 = somewhat low, 3 = neither high nor low, 4 = somewhat high, 5 = very high), with not a single country below midpoint trust. A recent study in the US found that participants almost always accepted the scientific consensus on basic, non-politicized knowledge questions (e.g. 'Are electrons are smaller than atoms?'), even those participants who said they do not trust science and who believed anti-science conspiracy theories, e.g. that the earth is flat (Pfänder, Kerzreho, and Mercier 2024).

How can we explain trust in science? The literature on the public understanding of science has generally focused not on explaining trust, but distrust, and more generally negative attitudes towards science (Bauer, Allum, and Miller 2007).

One explanation for distrust towards science is the *alienation model* (Gauchat 2011). According to this model, the "public disassociation with science is a symptom of a general disenchantment with late modernity, mainly, the limitations associated with codified expertise, rational bureaucracy, and institutional authority" (Gauchat 2011, 2). This explanation builds on the work of social theorists (Habermas 1989; Beck 1992; Giddens 1991; see Gauchat 2011 for an overview) who suggested that a modern, complex world increasingly requires expertise, and thus shapes institutions of knowledge elites. People who are not part of these institutions experience a lack of agency, resulting in a feeling of alienation. If the alienation model might account for some of the distrust towards science, its goal is not to explain why most people still trust science.

A second explanation which has dominated the literature on public understanding of science for decades (Bauer, Allum, and Miller 2007), is the deficit model. According to the deficit model, people do not trust science enough, because they do not know enough about it. The deficit model implies that when people trust science, it is because they are knowledgeable about it (Bauer, Allum, and Miller 2007).

In line with that view, large scale survey studies consistently identify education, and in particular science education, to be the strongest correlate of trust in science (Bak 2001; Noy and O'Brien 2019; Wellcome Global Monitor 2018, 2020): More educated people tend to trust science more. The deficit model provides a straightforward causal explanation for this correlation:

education fosters trust in science by transmitting knowledge and improving understanding of science.

Evidence for the role of knowledge in trust in science, however, is far from conclusive. While there is no consensus among researchers on how to measure understanding of science (Gauchat and Andrews 2018), a widely used measure consists of asking people a set of basic science knowledge questions (see e.g. Durant, Evans, and Thomas 1989; Miller 1998). Past research has found science knowledge to be alarmingly low (Miller 2004; National Academies of Sciences, Engineering, and Medicine 2016), contrasting with the relatively high levels of trust in science. Moreover, science knowledge is only weakly associated with attitudes towards science (Allum et al. 2008). Other research has shown that the correlation between education and positive attitudes towards science holds even after controlling for science knowledge, which led the researchers to conclude that only part of the effect of education could be explained by transmitting enduring knowledge of science (Bak 2001).

3.2 The rational impression model

Here, we seek to explain the following stylized facts: (i) Globally, most people have at least some trust in science, and essentially everyone, at least in the US, trusts nearly all of basic science; (ii) Education, and in particular science education, is the strongest predictor of trust in science; (iii) Levels of science knowledge are relatively low, and the association between science knowledge and trust in science is weak.

All of these observations are compatible with a model of trust in science which we dub the *rational impression model*. According to this model, people trust science because they are impressed by it, but they forget about the findings that impressed them.

For an information to be impressive, at least two criteria should be met: (i) it is perceived as hard to uncover and (ii) there is reason to believe it true. By (i), we mean information that most people would have no idea how to find out themselves. For example, most people would probably only be mildly impressed by someone telling them that a given tree has exactly 110,201 leaves. Even though obtaining this information implies an exhausting counting effort, everyone in principle knows how to do it. By contrast, finding out that it takes light approximately 100,000 years to travel from one end of the Milky Way to the other is probably impressive to most people, as they would not know how such a distance can be measured. Regarding (ii), people can sometimes judge the accuracy of scientific findings for themselves, for example when they are exposed to accessible and convincing explanations in school, e.g. simple explanations that have a broad explanatory scope (Read and Marcus-Newhall 1993; Lombrozo 2007; for a review, see Lombrozo 2006). When people cannot evaluate the quality of the information for themselves, they need to rely on other cues to establish whether it might be true or not. One such cue is the degree of consensus between different sources. It has been shown that when they lack relevant background knowledge, people rely on social consensus as a heuristic to judge the accuracy of a piece of information and the competence of its source

(Pfänder, De Courson, and Mercier 2025). This heuristic is rational, in the sense that it leads to sound inferences if the sources are independent and unbiased. As a result, unless we suspect a conspiracy among scientists, we should be impressed when they agree on something like the length of the Milky Way, and think of the scientists as competent, even if we have no idea how they got to agree on this measure.

Once people are impressed, that impression tends to stick, even if specific content is forgotten. We commonly form impressions of the people around us while forgetting the details of how we formed these impressions: If a colleague fixes our computer, we might forget exactly how they fixed it, yet remember that they are good at fixing computers. As an extreme example, patients with severe amnesia can continue to experience emotions linked to events they could not recall (Feinstein, Duff, and Tranel 2010). More relevantly here, a study has shown that while people find some science-related explanations more satisfying than others, this did not predict how well they could recall the explanations shortly after (Liquin and Lombrozo 2022), suggesting that that impressions and knowledge formation can be quite detached.

The rational impression model of trust in science reconciles the stylized facts outlined above: If most people trust science (i), it is because most people receive at least some basic education that exposes them to impressive scientific findings, setting up a baseline of trust in science. If education is the main predictor of trust in science (ii), it is because education is the primary mean through which people are exposed to science. If the levels of science knowledge are low while trust in science is high (iii), it is because specific knowledge of science is forgotten, while the impression of competence and trust is maintained.

While the rational impression model, we have argued, can reconcile well-established facts about trust in science and science knowledge, there is no experimental evidence directly testing its central mechanism. This is the goal of the present studies.

3.3 The present studies

We tested the predictions of the rational impression model in two experiments (total n = 696): Experiment 1 tests whether impressive scientific findings lead people to think of the scientists as more competent and more trustworthy. Experiment 2 tests whether that impression persists even if specific content is forgotten. Figure 3.1 provides an overview of the main results of the two experiments.

Both experiments were preregistered, and the choice of sample size was informed by power simulations. All materials, data, and code can be found on Open Science Framework project page (https://osf.io/j3bk4/). All analyses were conducted in R (version 4.2.2) using R Studio. We used two-sided tests for all hypotheses. Unless mentioned otherwise, we report unstandardized estimates that can be interpreted in units of the original scales.

As part of this project, we conducted two additional experiments which we present in detail in the ESM. The first experiment ('Experiment 1b' in the ESM) is almost identical to Experiment

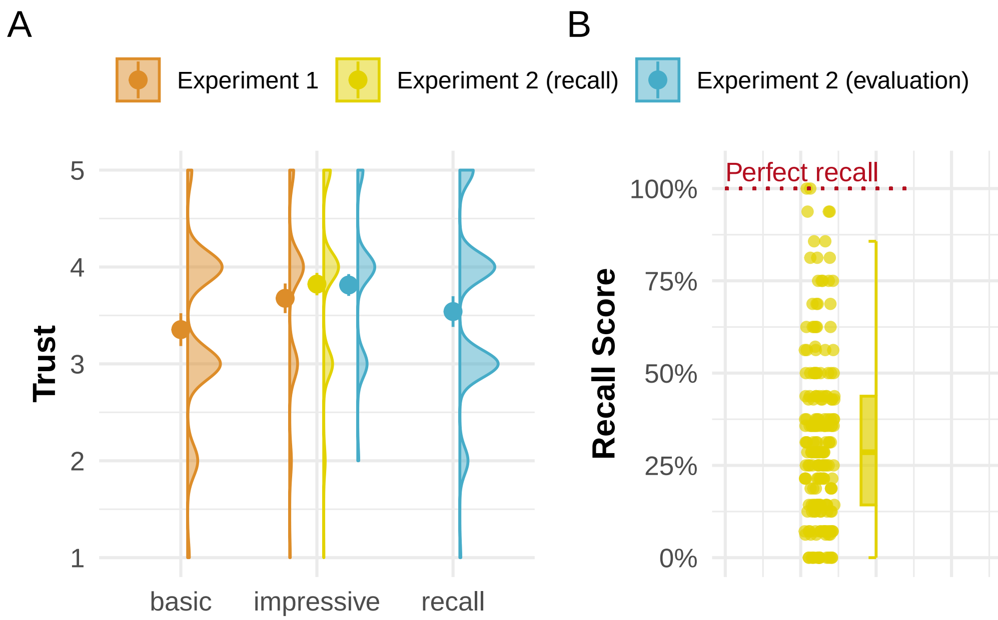

Figure 3.1: **A.** An overview of the differences in trust, according to whether the participants had read a basic, an impressive, or another participant's recalled version of the impressive vignette. The density plots show the distributions of participants' trust scores. The dots represent model estimates, and the vertical bars the 95% confidence intervals of the model predictions. **B.** The distribution of knowledge scores (ranging from 0% to 100% retained information) in the recall study of Experiment 2. Each dot corresponds to one participant. For the boxplot, the box represents the interquartile range (IQR), that is, the distance between the first and third quartiles, the center line indicates the median, and the outer lines (whiskers) extend to 1.5 times the IQR or the most extreme values within this range.

1, but suffered from a minor technical error during the implementation. Nonetheless, its findings are identical to those of Experiment 1. The second experiment ('Experiment 2b') was supposed to test the same hypothesis as Experiment 2, but the treatment–a very short distraction task–did not alter any of the outcome variables, suggesting that our outcome measures, a set of multiple-choice questions, was not sufficiently sensitive.

3.4 Experiment 1

The goal of Experiment 1 was to test whether exposure to impressive science content enhanced people's trust in scientists and their discipline. Participants were presented with vignettes about scientific findings in the disciplines of entomology and archaeology. The impressiveness of the texts was manipulated by creating one 'basic' and one 'impressive' version for each of the disciplines (see Table 3.1). Impressiveness was manipulated within participants, but between disciplines: each participant was randomly assigned to see an impressive version for one discipline, and a basic version for the other discipline. We tested the following hypotheses:

H1a: After having read an impressive text about a discipline's findings, compared to when reading a basic text, participants perceive that discipline's scientists as more competent.

H1b: Across both conditions, participants who are more impressed by the text about a discipline also tend to perceive the scientists of that discipline as more competent.

H2a: After reading an impressive text about a discipline's findings, compared to when reading a basic text, participants will trust the discipline more.

H2b: Across both conditions, participants who are more impressed by the text about a discipline also tend to trust the discipline more.

The results of two research questions, about perceptions of learning and of consensus, are reported in the ESM.

3.4.1 Methods

3.4.1.1 Participants

A power simulation (see OSF) suggested that the minimum required sample size to detect a statistically significant effect for all hypotheses with a power of 0.9 is 100 participants. We therefore recruited 100 participants from the UK via Prolific. One participant failed the attention check, resulting in a final sample of 99 participants (49 female, 50 male; age_{mean}: 42.071, age_{sd}: 12.857, age_{median}: 40).

3.4.1.2 Procedure

After providing their consent to participate in the study and passing an attention check (see ESM), participants read a short introductory text, and then two vignettes (one basic and one impressive) about scientific findings in the disciplines of entomology and archaeology, in a randomized order. After reading each vignette, participants were asked: “How much do you feel you’ve learnt about [human history/insects] by reading this text?” [1 - Nothing, 2 - A bit, 3 - Some, 4 - Quite a bit, 5 - A lot]); “How impressive do you think the findings of the [archaeologists/entomologists] described in the text are?” [1 - Not very impressive, 2 - A bit impressive, 3 - Quite impressive, 4 - Very impressive, 5 - Extremely impressive]); “Would you agree that reading this text has made you think of [archaeologists/entomologists] as more competent than you thought before?” [1 - Strongly disagree, 2 - Disagree, 3 - Neither agree nor disagree, 4 - Agree, 5 - Strongly agree]); and “Having read this text, would you agree that you trust the discipline of [archaeology/entomology] more than you did before?” [1 - Strongly disagree, 2 - Disagree, 3 - Neither agree nor disagree, 4 - Agree, 5 - Strongly agree]. Finally, we asked: “To which extent do you think the findings from the short text you just read reflect a minority or a majority opinion among archaeologists?” [1 - Small minority, 2 - Minority, 3 - About half, 4 - Majority, 5 - Large majority].

3.4.1.3 Materials

Table 3.1 shows the stimuli used in Experiment 1.

Table 3.1: Stimuli of Experiment 1

	Impressive	Basic
Archeology	Archaeologists, scientists who study human history and prehistory, are able to tell, from their bones, whether someone was male or female, how old they were, and whether they suffered from a range of diseases. Archaeologists can now tell at what age someone, dead for tens of thousands of years, stopped drinking their mother’s milk, from the composition of their teeth. Archaeologists learn about the language that our ancestors or cousins might have had. For instance, the nerve that is used to control breathing is larger in humans than in apes, plausibly because we need more fine-grained control of our breathing in order to speak. As a result, the canal containing that nerve is larger in humans than in apes – and it is also enlarged in Neanderthals. Archaeologists can also tell, from an analysis of the tools they made, that most Neanderthals were right-handed. It’s thought that handedness is related to the evolution of language, another piece of evidence suggesting that Neanderthals likely possessed a form of language.	Archaeology is the science that studies human history and prehistory based on the analysis of objects from the past such as human bones, engravings, constructions, and various objects, from nails to bits of pots. This task requires a great deal of carefulness, because objects from the past need to often be dug out from the ground and patiently cleaned, without destroying them in the process. Archaeologists have been able to shed light on human history in all continents, from ancient Egypt to the Incas in Peru or the Khmers in Cambodia. Archaeologists study the paintings made by our ancestors, such as those that can be found in Lascaux, a set of caves found in the south of France that have been decorated by people at least 30000 years ago. Archaeologists have also found remains of our more distant ancestors, showing that our species is just one among several that appeared, and then either changed or went extinct, such as Neanderthals, Homo erectus, or Homo habilis.

Table 3.1: Stimuli of Experiment 1 *(continued)*

	Impressive	Basic
Entomology	Entomologists are the scientists who study insects. Some of them have specialized in understanding how insects perceive the world around them, and they have uncovered remarkable abilities. Entomologists interested in how flies' visual perception works have used special displays to present images for much less than the blink of an eye, electrodes to record how individual cells in the flies' brain react, and ultra-precise electron microscopy to examine their eyes. Thanks to these techniques, they have shown that some flies can perceive images that are displayed for just three milliseconds (a thousandth of a second) – about ten times shorter than a single movie frame (of which there are 24 per second). Entomologists who study the hair of crickets have shown that these microscopic hairs, which can be found on antenna-like organs attached to the crickets' rear, are maybe the most sensitive organs in the animal kingdom. The researchers used extremely precise techniques to measure how the hair reacts to stimuli, such as laser-Doppler velocimetry, a technique capable of detecting the most minute of movements. They were able to show that the hair could react to changes in the motion of the air that had less energy than one particle of light, a single photon.	Entomologists are scientists who investigate insects, typically having a background in biology. They study, for example, how a swarm of bees organizes, or how ants communicate with each other. They also study how different insects interact with each other and their environment, whether some species are in danger of going extinct, or whether others are invasive species that need to be controlled. Sometimes entomologists study insects by observing them in the wild, sometimes they conduct controlled experiments in laboratories, to see for example how different environmental factors change the behavior of insects, or to track exactly the same insects over a longer period of time. An entomologist often specializes in one type of insect in order to study it in depth. For example, an entomologist who specializes in ants is called a myrmecologist.

3.4.2 Results and discussion

As a manipulation check, we find that participants perceived the impressive texts to be more impressive (mean = 4.01, sd = 0.87; $\hat{b}$ = 0.808 [0.569, 1.048], $p < .001$) than the basic texts (mean = 3.2, sd = 1.13).

Participants perceived scientists as more competent after having read an impressive text (H1a: $\hat{b}_{\text{Competence}}$ = 0.495 [0.314, 0.676], $p < .001$; mean = 3.76, sd = 0.9) than after having read a basic one (mean = 3.26, sd = 0.8). Pooled across both conditions, participants' impressiveness ratings were positively associated with competence: When participants reported being more impressed, they evaluated scientist#s as more competent (H1b: $\hat{b}$ = 0.407 [0.312, 0.502], $p < .001$).

Participants also trusted a discipline more after having read an impressive text (H2a: $\hat{b}_{\text{trust}}$ = 0.323 [0.17, 0.476], $p < .001$; mean = 3.68, sd = 0.87) than after having read a basic one. Participants' impressiveness ratings were positively associated with trust when pooling across all conditions (H2b: $\hat{b}$ = 0.296 [0.206, 0.385], $p < .001$).

3.5 Experiment 2

In Experiment 1, exposure to impressive scientific content increased trust in the relevant scientific discipline, and the perceived competence of the relevant scientists. Experiment 2 sought to test whether these perceptions were at least partly independent of being able to recall the specific content that had induced them. Experiment 2 consisted of a 'recall study' and an 'evaluation study.' In the recall study, each participant was assigned to read the impressive version of one of the vignettes from Experiment 1 (Table 3.1). As in Experiment 1, participants were asked about how impressed they were by the findings and whether they had changed their perception of the scientists' competence and their trust in the scientific discipline.

In line with the findings of Experiment 1, we expected participants to have increased perceptions of competence and trust in scientists after having read the impressive texts:

H1a: Participants perceive scientists as more competent after having read an impressive text about their discipline's findings.

H1b: Participants trust a discipline more after having read an impressive text about the discipline's findings.

In Experiment 2, participants were also given a recall task: They were asked to rewrite the text of the vignette they had just read, from memory. We predicted that participants would not be able to recall all the information presented in the short vignettes right after having read them (see methods section for details):

H2: Participants are not able to recall all the information of the original texts.

This hypothesis, however, only tested whether people forgot any of the content, potentially including non-impressive content. To address this issue, we asked participants to select elements of the vignettes they found impressive (see Table 3.2). We predicted that even for the subset of information that a participant said they found impressive, they forgot at least some of the content:

H3: Participants are not able to recall all the impressive information–as rated by themselves–contained in the original text.

H3 tests the hypothesis that participants immediately forget at least some of the information that has impressed them. However, it could be that participants in fact remember enough impressive information to justify the increase in perceived trust (in the discipline) and competence (in the scientists). To test whether that was the case, we conducted an evaluation study.

A new sample of participants was recruited, and they were randomly assigned to one of two conditions: In the 'original impressive text' condition, participants were assigned to read one of the two impressive texts of the recall study. In the 'recalled impressive text' condition, participants read one of the recall texts written by the participants of the recall study. We predicted that participants in the evaluation study would be less impressed by the texts recalled by the participants of the recall study, compared to the original impressive vignette texts (see Table 3.1) and, accordingly, would have less positive perceptions of the scientists' competence and the trustworthiness of their discipline:

H4a: The texts produced by participants of the experiment as a result of the recall task will be less impressive than the original texts, as rated by participants of the **evaluation study**.

H4b: Participants of the **evaluation study** perceive scientists as more competent after having read the original texts, compared to after having read the texts produced by participants of the experiment as a result of the recall task.

H4c: Participants of the **evaluation study** trust a discipline more after having read the original texts, compared to after having read the texts produced by participants of the experiment as a result of the recall task.

3.5.1 Methods

3.5.1.1 Participants

In a power simulation (see OSF), we varied the sample size of the recall study and effect sizes (assuming the same effect size for all hypotheses in each scenario). For all simulations, a constant evaluation study sample size of 400 participants was assumed (only relevant for H4a, b and c). The power simulation suggested that a power level of 90% would be reached when assuming a medium effect size of 0.5 with 50 participants. Due to uncertainty about our assumptions, we recruited a sample of 203 participants for the recall study, and a sample of 406 participants for the evaluation study. Although this was not the focus of the simulation, the results showed that for a medium effect size of 0.5, a sample size of 400 for the evaluation study yielded statistical power of greater than 90% for all hypotheses based on this sample (H4a, b and c). All participants were from the UK, recruited via Prolific, and paid to complete the experiment.

The final sample comprised 198 participants (four failed attention checks; 99 female, 99 male; age_{mean}: 42.646, age_{sd}: 15.547, age_{median}: 40) for the recall study, and 399 participants for the evaluation study (seven failed attention checks; 201 female, 198 male; age_{mean}: 41.709, age_{sd}: 13.483, age_{median}: 40).

3.5.1.2 Procedure

3.5.1.2.1 Recall study

In the recall study, after having consented to take part in the study and passing an attention check (see ESM), participants read the impressive version of one of the vignettes from Experiment 1 (Table 3.1). After reading the text, as in Experiment 1, participants were asked about changes in their perception of the scientists' competence, and trust in the scientists' discipline. They were also asked about the impressiveness of the text they read. The order of these questions was randomized.

Next, as an open-ended question, participants were asked to recall as much information as they could of the texts they had just read. They were told that their texts would be read by future participants. To further motivate participants, they were also told that they would get a bonus for recalling (without external aids) accurate information. They were not told how much that bonus would be. We paid them 5p per point gained in the recall task (see methods for how these points were assigned). This way, participants could reach a maximum bonus of 0.8 pound for archaeology (0.05p x 2 points x 8 content elements) and 0.7 pound (0.05p x 2 points x 7 content elements) for entomology. After that, participants were presented with the evaluation grid that we used to assess the open answers from the recall task (see Table 3.2). For each knowledge element, we asked participants to indicate whether they found it impressive or not ("Do you think this piece of information is impressive?" [Yes; No]). At the end of the recall study, participants were asked about their education level.

3.5.1.2.2 Evaluation study

After consenting to taking part in the study and passing an attention check (see ESM), participants read either one of the original vignettes, or a text produced as part of the recall task from the recall study. For those participants assigned to read a recall text, the text was randomly sampled (with replacement) from a pool of recall answers. Orthographic and grammatical mistakes in these texts were corrected with the help of ChatGPT beforehand. We had preregistered to sample among all recall answers from participants of the recall study. However, checking recall scores after the recall study and before launching the evaluation study, we found them to be very low on average (see ESM). To avoid having the answers of less motivated participants in our sample, we decided to only select answers that scored at least as well as the median in the recall measure of the evaluation study. This selection is conservative in that it makes it harder to confirm our predictions under H4.

After reading the text, just as in the recall study, participants were asked about changes in their perception of the scientists' competence, trust in the scientists' discipline, and the impressiveness of the text they read. The order of these questions was randomized. Finally, participants were asked about their education level.

3.5.1.3 Materials

For Experiment 2, besides the texts recalled by the participants of the recall study, the impressive version of the stimuli used in Experiment 1 was used (see Table 3.1).

Table 3.2: Recall evaluation grid

Archaeology	Entomology
1. Archaeologists can determine whether someone was male or female from their bones.	1. Entomologists use special displays to present images to flies for extremely short periods (less than the blink of an eye).
2. Archaeologists can determine how old someone was from their bones.	2. Entomologists can record how individual cells in flies' brains react using electrodes.
3. Archaeologists can determine whether someone suffered from a range of diseases from their bones.	3. Entomologists use ultra-precise electron microscopy to examine flies' eyes.
4. Archaeologists can determine at what age someone stopped drinking their mother's milk, based on the composition of their teeth.	4. Some flies can perceive images displayed for just three milliseconds. This duration is about ten times shorter than a single movie frame.
5. The nerve controlling breathing is larger in humans than in apes. The canal containing that nerve is also larger in humans and Neanderthals than in apes.	5. Crickets have microscopic hairs situated on antenna-like organs at their rear.
6. The fact that the nerve controlling breathing is larger in humans is possibly due to the need for fine-grained control of breathing to speak.	6. Crickets' hairs are possibly the most sensitive organs in the animal kingdom. They react to changes in air motion with less energy than one photon.
7. Archaeologists determined that most Neanderthals were right-handed, based on analysis of Neanderthals' tools.	7. Entomologists measured how cricket hairs react to stimuli, using laser-Doppler velocimetry, which can detect extremely minute movements.
8. Handedness is thought to be related to the evolution of language. This suggests that Neanderthals likely possessed a form of language.	

Note:

For each knowledge element, participants could score a maximum of two points.

3.5.1.3.1 Recall

As shown in Table 3.2, the texts were divided into a series of knowledge elements. For each participant, the extent to which they recalled each of the different elements was coded. A recall score based on how many of these elements they mentioned in their open-ended answer was then calculated.

The coding was done with the help of ChatGPT (see ESM for the exact prompt). The instructions were: to code 0 if a piece of knowledge was not mentioned or was mentioned with significant errors (e.g., writing "the nerve controlling fine hand movement is bigger in humans" instead of "the nerve controlling breathing is bigger in humans"); to code 1 if the piece of knowledge was mentioned, but some important elements were missing (e.g., writing "Archaeologists can determine at what age someone stopped drinking their mother's milk" instead of "Archaeologists can determine at what age someone stopped drinking their mother's milk from the composition of their teeth"), and/or there were some mistakes (e.g., writing "Archaeologists can determine at what age someone stopped drinking their mother's milk, based on the bones" instead of "Archaeologists can determine at what age someone stopped drinking their mother's milk based on the teeth"); to code 2 if the piece of knowledge was mentioned with all the main content, even if the participant had not used the precise technical words (e.g., "neanderthals," "laser-Doppler velocimetry") or had changed the phrasing in other ways. These instructions were intended to produce relatively generous recall scores.

Since the two vignettes contained a different number of total knowledge elements according to our evaluation grid (8 for archaeology, 7 for entomology), we used a relative measure for the final recall score, namely the share of obtained points among all possible points (possible range from 0 to 1, below the results are presented in percentages recalled for clarity). Practically, for all tests on recall, we computed a forgetting score (1-recall score), and tested whether it was statistically significantly different from zero.

To validate the scores assigned by ChatGPT, they were compared to scores assigned by two human coders for a subsample of 80 randomly chosen texts (half on archaeology, half on entomology). The human coders were unaware of the study context and the hypotheses. They were provided with the prompt given to ChatGPT. To measure the agreement between ChatGPT and the human coders, we calculated an intraclass correlation coefficient (ICC) for two-way designs with random raters (Ten Hove, Jorgensen, and Ark 2024). Following the guidelines in Heyman et al. (2014), we preregistered taking 0.7 as a threshold for acceptable reliability. In our sample, we observe an ICC of 0.838 [se = 0.032]. Human coders and ChatGPT, on average, assigned exactly the same score in 72.9% of all rating instances (compared to 73.5% of exact agreement between the two human coders). Following our preregistration, we consider the high agreement as suggested by the ICC a validation of the use of ChatGPT.

3.5.1.3.2 Recall of impressive items

After giving their post-recall evaluation of scientists' competence and trust in the discipline, participants were presented with the evaluation grid shown in Table 3.2 for the respective discipline they had been randomized to see. For each element in the evaluation grid, they were asked whether they found it impressive or not. Then, for each participant, the recall score was computed just as described above, but only on the subset of those elements they had subjectively rated as impressive.

3.5.2 Results and discussion

3.5.2.1 Recall study

First, as a kind of validation check, we note that most participants declared finding all knowledge elements to be impressive ($mean_{\text{Archeology}} = 7.04$, $median_{\text{Archeology}} = 8$, number of knowledge elements = 8; $mean_{\text{Entomology}} = 6.39$, $median_{\text{Entomology}} = 7$, number of knowledge elements = 7).

For the first set of hypotheses, we tested whether the average scores for perceived change in competence and trust were significantly different from their respective scale midpoints (which corresponds to no change in perception). We first tested the outcome variables' distributions for normality, using a Shapiro-Wilk test. In all cases, this test suggested that the data is considered non-normally distributed ($p < 0.05$). Following the preregistration, we therefore did not run a default one-sample t-test, but used a Wilcoxon signed-rank test instead. H1a and H1b were both supported: After having read an impressive text about the findings of a scientific discipline, participants saw the scientists as more competent (H1a: median = 4, W = 12388.5, $p < .001$), and their discipline as more trustworthy (H1b: median = 4, W = 9714.5, $p < .001$).

Despite having been impressed, a first descriptive analysis suggested that participants seemed to recall only very little information (*mean* = 32%, *sd* = 23%, *median* = 29%). Given these low average scores, we opted for a more conservative approach to testing H2 and H3: We defined the median as a cut-off and selected only the 50% of participants who had the highest recall scores, removing participants who may have put in less effort. Since we hypothesized that participants would forget content, this selection made it less likely that the hypotheses would be confirmed. Even for the 50% participants with the best recall, both hypotheses were supported: Participants did not perfectly recall all information (H2: median = 0.43, W = 5778, $p < .001$) and they did not recall all information contained in the elements they judged as impressive themselves (H3: median = 0.43, W = 5778, $p < .001$)[1].

3.5.2.2 Evaluation study

For H4a, b and c, trust, competence and impressiveness ratings were compared between the 'original impressive text' and the 'recalled impressive text' conditions of the evaluation study using independent sample t-tests. Participants who read on of the two original impressive texts reported being more impressed (H4a: $\hat{b}_{\text{Impressiveness}} = 0.387$, t = 5.411, $p < .001$), rated the scientists of the respective discipline as more competent (H4b: $\hat{b}_{\text{Competence}} = 0.274$, t =

[1] Only in Study 4 do we find evidence that changing one's mind towards the scientific consensus is associated with (more) trust in science (Studies 1: r = 0.064, p = 0.387; 2: r = 0.161, p = 0.051; 3: r = 0.044, p = 0.619; 4: r = 0.148, p = 0.037) and only in Study 2 evidence that it is associated with (less) conspiracy beliefs and (less) conspiracy thinking (Studies 1: r = -0.139, p = 0.061; 2: r = -0.225, p = 0.006; 3: r = -0.044, p = 0.631; 4: r = -0.053, p = 0.455).

3.074, p = 0.002), and had more trust in the respective discipline (H4c: $\hat{b}_{\text{Trust}} = 0.274$, t = 3.398, p < .001) than participants who read one of the texts of the recall task.

3.6 Discussion

It has long been a puzzle to the deficit model–which suggests that trust in science is primarily driven by science knowledge–that knowledge of science is at best weakly associated with science attitudes (Allum et al. 2008; National Academies of Sciences, Engineering, and Medicine 2016). The rational impression model makes sense of this: It predicts that people trust science primarily because they have been impressed by it, not because they remember much of the knowledge that impressed them.

Two experiments provide evidence for this model. Experiment 1 showed that impressive scientific findings lead people to think of scientists as more competent and trust science more. Experiment 2 showed that these impressions are formed even though participants forget the content that generated them almost immediately after reading it.

How can this model be rational, if it posits that trust is largely detached from knowledge? It is rational in that the heuristics it builds on lead to sound inferences in many contexts. If someone discovers something that is hard to know, such as the size of the Milky Way, and there appears to be a consensus, we should expect them to be competent, even without knowing the details of how they made this discovery. Even forgetting specific knowledge is not irrational: It has been argued that one of the main functions of episodic memory is to justify our beliefs in communication with others (Mahr and Csibra 2018). As a result, we should be particularly good at remembering things we might need to convince others of. In this regard, incentives of remembering science seem to be weak: Most exposure to science happens at school, and there is little reason for young learners' minds to anticipate having to convince others of the merits of specific scientific findings, which are typically of little practical relevance to them, and which appear to be consensually accepted.

Does this mean that scientific education should focus on impressive but potentially hard to grasp findings, instead of fostering a proper understanding of more basic discoveries? No, for at least two reasons. First, for some students at least, this knowledge will be remembered, and will prove important in their lives. Second, positive impressions of science can also be created by a proper understanding of its explanations, methods, etc. – even if, once again, that understanding is then often lost. However, our findings do suggest that science educators should not despair when they observe the low rates of science knowledge in adults, since the exposure to science they provided is arguably responsible for creating a bedrock of trust in science.

The present framework also stresses the vital role that can be played by science communication. The relationship between science communication and trust in science has already been explored in depth (e.g., Weingart and Guenther 2016; for a recent review see König et al. 2023), but we believe the present model might make a useful contribution. In particular, it suggests that, even though more understanding of the underlying methods is always preferable (König et al. 2023), even a relatively superficial exposure to impressive findings can bolster trust in science. An important caveat is that impressive findings that haven't yet gained the approval of the community might be particularly likely to backfire if people learn they have been disproven (on the importance of presenting a measure of consensus alongside scientific information, see König et al. 2023).

The present experiments have a number of limitations: First, they were conducted on convenience samples recruited in a single country, the UK. Second, they were conducted within a very short time frame. While we can show that participants almost immediately forget about impressive content, it is not clear from our study for how long the impressions persist (although in other contexts impressions formed on the basis of a much more superficial exposure have been shown to last for months, Gunaydin, Selcuk, and Zayas 2017). Future studies could extend our findings to other populations and to longer time frames.

3.6.0.1 Data availability

Data for all experiments and the simulations is available on the OSF project page (https://osf.io/j3bk4/). Note that on the project page, all materials related to what is referred to as "Experiment 1" in this paper are stored under "experiment_2", and all materials related to "Experiment 2" in this paper are stored under "experiment_4". This numbering is due to the original order in which experiments for this project were conducted. For a detailed report on the other experiments conducted as part of this project, see the ESM.

3.6.0.2 Code availability

The code used to create all results (including tables and figures) of this manuscript is also available on the OSF project page (https://osf.io/j3bk4/).

3.6.0.3 Competing interest

The authors declare having no competing interests.

4 How wise is the crowd: Can we infer people are accurate and competent merely because they agree with each other?

Are people who agree on something more likely to be right and competent? Evidence suggests that people tend to make this inference. However, standard wisdom of crowds approaches only provide limited normative grounds. Using simulations and analytical arguments, we argue that when individuals make independent and unbiased estimates, under a wide range of parameters, individuals whose answers converge with each other tend to have more accurate answers and to be more competent. In 6 experiments (UK participants, total N = 1197), we show that participants infer that informants who agree have more accurate answers and are more competent, even when they have no priors, and that these inferences are weakened when the informants were systematically biased. In conclusion, we speculate that inferences from convergence to accuracy and competence might help explain why people deem scientists competent, even if they have little understanding of science.

published as:

Pfänder, J., De Courson, B., & Mercier, H. (2025). *How wise is the crowd: Can we infer people are accurate and competent merely because they agree with each other?* **Cognition**, 255, 106005. https://doi.org/10.1016/j.cognition.2024.106005
For supplementary materials, please refer either to the published version, or the publicly available preprint via github.

4.1 Introduction

Imagine that you live in ancient Greece, and a fellow called Eratostenes claims the circumference of the earth is 252000 stades (approximately 40000 kilometers). You know nothing about this man, the circumference of the Earth, or how one could measure such a thing. As a result, you discard Eratostenes' opinion and (mis)take him for a pretentious loon. But what if other scholars had arrived at very similar measurements, independently of Eratosthenes?

Or even if they had carefully checked his measurement, with a critical eye? Wouldn't that give you enough ground to believe not only that the estimates might be correct, but also that Eratosthenes and his fellow scholars must be quite bright, to be able to achieve such a feat as measuring the Earth?

In this article, we explore how, under some circumstances, we should, and we do infer that a group of individuals whose answers converge are likely to be correct, and to be competent in the relevant area, even if we had no prior belief about either what the correct answer was, or about these individuals' competence.

We begin by reviewing existing studies showing that people infer that competent informants who converge in their opinions are likely to be accurate. The wisdom of crowds literature provides normative grounds for this inference. We then argue that both the experimental and theoretical literature have paid little attention to extending this inference to cases in which there is no information about the informants' competence, and to inferences about the competence of the informants. We first develop normative models, both analytically and with simulations, to show that inferences from convergence to accuracy and to competence are warranted under a wide range of parameters. Second, we present a series of experiments in which participants evaluate both the accuracy and competence of informants as a function of how much their answers converge on a given problem, in the absence of any priors about these individuals' competence, or what the correct answer is.

4.2 Do people infer that individuals whose answers converge tend to be right, and to be competent?

The literature on the wisdom of crowds has treated separately situations with continuous answers, such as the weight of an ox in Galton's famous observation (Galton 1907), and with categorical answers, as when voters have to choose between two options, in the standard Condorcet Jury Theorem (De Condorcet 2014). The continuous and the categorical case are typically modeled with different tools, and they have usually been studied in different empirical literatures (see below). Given that they both represent common ways for answers to converge more or less (e.g. when people give numerical estimates vs. vote on one of a limited number of options), we treat them both here, with different simulations and experiments.

In the continuous case, the most relevant evidence comes from the literature on 'advice-taking' (for review, see, Kämmer et al. 2023). In these experiments, participants are called 'judges' who need to make numerical estimates–sometimes on factual knowledge, e.g. 'What year was the Suez Canal opened first?' (Yaniv 2004), sometimes on knowledge controlled by the experimenters, e.g. 'How many animals were on the screen you saw briefly?' (Molleman et al. 2020). To help answer these questions, participants are given estimates from others, the 'advisors'.

Most of this literature is irrelevant to the point at hand since participants are presented with single estimates, either from a single advisor (e.g. Bednarik and Schultze 2015; Soll and

Larrick 2009; Yaniv 2004; Yaniv and Kleinberger 2000; Harvey and Fischer 1997), or as an average coming from a group of advisors (e.g. Jayles et al. 2017; Mannes 2009), but without any information about the distribution of initial estimates, so that we cannot tell whether participants put more weight on more convergent answers.

Some advice-taking studies provide participants with a set of individual estimates. One subset of these studies manipulates the degree of convergence between groups of advisors, through the variance of estimates (Molleman et al. 2020; Yaniv, Choshen-Hillel, and Milyavsky 2009), or their range (Budescu and Rantilla 2000; Budescu et al. 2003; Budescu and Yu 2007). These studies find that participants are more confident about, or rely more on, estimates from groups of advisors that converge more.

Other studies manipulated the degree of convergence within a group of advisors. These studies present participants with a set of estimates, some of which are close to each other, while others are outliers (Harries, Yaniv, and Harvey 2004; Yaniv 1997, study 3 & 4). These studies find that participants discount outliers when aggregating estimates.

Studies on advice taking thus suggest participants believe that more convergent opinions are more likely to be correct. None of these studies investigated whether participants also believe that those whose opinions converge are also more likely to possess greater underlying competence.

In categorical choice contexts, there is ample and long-standing (e.g. Crutchfield 1955) evidence from experimental psychology that participants are more likely to be influenced by majority opinions, and that this influence is stronger when the majority is larger, both in absolute and in relative terms (e.g., Morgan et al. 2012; for review, see Mercier and Morin 2019). This is true even if normative conformity (when people follow the majority because of social pressure rather than a belief that the majority is correct) is unlikely to play an important role (e.g. because the answers are private, see Asch 1956). Similar results have been obtained with young children (e.g. Fusaro and Harris 2008; Corriveau, Fusaro, and Harris 2009; Bernard et al. 2015; Bernard, Proust, and Clément 2015; E. E. Chen, Corriveau, and Harris 2013; Herrmann et al. 2013; Morgan, Laland, and Harris 2015).

If many studies have demonstrated that participants tend to infer that more convergent answers are more likely to be correct, few have examined whether participants thought that this convergence was indicative of the informants' competence. One potential exception is a study with preschoolers in which the children were more likely to believe the opinion of an informant who had previously been the member of a majority over that of an informant who had dissented from the majority (Corriveau, Fusaro, and Harris 2009). However, it is not clear whether the children thought the members of the majority were particularly competent, since their task–naming an object–was one in which children should already expect a high degree of competence from (adult) informants. This result might thus indicate simply that children infer that someone who disagrees with several others on how to call something is likely wrong, and thus likely less competent at least in that domain.

4.3 What inferences from convergence should we expect people to draw?

Should we expect that people be able to infer that more convergent answers likely indicate not only more accurate answers, but also that those who gave the answers were competent? In order to make the best of communicated information, humans have to be able to evaluate it, so as to discard inaccurate or harmful information, while accepting accurate and beneficial information (J. M. Smith and Harper 2003). It has been argued that a suite of cognitive mechanisms–mechanisms of epistemic vigilance–evolved to serve this function (Sperber et al. 2010; Mercier 2020). Since the opinion of more than one individual is often available to us, there should be mechanisms of epistemic vigilance dedicated to processing such situations. It would be these mechanisms that lead us to put more weight on an opinion that is shared by a larger majority (in relative or absolute terms), and, in some cases at least, to discount majority opinion when the opinions haven't been formed independently of each other (Mercier and Miton 2019). Evidence suggests that these mechanisms rely on heuristics which become more refined with age (Morgan, Laland, and Harris 2015), and which are far from perfect (in particular, they ignore many cases of informational dependencies, see, e.g. Yousif, Aboody, and Keil 2019; Mercier and Miton 2019). As mentioned above, in the experiments evaluating how people process convergent information, the participants had grounds to believe that the information came from competent informants. However, the same mechanisms could lead people to perform the same inference when they do not have such information (especially since, as is shown presently, such an inference is warranted).

Regarding the inference from convergence to competence, other cognitive mechanisms allow us to infer how competent people are, based on a variety of cues, from visual access (Pillow 1989), to the time it takes to answer a question (Richardson and Keil 2022). One of the most basic of these mechanisms infers from the fact that someone was right, that they possess some underlying competence (e.g. Koenig, Clément, and Harris 2004). As a result, if participants infer that convergent opinions are more likely to be accurate, they should also infer that the informants who provided the opinions are competent. Before testing whether participants infer accuracy and competence from convergence, we show that these inferences are normatively warranted, both in the categorical and in the continuous case.

4.3.1 Analytical argument

Regarding the analytical answer, the question can be broken down into two questions. First, can we infer that a population of informants whose answers converge more is, on average, more competent? Second, can we infer that, within this population, individuals who are closer to the consensus or to the average answer are more competent?

In the continuous case, let us imagine a population of informants. Individual opinions are drawn from a normal distribution, centered on the correct answer (i.e., informants have no systematic bias, as for instance in Galton's classic demonstration, Galton 1907). The variance of this distribution represents the individuals' competence: the larger the variance, the lower the competence. We observe the individual answers. In this setting, it is well known in statistics (see also Electronic Supplementary Materials, ESM) that the sample mean (i.e. the mean of the answers of all the informants) is the best estimator of the correct answer, and the sample variance (i.e. the mean squared distance between the answers and the sample mean) is the best estimator of the population's average competence (best understood here as the least volatile estimator). This means that a population of informants whose answers converge more (lower sample variance) is, on average, more competent (informants tend to answer with a lower variance).

We can extend this argument from populations to individuals: Consider that competence varies within a population, such that each informant's answer is drawn from their own distribution – always centered on the correct answer, but with different variances. In this case, the distance between each informant's answer and the sample mean provides the best estimate of that informant's competence (i.e. individuals whose answers are further away from the sample mean tend to be less competent).

In the categorical case, we define the competence of an informant as the probability of choosing the correct answer. The law of large numbers implies that the relative size of the majority – the share of informants who choose the most chosen answer – is, when the population is large enough, a good approximation of the average competence of the population (e.g. if the average competence is .66, and there are enough informants, then approximately 66% of informants will select the right answer). For smaller populations, the relationship holds, with some degree of noise. In other words, the more the answers converge, the more the population can be inferred to be competent (and the larger the population, the more reliable the inference).

Now, can we infer that informants belonging to the majority are more competent? To do so, we must assume some distribution of competence in the population. Using Bayes theorem, we can then show (see ESM) that the more informants agree with a focal informant, the more the focal informant can be inferred to be competent (Figure 4.1). We also find that, the more competence varies in the population, the more the degree of convergence is indicative of an informant's competence. For example, if competence is roughly uniform in the population, then being part of a minority likely reflects bad luck, rather than incompetence. More generally, the fact that the focal individual is right (their 'accuracy') can be inferred more strongly than their competence, as there is noise in the answer choice (an incompetent individual can always pick the correct answer by chance and vice versa). Figure 4.1) provides an example with a specific categorical choice scenario, under two different population distributions of competence.

4.3.2 Simulations

In order to better understand the influence of different parameters (e.g. degree of convergence, number of individuals) on the relationship between convergence, accuracy, and competence of

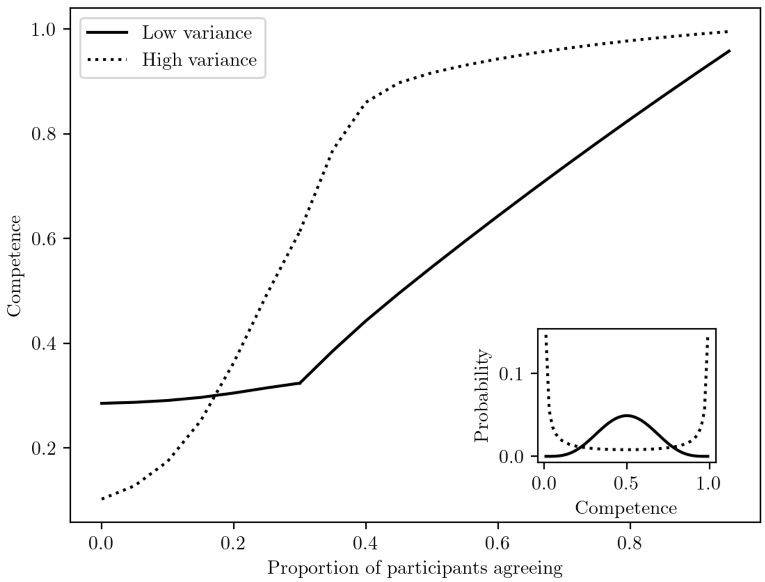

Figure 4.1: Results of the analytic argument. The figure shows the average estimated competence for a focal individual, depending on the proportion of individuals agreeing with them, in a categorical choice scenario. Here, we assume that 20 individuals independently pick one out of 5 choice options. Two situations are represented: one in which there is low variance in the distribution of competence in the population (the bell-shaped curve; solid line), and one in which there is high variance in this distribution (the u-shaped curve; dotted line).

informants, we conducted simulations. In all simulations, we assume agents to be unbiased and independent in their answers, but varying in their competence. All code and data regarding the simulations can be found on Open Science Framework project page (https://osf.io/6abqy/).

4.3.2.1 Continuous case

Groups of agents, with each agents' competence varying, provide numerical answers. We measure how accurate these answers are, and how much they converge (i.e. how low their variance is). We then look at the relationship between convergence and both the accuracy of the answers and the competence of the agents.

More specifically, agents provide an estimate on a scale from 1000 to 2000 (chosen to match the experiments below). Each agent is characterized by a normal distribution of possible answers. All of the agents' distributions are centered around the correct answer, but their standard deviation varies, representing varying degrees of competence. The agents' standard deviation varies from 1 (highest competence) to 1000 (lowest competence). Each agent's competence is drawn from a population competence distribution, expressed by a beta distribution, which can take different shapes. We conducted simulations with a variety of beta distributions which cover a wide range of possible competence populations (see Figure 4.2 A).

A population of around 990000 agents (varying slightly as a function of group sizes) with different competence levels is generated. An answer is drawn for each agent, based on their respective competence distribution. The accuracy of this answer is defined as the squared distance to the true answer. Having a competence and an accuracy value for each agent, we randomly assign agents to groups of, e.g., three. For each group, we calculate the average of the agents' competence and accuracy. We measure the convergence of a group's answers by calculating the standard deviation of the agents' answers. We repeat this process for different sample sizes for the groups, and different competence distributions. Figure 4.2 C displays the resulting relation of convergence with accuracy (left), and competence (right) for different underlying competence distributions and group sizes. We draw broad conclusions from these results after reporting the outcome of the simulations with categorical answers.

4.3.2.2 Categorical case

In the case of categorical answers, convergence can be measured as the relative size of informants picking the same option. The Condorcet jury theorem (De Condorcet 2014; for a recent treatment, see Dietrich and Spiekermann 2013) and its extensions to situations with more than two options (e.g., Hastie and Kameda 2005) already shows that the answer defended by the majority/plurality is more likely to be correct. Regarding competence, Romeijn and Atkinson (2011) have shown that when individuals are more competent, the share of the majority vote tends to increase. However, they have only studied a case with a uniform distribution of competence. By contrast, here, we investigate a wide range of distributions of competence, and we do not assume that all individuals are equally competent, meaning that, from an observed set of answers, we can attribute an average competence to individuals whose answers form a majority/plurality vs. individuals whose answers form a minority.

We simulate agents whose competence varies and who have to decide between a number of options, one of which is correct. Competence is defined as a value p which corresponds to the probability of selecting the right answer (the agents then have a probability (1-p)/(m-1), with m being the number of options, of selecting any other option). Competence values range from chance level (p = 1/m) to always selecting the correct option (p = 1). Individual competence levels are drawn from the same population competence beta distributions as in the numerical case (see Figure 4.2 A). Based on their competence level, we draw an answer for each agent. We measure an agent's accuracy as a binary outcome, namely whether they selected the correct option or not. In each simulation around 99900 agents (varying slightly as a function of the group size) are generated, and then randomly assigned to groups (of varying size in different simulations). Within these groups, we first calculate the share of individuals voting for each answer, allowing us to measure convergence. For example, in a scenario with three choice options and three individuals, two might vote for option A, and one for option C, resulting in two levels of convergence, 2/3 for A and 1/3 for C. For each level of convergence occurring within a group, we then compute (i) the accuracy (either 1 if the correct option or 0 else), (ii) the average competence of agents. Across all groups, we then compute the averages of these values, for each level of convergence.

We repeat this procedure varying population competence distributions and the size of informant groups, holding the number of choice options constant at n = 3 (for simulations with varying choice options, see ESM). Figure 4.2 B shows the average accuracy (left), and the average competence (right) value as a function of convergence, across different underlying competence distributions and group sizes.

The simulations for the numerical and categorical case demonstrate a similar pattern, which can be summarized as follows:

1. Irrespective of group size (and number of choice options) and of the competence distribution, there is a very strong relation between convergence and accuracy: more convergent answers tend to be more accurate.
2. For any group size and any competence distribution, there is a relation between convergence and the competence of the agents: more convergent answers tend to stem from more competent agents. The strength of this relation is not much affected by the number of agents whose answers are converging, but, although it is always positive, it ranges from very weak to very strong depending on the population's competence distribution.
3. The relation between convergence and accuracy is always much stronger than the relation between convergence and competence of the agents.

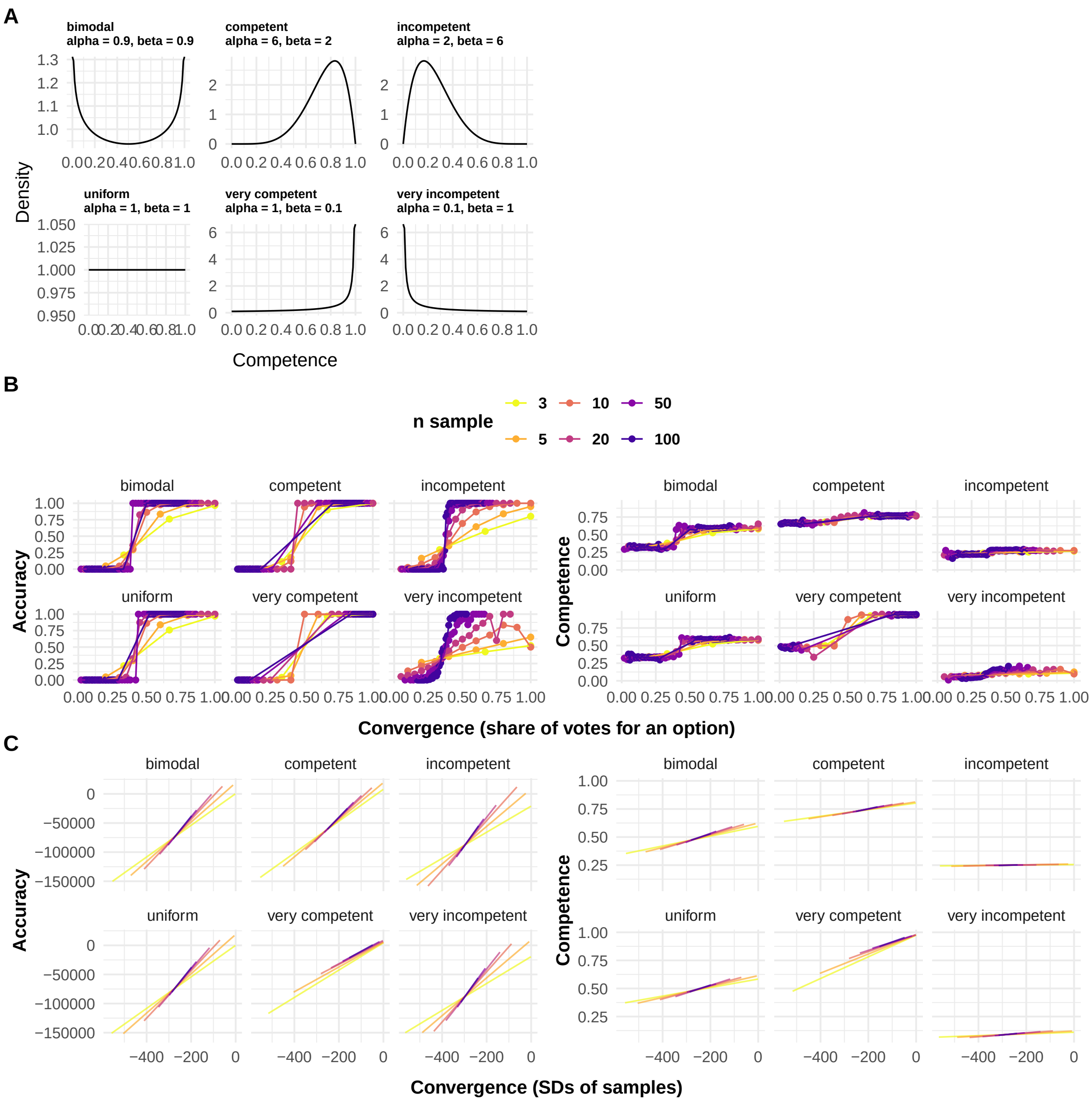

Figure 4.2: Results of the simulations. **A** Shows the different population competence distributions we considered in our simulations. In the continuous simulations, competence values of 0 correspond to a very large standard deviation (1000, with a mean of 1500, on a scale from 1000 to 2000), thereby practically taking the form of a uniform distribution, while competence of 1 corresponds to a very small standard deviation (1, on the same scale). In the categorical simulations, a competence value of 0 corresponds to chance (e.g. in a 3-choice-options scenario, an individual picking the correct answer with a probability of 1/3), while a competence value of 1 corresponds to definitely picking the correct answer. **B** Shows the results of simulations in a categorical setting with three choice options. Points represent average accuracy (left)/competence (right) values by degree of convergence (measured by the share of votes for an option), for different population competence distributions (panels) and sample sizes (colors). **C** Shows the results in a continuous setting. Regression lines represent the correlation between accuracy (left; measured by squared distance to true mean and reversed such that greater accuracy corresponds to being further up on the y-axis) or competence (right), respectively, and convergence (reversed such that greater convergence corresponds to being more right on the x-axis).

4.4 Overview of the experiments

Our models indicate that groups of informants are more likely to have given accurate answers, and to be competent, when their answers converge. In a series of experiments, we test whether people draw these inferences both in numerical tasks (Experiments 1, 2, 3), and in categorical tasks (Experiments 4, 5, 6). By contrast with previous studies, participants were not given any information about the tasks–how difficult they were–and the informants–how competent they might be. There has recently been much interest in investigating whether participants are able to take informational dependencies into account when evaluating convergent information (Yousif, Aboody, and Keil 2019; Mercier and Miton 2019; Desai, Xie, and Hayes 2022; Xie and Hayes 2022; Yin et al. 2024). To test whether participants took into account dependencies between the informants in the present context, this factor was manipulated in two different ways (in Experiment 2, some informants had discussed with each other; in Experiments 3 and 5, some informants had an incentive to provide the same answer). We predicted that dependencies between the informants' answers should reduce participants' reliance on the convergence of the answer as a cue to accuracy and to competence.

All experiments were preregistered. All documents, data and code can be found on Open Science Framework project page (https://osf.io/6abqy/). All analyses were conducted in R (version 4.2.2) using R Studio. For most statistical models, we relied on the `lme4` package and its `lmer()` function. Unless mentioned otherwise, we report unstandardized model coefficients that can be interpreted in units of the scales we use for our dependent variables.

4.5 Experiment 1

In Experiment 1, participants were provided with a set of numerical estimates which were more or less convergent, and asked whether they thought the estimates were accurate, and whether the informants making the estimates were competent. Perceptions of accuracy were measured as the confidence in what the participants thought the correct answer was, on the basis of the numerical estimates provided: the more participants think they can confidently infer the correct answer, the more they must think the estimates accurate, on average (the results replicate with a more direct measure of accuracy, see Experiment 3). Our hypotheses were:

H1: When making a guess based on the estimates of (independent) informants, participants will be more confident about their guess when these estimates converge compared to when they diverge.

H2: Participants perceive (independent) informants whose estimates converge more as more competent than informants whose estimates diverge.

We had three research questions regarding the number of informants which report in the ESM.

4.5.1 Methods

4.5.1.1 Participants

We recruited 200 participants from the UK via Prolific (100 female, 100 male; age_{mean}: 39.73, age_{sd}: 15.39, age_{median}: 35.5). Not a single participant failed our attention check. The sample size was determined on the basis of a power analysis for a t-test to detect the difference between two dependent means ("matched pairs") run on G*Power3. The analysis suggested that a combined sample of 199 would provide us with 80% power to detect a true effect size of Cohen's d ≥ 0.2 (alpha = .05, two-tailed).

4.5.1.2 Procedure

After providing their consent to participate in the study and passing an attention check, participants read the following introduction: "Some people are playing games in which they have to estimate various quantities. Each game is different. You have no idea how competent the people are: they might be completely at chance, or be very good at the task. It's also possible that some are really good while others are really bad. Some tasks might be hard while others are easy. Across various games, we will give you the answers of several players, and ask you questions about how good they are. As it so happens, for all the games, the estimates have to be between 1000 and 2000, but all the games are completely different otherwise, and might require different skills, be of different difficulties, etc. Each player in the game makes their own estimate, completely independent of the others". After being presented with the results of a game (Figure 4.3), participants had to (i) make a guess about the correct answer based on the estimates they see, "What would you guess is the correct answer, if there is one?", (ii) estimate their confidence in this guess, "How confident are you that your answer is at least approximately correct?" on a 7-point Likert scale ("not confident at all" to "extremely confident"), (iii) estimate the competence of the group of players whose estimates they saw, "On average, how good do you think these players are at the game?", also on a 7-point Likert scale (from "not good at all" to "extremely good").

4.5.1.3 Design

We manipulated two experimental factors, with two levels each: the convergence of the estimates (how close they were to each other; levels: divergent/convergent), and the number of estimates (how many players there were; levels: three/ten). This latter factor was chiefly included to make our results more robust, and was not attached to specific hypotheses. We used a 2 (convergence: divergent/convergent) x 2 (number: three/ten) within-participant design, with each participant seeing all the conditions. Participants saw two different sets of estimates per condition, for a total of eight sets of estimates per participant.

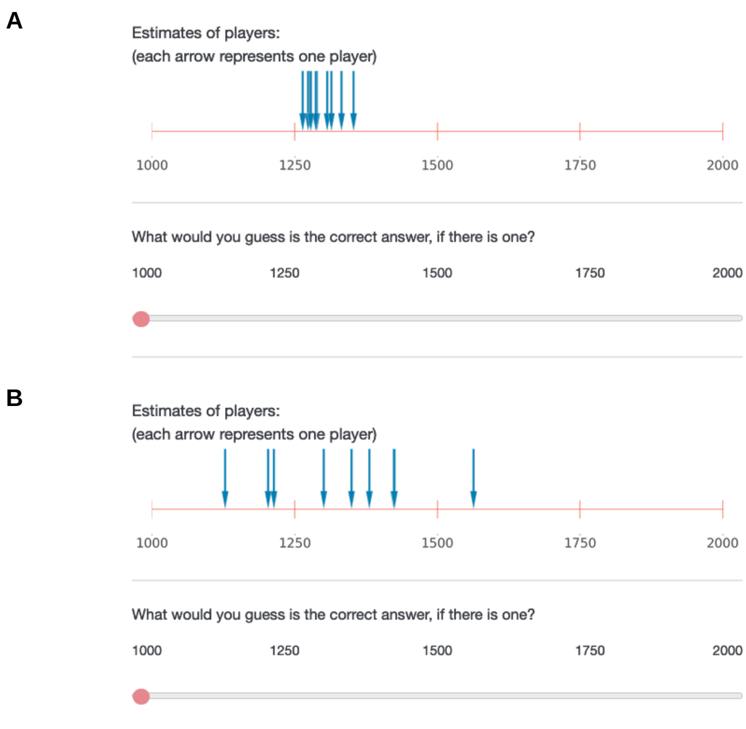

Figure 4.3: Example of two stimuli from Experiment 1, both in the 10 players condition, **A** corresponding to the convergent, **B** to the divergent condition. Similar stimuli are used in Experiments 2 and 3.

4.5.1.4 Materials

We generated sets of estimates with random draws from normal distributions. First, we varied the standard deviation of these distributions to simulate the degree of convergence (150 for divergence, 20 for convergence; estimate scale ranged from 1000 to 2000). Second, we varied the number of draws (either three or ten) from these distributions. For each of the four possible resulting conditions, we generated two random draws. We repeated this process for three different sets of estimates, and participants were randomly assigned to one of these sets. More information on how the stimuli were created can be found in the ESM.

4.5.2 Results and discussion

To account for dependencies of observations due to our within-participant design, we ran mixed models, with a random intercept and a random slopes of convergence for participants. In the models for our hypotheses, we control for the number of estimates provided to the participants (three or ten). Visualizations and descriptive statistics can be found in ESM. We find a positive effect of convergence on accuracy: Participants were more confident about their estimate in convergent scenarios (mean = 4.56, sd = 1.448) than in divergent ones (mean = 3.192, sd = 1.392; $\hat{b}_{\text{Accuracy}}$ = 1.368 [1.225, 1.51], $p < .001$). We also find a positive effect of convergence on competence: participants rated players as more competent in convergent scenarios (mean = 4.75, sd = 1.237) than in divergent ones (mean = 3.518, sd = 1.266; $\hat{b}_{\text{Competence}}$ = 1.232 [1.065, 1.4], $p < .001$).

In an exploratory, non-preregistered analysis, we tested whether the effect of convergence is larger on accuracy than on competence. To this end, we regressed the outcome score on convergence and its interaction with a binary variable indicating which outcome was asked for (accuracy or competence), while controlling for the number of informants. We do not find a statistically significant interaction that would indicate a difference of the effect of convergence ($\hat{b}$ = 0.135 [-0.001, 0.271], = 0.052). Pooled across divergent and convergent conditions, we find that participants reported lower perceived accuracy than competence ($\hat{b}$ = -0.257 [-0.359, -0.156], < .001)

In summary, as predicted, when the informants' answers were more convergent, participants were more confident that their answers were correct, and they believed the informants to be more competent. This was true both when there were three informants and when there were ten informants.

4.6 Experiment 2

We have shown that it is rational to infer that convergent estimates are more likely to be accurate, and to have been made by competent individuals, only if these individuals were independent and unbiased. However, convergence could come about differently. If the individuals do not make their estimates independently of each other, a single individual might exert a strong influence on the others, making their convergence a poor cue to their accuracy. Alternatively, all individuals might have an incentive to provide a similar, but not accurate answer. In Experiment 2, we investigate the first possibility, and the second in Experiment 3. In particular, for Experiment 2 we rely on past results showing that participants, under some circumstances, put less weight on opinions that have been formed through discussion, by contrast with more independent opinions (Harkins and Petty 1987; see also Lopes, Vala, and Garcia-Marques 2007; Hess and Hagen 2006; Einav 2018). We sought to replicate this finding in the context of convergent estimates, formulating the following hypotheses:

H1: When making a guess based on convergent estimates of informants, participants will be more confident about their guess when informants were independent compared to when they weren't (i.e. they could discuss before).

H2: Participants perceive informants whose estimates converge as more competent when they are independent, compared to when they weren't (i.e. they could discuss before).

Note that these predictions only stem from past empirical results, and are not necessarily normatively justified (modeling the effects of discussion would be beyond the scope of this paper, but see Dietrich and Spiekermann 2024).

4.6.1 Methods

4.6.1.1 Participants

We recruited 200 participants from the UK via Prolific (100 female, 99 male, 1 not-identified; age_{mean}: 40.545, age_{sd}: 13.561, age_{median}: 38.5). Not a single participant failed our attention check. As for experiment 1, the sample size was determined on the basis of a power analysis for a t-test to detect the difference between two dependent means ("matched pairs") run on G*Power3. The analysis suggested that a combined sample of 199 would provide us with 80% power to detect a true effect size of Cohen's d ≥ 0.2 (alpha = .05, two-tailed).

4.6.1.2 Design

In a within-participants design, participants saw both an independence condition, in which they were told "Players are asked to make completely independent decisions – they cannot see each other's estimates, or talk with each other before giving their estimates," and a dependence condition, in which they were told "Players are asked to talk with each other about the game at length before giving their estimates."

4.6.1.3 Materials

We used the materials generated for the convergent condition of Experiment 1. By contrast to Experiment 1, participants saw only two stimuli in total (one set of estimates per condition), and we only used stimuli involving groups of three informants. Otherwise, we proceeded just as in Experiment 1: we randomly assigned individual participants to one of the three series of stimuli, and for each participant, we randomized the order of appearance of conditions.

4.6.2 Results and discussion

To account for dependencies of observations due to our within-participant design, we ran mixed models, with a random intercept for participants. Visualizations and descriptive statistics can be found in ESM. The data does not support our hypotheses. Participants were slightly less confident about their estimates when the converging informants were independent (mean = 3.775, sd = 1.502), compared to when they discussed (mean = 4.03, sd = 1.389; $\hat{b}_{\text{Accuracy}}$ = -0.255 [-0.462, -0.048], p = 0.016). The effect is small, but in the opposite direction of what we had predicted. We do not find an effect regarding competence ($\hat{b}_{\text{Competence}}$ = -0.12 [-0.272, 0.032], p = 0.120).

Contrary to the hypotheses, participants did not deem convergent estimates made after a discussion, compared to independently made estimates, to be less accurate, or produced by less competent individuals. This might stem from the fact that participants, in various situations, neglect informational dependencies (Yousif, Aboody, and Keil 2019), or from the fact that discussing groups actually perform better than non-discussing groups in a range of tasks (for review, see, e.g., Mercier 2016), including numerical estimates (e.g. Mercier and Claidière 2022). As a result, the participants in the current experiment might have been behaving rationally when they did not discount the estimates made after discussion.

4.7 Experiment 3

Experiment 3 tests whether participants are sensitive to another potential source of dependency between convergent estimates: when the individuals making the estimate share an incentive to bias their estimates and disregard accuracy. Even though Experiment 3 is formally similar to Experiment 1, the setting is different, as participants were told that they would be looking at (fictional) predictions of experts for stock values, instead of the answers of individuals in abstract games. In the conflict of interest condition, the experts had an incentive to value the stock in a given way, while they had no such conflict of interest in the independence condition. We tested for an interaction, namely whether the positive effect of convergence is reduced when informants are systematically biased, compared to when they are not. On this basis, we formulate four hypotheses, two of which are identical to those of Experiment 1, and only apply in the independent condition, and two that bear on the interaction:

H1a: Participants perceive predictions of independent informants as more accurate when they converge compared to when they diverge.

H1b: Participants perceive independent informants as more competent when their predictions converge compared to when they diverge.

H2a: The effect of convergence on accuracy (H1a) is more positive in a context where informants are independent compared to when they are in a conflict of interest.

H2b: The effect of convergence on competence (H1b) is more positive in a context where informants are independent compared to when they are in a conflict of interest.

We have not conducted simulations to validate these predictions. However, given the operationalization chosen, in which convergence should provide little evidence of either accuracy or competence, we believe the predictions regarding the superiority of the independent informants stem naturally from the model of independent informants presented above.

4.7.1 Methods

4.7.1.1 Participants

The interaction design of our third experiment made the power analysis more complex and less standard than for experiments one and two. Because we could build upon data from the first experiment, we ran a power analysis by simulation. The simulation code is available on the OSF, and the procedure is described in the preregistration document. The simulation suggested that 100 participants provide a significant interaction term between 95% and 97% of the time, given an alpha threshold for significance of 0.05. Due to uncertainty about our effect size assumptions and because we had resources for a larger sample, we recruited 199 participants for this study – again, from the UK and via Prolific (99 female, 100 male; age_{mean}: 40.296, age_{sd}: 12.725, age_{median}: 38).

4.7.1.2 Procedure

After providing their consent to participate in the study and passing an attention check, participants read the following introduction: "You will see four scenarios in which several experts predict the future value of a stock. You have no idea how competent the experts are. It's also possible that some are really good while others are really bad. As it so happens, in all scenarios, the predictions for the value of the stock have to lie between 1000 and 2000. Other than that, the scenarios are completely unrelated: it is different experts predicting the values of different stocks every time." Participants then saw the four scenarios, each introduced by a text according to the condition the participant was assigned to. To remove any potential ambiguity about participants' inferences on the accuracy of the estimates, we replaced the question about confidence to one bearing directly on accuracy: "On average, how accurate do you think these three predictions are?" on a 7-point Likert scale ("not accurate at all" to "extremely accurate"). The question about competence read: "On average, how good do you think these three experts are at predicting the value of stocks?", also assessed on a 7-point Likert scale (from "not good at all" to "extremely good").

4.7.1.3 Design

We manipulated two factors: informational dependency (two levels, independence and conflict of interest; between participants) and convergence (two levels, convergence and divergence; within participants). In the independence condition, the participants read "Experts are independent of each other, and have no conflict of interest in predicting the stock value - they do not personally profit in any way from any future valuation of the stock." In the conflict of interest condition, the participants read "All three experts have invested in the specific stock whose value they are predicting, and they benefit if other people believe that the stock will be valued at [mean of respective distribution] in the future."

4.7.1.4 Materials

The distributions presented were similar to those of Experiment 1, although generated in a slightly different manner (see ESM). Each participant rated four scenarios, two for each level of convergence. By contrast to Experiment 1, all scenarios only involved groups of three informants.

4.7.2 Results and discussion

To account for dependencies of observations due to our within-participant design, we ran mixed models, with a random intercept and a random slope for convergence for participants. We find evidence for all four hypotheses. As for the first set of hypotheses, to match the setting of experiment one, we reduced the sample of Experiment 3 to half of the participants, namely those who were assigned to the independence condition. On this reduced sample, we ran the exact same analyses as in Experiment 1 and replicated the results. As for accuracy, participants rated informants in convergent scenarios (mean = 5.28, sd = 1.052) as more accurate than in divergent ones (mean = 3.4, sd = 1.08; $\hat{b}_{\text{Accuracy}} = 1.88$ [1.658, 2.102], $p < .001$). As for competence, participants rated informants in convergent scenarios (mean = 5.235, sd = 0.992) as more competent than in divergent ones (mean = 3.61, sd = 1.111; $\hat{b}_{\text{Competence}} = 1.625$ [1.411, 1.839], $p < .001$).

The second set of hypotheses targeted the interaction of informational dependency and convergence (Figure 4.4). In the independence condition, the effect of convergence on accuracy was more positive ($\hat{b}_{\text{interaction, Accuracy}}$ = 0.991 [0.634, 1.348], p < .001) than in the conflict of interest condition ($\hat{b}_{\text{baseline}}$ = 0.889 [0.636, 1.142], p < .001). Likewise the effect of convergence on competence is more positive ($\hat{b}_{\text{interaction, Competence}}$ = 0.802 [0.474, 1.13], p < .001) than in the conflict of interest condition ($\hat{b}_{\text{baseline}}$ = 0.823 [0.591, 1.056], p < .001).

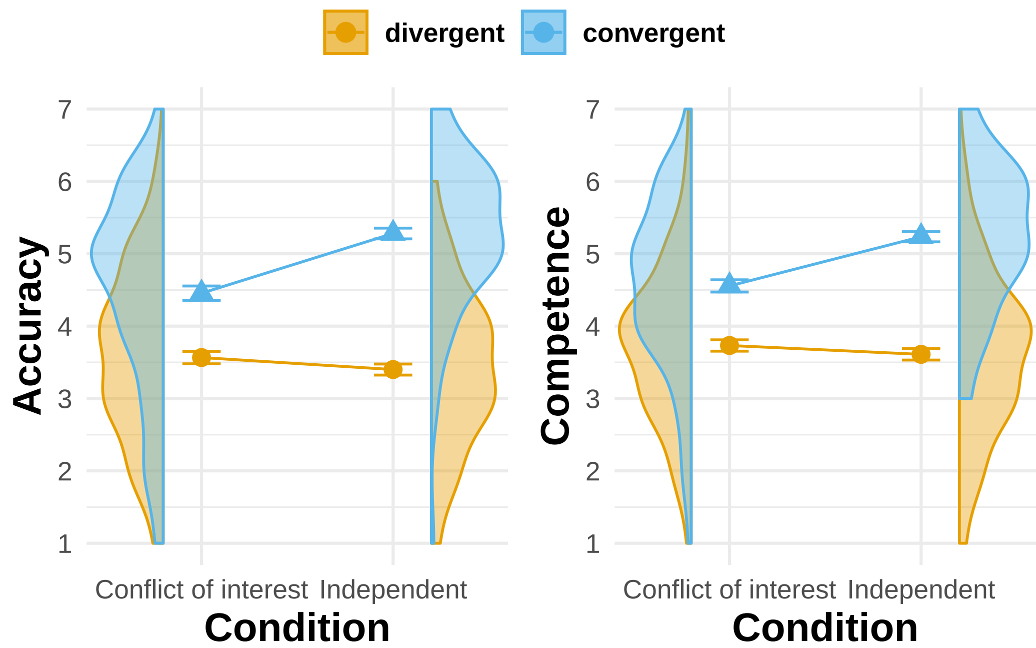

Figure 4.4: Results of Experiment 3, showing the distributions of accuracy and competence ratings by convergence and informational dependency.

In an exploratory, non-preregistered analysis, we tested whether the effect of convergence is larger on accuracy than on competence. To this end, we regressed the outcome score on convergence and its interaction with a binary variable indicating which outcome was asked for (accuracy or competence), while controlling for informational dependency. We find a negative interaction effect, indicating that pooled across independent and conflict of interest conditions, the effect of convergence had as smaller effect on competence than on accuracy ($\hat{b}$ = -0.161 [-0.294, -0.028], = 0.018). Pooled across all conditions, participants reported higher perceived competence than accuracy ($\hat{b}$ = 0.108 [0.025, 0.191], = 0.011).

Experiment 3 shows that, when the individuals making the estimates are systematically biased, participants put less weight on the convergence of their estimates to infer that the estimates are accurate, and that the individuals making them are competent.

4.8 Experiment 4

In a second series of experiments, we test similar predictions to those of the previous experiments, but in a categorical choice context. The set-up is similar to that of Experiment 1, except that the outcomes seen by the participants are not numerical estimates, but choices made between a few options. An additional difference is that participants rate a focal informant, and not a group of informants. There were two reasons for this choice: First, accuracy is not on a continuum as in the first three experiments (an option was either correct or not), so forming an average across informants who chose different options was less sensible. Second, rating a focal individual allowed us to have a minority condition, which would not have been possible when providing an average rating for a group. Experiment 4 tests hypotheses that are analogous to those of Experiment 1:

H1: Participants perceive an estimate of an independent informant as more accurate the more it converges with the estimates of other informants.

H2: Participants perceive an independent informant as more competent the more their estimate converges with the estimates of other informants.

4.8.1 Methods

4.8.1.1 Participants

We ran a power simulation to inform our choice of sample size. All assumptions and details on the procedure can be found on the OSF. We ran two different power analyses, one for each outcome variable. We set the power threshold for our experiment to 90%. The power simulation for accuracy suggested that even for as few as 10 participants (the minimum sample size we simulated data for), we would have a power of close to 100%. The simulation for competence suggested that we achieve statistical power of at least 90% with a sample size of 30. Due to uncertainty about our assumptions and because it was within our budget, we recruited 100 participants, from the UK and via Prolific (50 female, 50; age_{mean}: 37.32, age_{sd}: 11.526, age_{median}: 36).

4.8.1.2 Procedure

After providing their consent to participate in the study and passing an attention check, participants read the following introduction: “To be able to understand the task, please read the following instructions carefully: Some people are playing games in which they have to select the correct answer among three answers. You will see the results of several of these games. Each game is different, with different solutions and involving different players. All players answer independently of each other. At first, you have no idea how competent each individual player is: they might be completely at chance, or be very good at the task. It’s also possible that some players are really good while others are really bad. Some games might be difficult while others are easy. Your task will be to evaluate the performance of one of the players based on what everyone's answers are." They were then presented to the results of eight such games (Figure 4.5). To assess perceived accuracy, we asked: "What do you think is the probability of player 1 being correct?". Participants answered with a slider on a scale from 0 to 100. To assess perceived competence, we asked participants: "How competent do you think player 1 is in games like these?" Participants answered on a 7-point Likert scale (from (1)"not competent at all" to (2)"extremely competent").

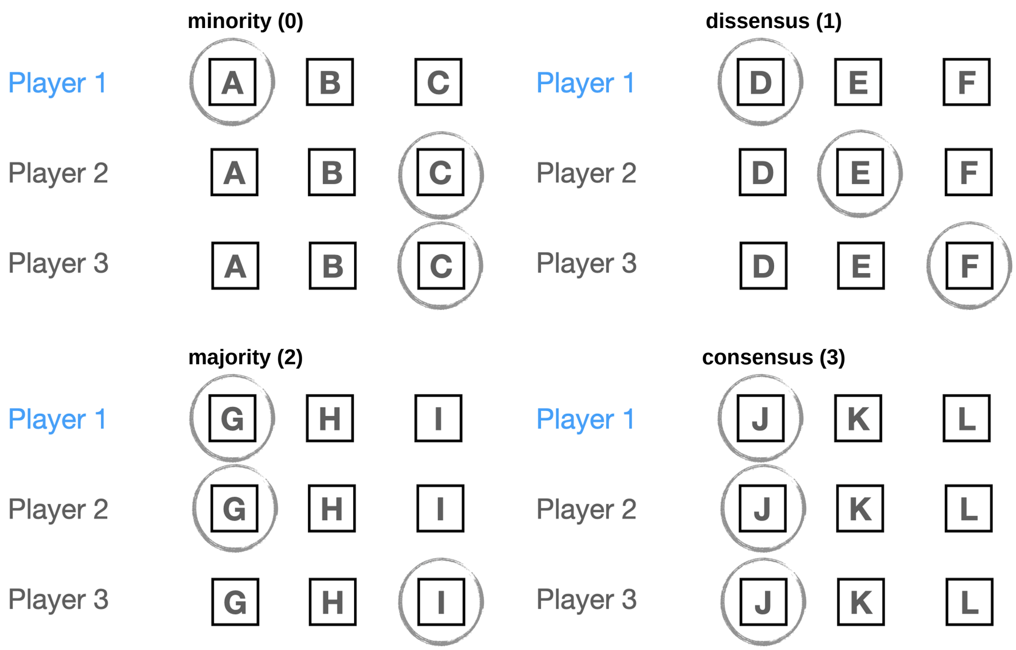

Figure 4.5: One set of stimuli by level of convergence, in Experiment 4 (similar stimuli are used in Experiments 5 and 6). A full set of stimuli can be found in the ESM.

4.8.1.3 Design

We manipulated convergence within participants, by varying the ratio of players choosing the same response as a focal player (i.e. the one that participants evaluate). The levels of convergence are: (i) consensus, where all three players pick the same option [coded value = 3]; (ii) majority, where either the third or second player picks the same option as the first player [coded value = 2]; (iii) dissensus, where all three players pick different options [coded value = 1]; (iv) minority, where the second and third player pick the same option, but one that is different from the first player's choice [coded value = 0]. In our analysis, we treat convergence as a continuous variable, assigning the coded values in squared parenthesis here.

4.8.1.4 Materials

All participants saw all four conditions, with two stimuli per condition. Each participant therefore saw eight stimuli in total (4 convergence levels x 2 stimuli).

4.8.2 Results and discussion

To account for dependencies of observations due to our within-participant design, we ran mixed models, with a random intercept and a random slope for participants.

As in the numerical setting, we found a positive effect of convergence on both accuracy ($\hat{b}_{\text{Accuracy}}$ = 16.838 [15.009, 18.668], $p < .001$; on a scale from 0 to 100) and competence ($\hat{b}_{\text{Competence}}$ = 0.683 [0.578, 0.788], $p < .001$; on a scale from 1 to 7).

In the ESM, we show that compared to what the normative models would predict, participants underestimate the effect of convergence on both accuracy and competence, but especially on accuracy.

In an exploratory, non-preregistered analysis, we tested whether the effect of convergence is larger on accuracy than on competence. To do so, we first standardized both outcome scores to account for the different scales. We then regressed the outcome score on convergence and its interaction with a binary variable indicating which outcome was asked for (accuracy or competence). We find a negative interaction, indicating that convergence had a smaller effect on competence than on accuracy ($\hat{b}$ = -0.095 [-0.136, -0.053], < .001; units in standard deviations).

4.9 Experiment 5

Experiment 5 is a conceptual replication of Experiment 3 in a categorical instead of a numerical case: are participants less likely to infer that more convergent estimates are more accurate, and the individuals who made them more competent, when the estimates are made by individuals with a conflict of interest pushing them to all provide a given answer, compared to when they are made by independent individuals? The independence condition of Experiment 5 also serves as a replication of Experiment 4, leading to the following hypotheses:

H1a: Participants perceive an estimate of an independent informant as more accurate the more it converges with the estimates of other informants.

H1b: Participants perceive an independent informant as more competent the more their estimate converges with the estimates of other informants.

H2a: The effect of convergence on accuracy (H1a) is more positive in a context where informants are independent compared to when they are biased (i.e. share a conflict of interest to pick a given answer).

H2b: The effect of convergence on competence (H1b) is more positive in a context where informants are independent compared to when they are biased (i.e. share a conflict of interest to pick a given answer).

4.9.1 Methods

4.9.1.1 Participants

We ran a power simulation to inform our choice of sample size. All assumptions and details on the procedure can be found on the OSF. We ran two different power analyses, one for each outcome variable. We set the power threshold for both to 90%.

The power simulation for accuracy suggested that for 80 participants, we would have a power of at least 90% for the interaction effect. The simulation for competence suggested that with already 40 participants, we would detect an interaction, but only with 60 participants would we also detect an effect of convergence. Due to uncertainty about our assumptions and because resources were available for a larger sample, we recruited 200 participants, in the UK and via Prolific (99 female, 100, 1 non-identified; age_{mean}: 41.88, age_{sd}: 13.937, age_{median}: 39).

4.9.1.2 Procedure

After providing their consent to participate in the study and passing an attention check, participants read the following introduction: "We will show you three financial advisors who are giving recommendations on investment decisions. They can choose between three investment options. Their task is to recommend one. You will see several such situations. They are completely unrelated: it is different advisors evaluating different investments every time. At first you have no idea how competent the advisors are: they might be completely at chance, or be very good at the task. It's also possible that some are really good while others are really bad. Some tasks might be difficult while others are easy. Your task will be to evaluate the performance of one of the advisors based on what everyone's answers are." To assess perceptions of accuracy, we asked: "What do you think is the probability of advisor 1 making the best investment recommendation?". Participants answered with a slider on a scale from 0 to 100. To assess perceptions of competence, we asked: "How competent do you think advisor 1 is regarding such investment recommendations?" Participants answered on a 7-point Likert scale (from (1)"not competent at all" to (7)"extremely competent").

4.9.1.3 Design

We manipulated convergence within participants, and conflict of interest between participants. In the conflict of interest condition, experts were introduced this way: "The three advisors have already invested in one of the three options, the same option for all three. As a result, they have an incentive to push that option in their recommendations." Participants assigned to the independence condition read: "The three advisors are independent of each other, and have no conflict of interest in making investment recommendations."

4.9.1.4 Materials

We used the same stimuli as in Experiment 4. Identical to Experiment 4, participants saw all four convergence conditions, with two stimuli (i.e. expert predictions) per condition. Each participant therefore saw eight stimuli in total (4 convergence levels x 2 stimuli).

4.9.2 Results and discussion

To account for dependencies of observations due to our within-participant design, we ran mixed models, with a random intercept and a random slope for convergence for participants.

We find evidence for all four hypotheses (see Figure 4.6). To test H1a and H1b, we use the same analyses as in Experiment 4, restricted on the independence condition, and replicate the results. We find a positive effect of convergence on both accuracy ($\hat{b}_{\text{Accuracy}}$ = 12.337 [10.362, 14.311], p < .001) and competence ($\hat{b}_{\text{Competence}}$ = 0.562 [0.459, 0.665], p < .001).

The second set of hypotheses targeted the interaction of informational dependency and convergence (Figure 4.6). In the independence condition, the effect of convergence on accuracy was more positive ($\hat{b}_{\text{interaction, Accuracy}}$ = 3.008 [0.027, 5.988], p = 0.048) than in the conflict of interest condition ($\hat{b}_{\text{baseline}}$ 9.329 [7.232, 11.426], p < .001). Likewise, the effect of convergence on competence was more positive ($\hat{b}_{\text{interaction, Competence}}$ = 0.165 [0.014, 0.316], p = 0.032) than in the conflict of interest condition ($\hat{b}_{\text{baseline}}$ = 0.397 [0.291, 0.503], p < .001).

In an exploratory, non-preregistered analysis, we tested whether the effect of convergence is larger on accuracy than on competence. To do so, we first standardized both outcome scores to account for the different scales. We then regressed the outcome score on convergence and its interaction with a binary variable indicating which outcome was asked for (accuracy or competence), while controlling for informational dependency. We find a negative interaction, indicating that convergence had a smaller effect on competence than on accuracy ($\hat{b}$ = -0.079 [-0.113, -0.044], < .001; units in standard deviations).

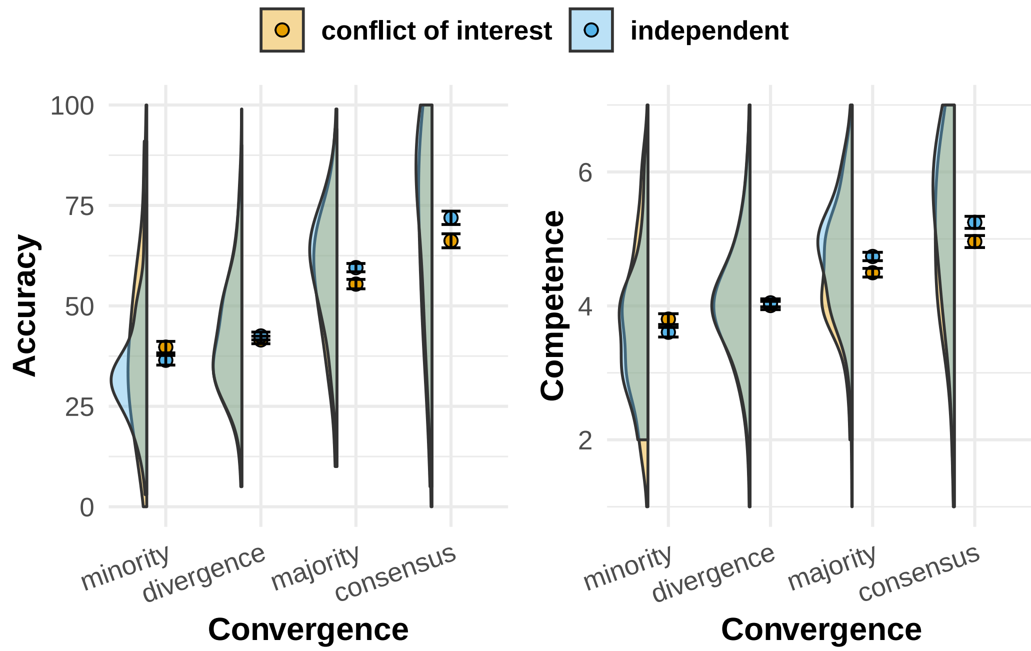

Figure 4.6: Interaction of convergence and informational dependency.

4.10 Experiment 6

Experiment 6 is a replication and extension of Experiment 4 in which we test the effect of the number of choice options (three and ten, instead of only three). Our simulations suggested that, at least for some underlying population competence distributions, consensus should be more indicative of competence when there are more choice options, compared to fewer (see ESM).

First, considering only the three options condition, we ran a direct replication of experiment 4. Second, following the results from our model, we predict that

H1: The effect of convergence on accuracy (H1a) is more positive in a context when informants can choose among ten response options compared to when they can choose among only three.

H2: The effect of convergence on competence (H1b) is more positive in a context when informants can choose among ten response options compared to when they can choose among only three.

4.10.1 Methods

4.10.1.1 Participants

We ran a power simulation to inform our choice of sample size. All assumptions and details on the procedure can be found on the OSF. We used previous experiments and estimates of our models to inform our choice of parameter values. We ran two different power analyses, one for each outcome variable. We set the power threshold for our experiment to 90%. The power simulation for accuracy suggested that for 140 participants we would cross the power threshold of 90% for the interaction effect (power = 0.928). The simulation for competence suggested that with 300 participants, we would detect an interaction with a power of 87%. Due to budget constraints, we considered aiming for a sample of 300 participants as good enough, although slightly below our threshold. Due to two failed attention checks, our final sample consisted of 298 subjects, recruited, as in all experiments, in the UK and via Prolific (149 female, 149; age_{mean}: 42.091, age_{sd}: 13.065, age_{median}: 40).

4.10.1.2 Procedure

We used the same procedure as in Experiment 4, with the addition of one condition described below.

4.10.1.3 Design

The number of choice options was manipulated between participants. Participants were randomly assigned to either see stimuli with three options (as in Experiment 4), or stimuli with ten options. Participants assigned to the ten options condition were divided into one of two distinct sub-conditions: one in which the range of the answers corresponds to the range of the three options condition, and another with increased range (see ESM). We found no differences between the two sub-conditions and collapsed them into a single ten options condition.

4.10.1.4 Materials

For the three options condition, we used the same stimuli as in Experiments 4 and 5. For the ten options condition, we created new sets of stimuli (see ESM). Identical to Experiments 4 and 5, participants saw all four convergence conditions, with two stimuli per condition. Each participant therefore saw eight stimuli in total.

4.10.2 Results and discussion

To account for dependencies of observations due to our within-participant design, we ran mixed models, with a random intercept and a random slope for convergence for participants.

We replicate the results of experiment 4, but do not find evidence for an interaction between convergence and the number of choice options. To match the setting of experiment one, we reduced the sample to half of the participants, namely those who were assigned to the three options condition. On this reduced sample, we ran the exact same analyses as in experiment 4 and replicated the results. We find a positive effect of convergence on both accuracy ($\hat{b}_{\text{Accuracy}}$ = 15.679 [14.112, 17.246], p < .001) and competence ($\hat{b}_{\text{Competence}}$ = 0.65 [0.564, 0.736], p < .001). This finding also holds across the entire sample, pooling three and ten choice option observations ($\hat{b}_{\text{Accuracy}}$ = 16.305 [15.124, 17.485], < .001; $\hat{b}_{\text{Competence}}$ = 0.675 [0.611, 0.739], < .001).

We do not find evidence of an interaction, i.e. evidence that the number of choice options changes the effect of convergence ($\hat{b}_{\text{interaction, Accuracy}}$ = 1.252 [-1.11, 3.613], = 0.298; $\hat{b}_{\text{interaction, Competence}}$ = 0.05 [-0.078, 0.178], = 0.442).

We tested whether the effect of convergence is larger on accuracy than on competence in an exploratory, non-preregistered analysis. To this end, we first standardized both outcome scores to account for the different scales. We then regressed the outcome score on convergence and its interaction with a binary variable indicating which outcome was asked for (accuracy or competence), while controlling for the number of choice options. In line with all previous experiments, we find a negative interaction, indicating that convergence had a smaller effect on competence than on accuracy ($\hat{b}$ = -0.049 [-0.075, -0.023], < .001).

Experiment 6 replicates Experiments 3 and 4 (independence condition), but suggests that the number of options the informants choose from does not powerfully affect participants' estimates of the informants' accuracy or competence. In the ESM, we show that, compared to our model, participants underestimate the effect of the number of choice options.

4.11 General discussion

Using both analytical arguments and simulations, we have shown that, under a wide range of parameters, more convergent answers tend to be more accurate, and to have been made by more competent informants.

In two experiments (Experiment 1, and independence condition of Experiment 3), we find that participants presented with a set of more (rather than less) convergent numerical estimates find the estimates more accurate, and the individuals making the estimates more competent, thus drawing normatively justified inferences. Experiment 2 suggests that participants do not think that a discussion between the individuals makes their convergence less indicative of accuracy or their competence. By contrast, Experiment 3 reveals that, when the individuals making the estimates are systematically biased by a conflict of interest, participants put less weight on the convergence of their estimates.

Similar results are obtained in a categorical choice context, in which participants see the answers of individuals made within a limited set of options. Experiments 4, 5 (independence condition), and 6 show that, the more the answers converge, the more participants believe them to be accurate and the individuals who made to be competent, again drawing normatively justified inferences. Experiment 5 shows that these inferences are weakened when the convergence can be explained by a conflict of interest (as in Experiment 3). Experiment 6 fails to find an effect of the number of options.

We also observe that, in line with our simulations, participants draw stronger inferences from convergence to accuracy than to competence in five of the six experiments.

On the whole, participants thus appear to draw normatively justified inferences: 1. They infer that more convergent answers are more accurate; 2. They infer that more convergent answers are coming from more competent informants; 3. The inference on accuracy tends to be stronger than the inference on competence ; 4. Both the inference on accuracy and on the inference on competence are weaker when the informants have a conflict of interest, but not when they are merely discussing with each other. The only exception appears to be that participants do not take into account the number of options in categorical choices scenarios.

4.12 Conclusion

When people see that others agree with each other, they tend to believe that they are right. This inference has been evidenced in several experiments, both for numerical estimates (e.g. Molleman et al. 2020; Yaniv, Choshen-Hillel, and Milyavsky 2009; Budescu and Rantilla 2000; Budescu et al. 2003; Budescu and Yu 2007), and for categorical choices (e.g., Morgan et al. 2012; for review, see Mercier and Morin 2019). However, these experiments do not test whether this inference of accuracy of the information extends to an inference of competence of the informants. Moreover, by their design, participants arguably assumed a degree of competence among the informants. For instance, when children are confronted with several individuals who agree on how to name a novel object (e.g. Corriveau, Fusaro, and Harris 2009), they can assume that these (adult) individuals tend to know what the names of objects are. If a certain competence of the informants is assumed, then well-known results from the literature on judgment aggregation–the wisdom of crowds–show that the average opinion of a set of individuals is, in a wide range of circumstances, more likely to be accurate than that of a single individual (e.g. Larrick and Soll 2006).

Here, we assumed no prior knowledge of individual competence, asking the question: if we see informants, whose competence is unknown, converge on an answer, is it rational to infer that this answer is more likely to be correct, and that the informants are likely to be competent?

We have shown that the answer is yes on both counts–assuming there is no systematic bias among the informants. An analytical argument and a series of simulations revealed that, for both the numerical choice context and the categorical choice context, the more individuals agree on an answer, the more likely the answer is correct, and the more likely the individuals are competent, with the former effect being stronger than the latter. Moreover, this is true for a wide range of assumed population distributions of competence. In a series of experiments, we have shown that participants (UK) draw these inferences, but that they do so less when there is reason to assume a bias among informants.

The results–both simulations and experiments–are a novel contribution to the wisdom of crowds literature. In this literature–in particular that relying on the Condorcet Jury Theorem–a degree of competence is assumed in the individuals providing some answers. From that competence, it can be inferred that the individuals will tend to agree, and that their answers will tend to be accurate. Here, we have shown that the reverse inference–from agreement to competence–is also warranted, and that it is warranted under a wide range of circumstances: If one does not suspect any systematic bias, convergence alone can be a valid cue when determining who tends to be an expert. This finding qualifies work suggesting that people need a certain degree of expertise themselves, in order to figure out who is an expert (Nguyen 2020; Hahn, Merdes, and Sydow 2018).

The present experiments show that people draw inferences from convergence to accuracy and to competence, but then do not precisely show how they do it. As suggested in the introduction, it is plausible that participants use the combination of two heuristics: one leading them from convergence to accuracy, and one from accuracy to competence. This two step process is coherent with the observation that the inference from convergence to accuracy (one step) is stronger than that from convergence to competence (two steps). However, our results are also compatible with participants performing two inferences from convergence, one to accuracy and one to competence, the latter being weaker because the inferences roughly follow the normative model. Our results are coherent with work on other mechanisms of epistemic vigilance that process the combination of several informants' opinions, showing that these mechanisms rely on sensible cues, but also systematically fail to take some subtle cues into account (Mercier and Morin 2019).

A most prominent context in which the inferences uncovered here might play a role is that of science. Much of science is counterintuitive, and most people do not have the background knowledge to evaluate most scientific evidence. However, science is, arguably, the institution in which individuals end up converging the most in their opinions (on consensus being the defining trait of science by contrast with other intellectual enterprises, see Collins 2002). For instance, scientists within many disciplines agree on things ranging from the distance between the solar system and the center of the galaxy to the atomic structure of DNA. This represents an incredible degree of convergence. When people hear that scientists have measured the distance between the solar system and the center of the galaxy, if they assume that there is a broad agreement within the relevant experts, this should lead them to infer that this measure is accurate, and that the scientists who made it are competent. Experiments have already shown that increasing the degree of perceived consensus among scientists tends to increase acceptance of the consensual belief (Van Stekelenburg et al. 2022), but it hasn't been shown yet that the degree of consensus also affects the perceived competence of scientists.

In the case of science, the relationship between convergence and accuracy is broadly justified. However, at some points of history, there has been broad agreement on misbeliefs, such as when Christian theologians had calculated that the Earth was approximately six thousand years old. To the extent that people were aware of this broad agreement, and believed the theologians to have reached it independently of each other, this might have not only fostered acceptance of this estimate of the age of the Earth, but also a perception of the theologians as competent.

The current study has a number of limitations. In our simulations, we assume agents to be independent and unbiased. Following previous work generalizing the Condorcet Jury Theorem to cases of informational dependency (Ladha 1992), more robust simulations would show that–while still assuming no systematic bias–our results hold even when agents influence each others' answers. Regarding our experiments, if the very abstract materials allow us to remove most of the priors that participants might have, they might also reduce the ecological validity of the results. Although the main results replicate well across our experiments, and we can thus be reasonably certain of their robustness, it's not clear how much they can be generalized. Experimental results with convenience samples can usually be generalized at least to the broader population the samples were drawn from–here, UK citizens (Coppock 2019). However, we do not know whether they would generalize to other cultures.

These limitations could be overcome by replicating the present results in different cultures, using more ecologically valid stimuli. For instance, it would be interesting to test whether the inference described here, from convergence to competence, might be partly responsible for the fact that people tend to believe scientists to be competent (Cologna et al. 2024). Finally, future studies could also attempt to systematically model cases of informational dependencies more subtle than those used in Experiments 3 and 5, to derive and test normative predictions.

4.12.0.1 Data availability

Data for all experiments and the simulations is available on the OSF project page (https://osf.io/6abqy/).

4.12.0.2 Code availability

The code used to create all results (including tables and figures) of this manuscript is also available on the OSF project page (https://osf.io/6abqy/).

4.12.0.3 Competing interest

The authors declare having no competing interests.

5 The French trust more the sciences they perceive as precise and consensual

Past research has shown that in the US, people's trust in science varies considerably between disciplines. Here, we show that this is the case also for France: A representative sample of the French population (N = 1,012) trusted researchers in biology and physics more than researchers studying climate science, economics, or sociology. We further show that trust differences within and between disciplines are associated with perceptions of consensus and precision: the more precise and consensual people perceive findings of a discipline to be, the more they tend to trust it. While these findings remain correlational, they are predicted by the rational impression account of trust in science, and they could have practical implications in terms of science communication.

> **i** available as a preprint here:
>
> Pfänder, J., & Mercier, H. (2025). *The French trust more the sciences they perceive as precise and consensual.* https://doi.org/10.31219/osf.io/k9m6e_v1

5.1 Introduction

Past research has shown that in the US, people's trust in science varies considerably between disciplines. For example, Altenmüller, Wingen, and Schulte (2024) found that among 20 different scientific disciplines, participants trusted (for instance) mathematicians, physicists and biologists more than climate scientists or historians, who in turn were more trusted than economists or political scientists. Analyzing trust in 45 different disciplines, Gligorić, Kleef, and Rutjens (2024) found that participants trusted neuroscientists the most, closely followed by various specialists in physics and biology. Psychologists, sociologists, economists and political scientists were trusted less.

Several explanations have been put forward to account for these differences. Altenmüller, Wingen, and Schulte (2024) showed that people assign different political inclinations to scientists in different disciplines. For example, people consider climate scientists to be liberal while they tended to see economists as more conservative (although all scientists of all disciplines were perceived to be rather liberal, with some as disciplines seen as moderate, but none as clearly conservative). Conservative participants trusted conservative scientists more and liberal participants would trust liberal scientists more. Gligorić, Kleef, and Rutjens (2024) have shown that people perceive some scientists–such as physicists–as more competent (than other scientists), while some–such as zoologists–as warmer, with both dimensions contributing to overall trust (see also Fiske and Dupree 2014). Laypeople also appear to believe that some disciplines are more 'scientific' than others–for example physics and biology by comparison with economics or sociology (Gauchat and Andrews 2018). On the whole, these beliefs about how scientific a discipline is track how much people trust it. However, it's unclear what leads to these beliefs in the first place.

The current research makes two contributions. First, past research on trust differences between scientific disciplines has focused on the US. We seek to extend the findings above to a different country: France. Although we expect to see broadly similar patterns, some differences with US results might also be obtained, for instance with regards to climate science, which tends to be particularly polarizing in the US (E. K. Smith, Bognar, and Mayer 2024; Huang 2021).

Second, we seek to better understand why some disciplines are more trusted than others, in particular by relying on the rational impression account of trust in science. This account suggests that people come to trust science because they are impressed by scientific findings, and that even if they then forget the specific findings, the impression of competence and trust persists (Pfänder, Rouilhan, and Mercier 2025). Different features of a scientific result can make it impressive: for instance, how difficult the objects being studied are to perceive (elementary particles, distant galaxies) and to understand (the human brain). However, there are two dimensions that should make scientific results more impressive across the board: being precise, and being consensual. All else being equal, providing a more precise estimate or understanding of one's object of study should be deemed more impressive–compare someone who tells you the universe is several billion years old, versus someone who tells you it is 13.7 billion years old. However, precision is only impressive if it is accurate–saying that the big bang happened on a Tuesday at 3.35pm is precise, but implausibly precise. What makes the figure of 13.7 billion years plausible, for someone who doesn't evaluate the relevant evidence first-hand? The fact that relevant scientists agree on this figure Vaupotič, Kienhues, and Jucks (2021). Even outside of a scientific context, people appear to have that intuition: people who converge more precisely on an estimate are deemed more likely to be correct, and more competent (Pfänder, De Courson, and Mercier 2025).

If people trust more scientific results they deem more impressive, and that to be impressive a result must be precise and consensual, the following two hypotheses should be true (both across and within disciplines, as well as for science in general):

H1: People trust scientists more when they are perceived as more consensual.

H2: People trust scientists more when they are perceived as more precise.

The present study has been conducted alongside a manylabs project on trust in science (Cologna et al. 2025; Mede et al. 2025). In addition to the main survey identical in all countries, we asked a representative sample of the French population about their trust in different scientific disciplines, as well as their impressions of these disciplines in terms of precision and consensus. Unless it is mentioned otherwise, the analyses (along with the hypotheses and materials) were pre-registered (https://osf.io/t6cbw). All data and code are publicly available via this article's OSF project page (https://osf.io/u8f3v/).

5.2 Method

5.2.1 Participants

We recruited a sample of 1012 participants representative of the French population (see Table 5.1)[1]. Participants were recruited from online panels by the market research company *Bilendi & respondi*. The survey was programmed with the survey software *Qualtrics*. Participants received vouchers or credit points for finishing the full survey. Participants had to be at least 18 years old and agree with the terms and conditions of the consent form. We excluded participants who failed to pass at least one of two attention checks, the first one consisting of of writing "213" into a text box at the beginning of the survey, the second of of selecting "strongly disagree" for an extra item in a scale of science-related populist attitudes towards the middle of the survey (for more details on the data collection, see Mede et al. 2025). Data was collected between February 21, 2023 and March 09, 2023.

Table 5.1: Summary table of sample demographics

	Overall, 1012 (100%)[1]
Gender	
Woman	502 (50%)
Man	500 (49%)
Prefer to self-describe	6 (0.6%)
Prefer not to say	4 (0.4%)
Age	
Mean (SD)	44.6 (14.9)

[1]Only in Study 4 do we find evidence that changing one's mind towards the scientific consensus is associated with (more) trust in science (Studies 1: r = 0.064, p = 0.387; 2: r = 0.161, p = 0.051; 3: r = 0.044, p = 0.619; 4: r = 0.148, p = 0.037) and only in Study 2 evidence that it is associated with (less) conspiracy beliefs and (less) conspiracy thinking (Studies 1: r = -0.139, p = 0.061; 2: r = -0.225, p = 0.006; 3: r = -0.044, p = 0.631; 4: r = -0.053, p = 0.455).

	Overall, 1012 (100%)[1]
Median	45.0
Min, Max	18.0, 88.0
Education	
Did not attend school	2 (0.2%)
Primary education	6 (0.6%)
Secondary education (e.g., high school)	350 (35%)
Higher education (e.g., university degree or higher education diploma)	654 (65%)
Income	
Mean (SD)	28,728.2 (39,415.5)
Median	25,000.0
Min, Max	0.0, 960,000.0
Residence	
Rural	484 (48%)
Urban	528 (52%)
Religious	
Not religious at all (1)	502 (50%)
(2)	169 (17%)
(3)	203 (20%)
(4)	103 (10%)
Very strongly religious (5)	34 (3.4%)

[1] n (%)

5.2.2 Measures

The complete questionnaires in English and French are available on the OSF project page. Here, we present only the measures used for the study at hand.

5.2.2.1 Trust in scientists in general

To measure trust in scientists in general, we follow Mede et al. (2025) who use a composite measure of twelve questions that are based on Besley, Lee, and Pressgrove (2021b) and cover four essential dimensions of trust in scientists: competence, integrity, benevolence, and openness (see Table 5.2). For more information on the psychometric properties of the trustworthiness scale, see Mede et al. (2025).

Using this scale presents a deviation from the preregistration, in which we had mentioned building a composite trust measure based on the benevolence and competence dimensions only. Since the pre-registration, however, the more faceted measure has been well validated and it would have been arbitrary to rely on only two dimensions. For full transparency, we show that all analyses replicate using our preregistered, reduced trust measure (Section 5.7.1, Table 5.5 and Table 5.6). We also replicate our findings using the same single item trust measure we use to assess trust in scientists of specific disciplines (Section 5.8.1, Table 5.7 and Table 5.8).

5.2.2.2 Perceptions of scientists across different disciplines

To measure trust in scientists across disciplines, participants were asked "How much do you trust the researchers working in these disciplines? Biology/Physics/Climate science/ Economics/Sociology/Science in general [1 = Do not trust at all, 5 = Trust very much]". We proceeded similarly for perceived precision ("How precise do you think the results obtained by researchers in these disciplines are? [1 = Not precise at all, 5 = Extremely precise]") and consensus ("How much do you think researchers in these disciplines agree on fundamental findings? [1 = They do not agree at all, 5 = They agree very much]").

5.2.3 Procedure

Participants first filled out a questionnaire containing 111 variables from the main manylabs survey (Mede et al. 2025), among which the questions about the trustworthiness of scientists in general. After that, they responded to the questions about differences between scientific disciplines.

5.3 Results

Figure 5.1 provides an overview of the results. On average, participants trusted scientists in general above the scale midpoint (of 3, mean = 3.457, sd = 0.59). They tended to trust biologists most (mean = 3.779, sd = 0.914), followed by physicists (mean = 3.682, sd = 0.922) and climate scientists (mean = 3.471, sd = 1.014). Below the score of trust in scientists in general was trust in sociologists (mean = 3.138, sd = 0.986) and economists (mean = 2.941, sd = 1.035).

To test the hypotheses we calculated three different types of models, all with trust as an outcome variable and either precision or consensus as predictor variable. The first type of model assessed associations of these three variables across all disciplines. Measures for all the scientific disciplines were pooled together and mixed effects models with random intercept and slope for participants were conducted. These models account for the fact that each participant provides several ratings, but they do not differentiate between disciplines. The second type of model assessed associations within disciplines. We added random effects for discipline to the previous models, both for the intercept and the slope. These models provide an average estimate of the within-discipline association. Additionally, they provide an estimate of how much this association varies between disciplines (the standard deviation of the random effects). For these two kinds of models, we restricted the data to only discipline-specific measures, excluding perceptions of scientists in general. The third type of model was used to measure the associations regarding science in general. For this purpose, we relied on perceptions of science in general, using the composite measure of trust described above. Since for perceptions of trust, precision, and consensus of scientists in general there was only one measure per participant, we calculated simple linear regression models.

Pooling across disciplines, people tended to trust the scientists of these disciplines more, the more they perceived the results of the scientists' discipline to be precise ($\hat{\beta}_{\text{across disciplines}} = 0.659$, $p < .001$), and consensual ($\hat{\beta}_{\text{across disciplines}} = 0.605$, $p < .001$). These results hold when looking at within discipline variation (precision: $\hat{\beta}_{\text{within disciplines}} = 0.588$, $p < .001$; consensus: $\hat{\beta}_{\text{within disciplines}} = 0.514$, $p < .001$). We also obtain a statistically significant associations when focusing on perceptions of scientists in general (precision: $\hat{\beta}_{\text{general}} = 0.383$, $p < .001$; consensus: $\hat{\beta}_{\text{general}} = 0.334$, $p < .001$).

In an exploratory, non-preregistered analysis, we investigated whether people's perceptions of precision and consensus differed. To do so, we pooled all consensus and precision ratings participants gave for the different disciplines (but not science in general) in a long-format data. We then ran a mixed model of a binary predictor variable (indicating whether precision of consensus was measured), on the outcome score. We added by-participant and by-discipline random intercepts. Since both variables were assessed on a 5-point Likert scale, we did not standardize the scores. We find that participants, on average, rated scientists as slightly more precise than consensual ($\hat{\beta}_{\text{precision vs. consensus}} = 0.044$, $p = 0.001$; on a scale from 1 to 5).

5.4 Discussion and conclusion

French participants trust more some scientific disciplines (biology and physics) than others (economics and sociology), with climate science falling in between. This is in line with the

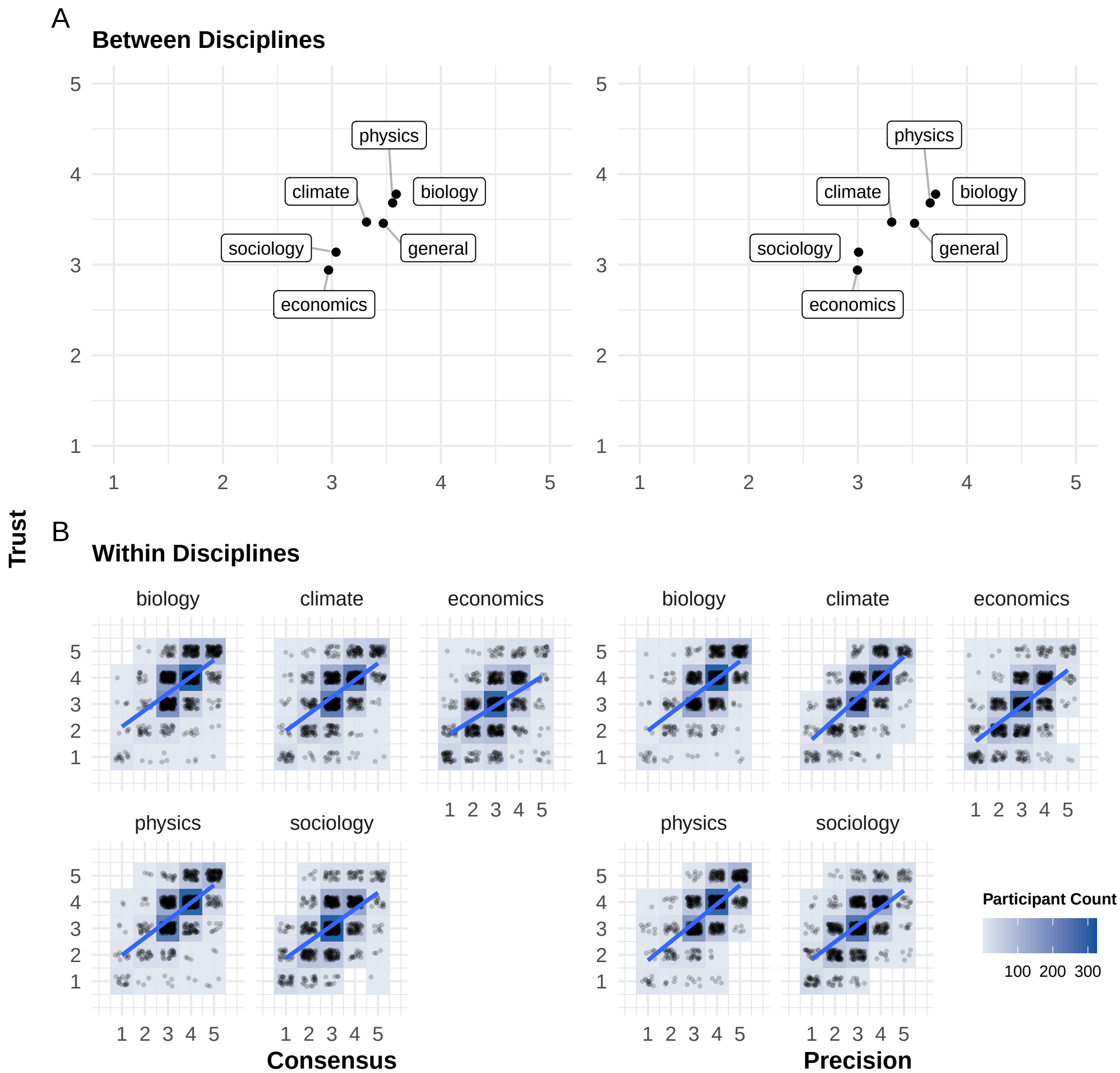

Figure 5.1: **A.** Plots show the association of consensus (left) and precision (right) with trust between the different disciplines. Dots represent average scores. For reference, we adde the score of scientists in general ('general'). **B.** Plots show the same associations but separately for all disciplines. Each dot represents a participant. Note that all scores are integer values ranging from 1 to 5, but dots are jittered within their respective quadrant for better visibility.

pattern observed in the US Gligorić, Kleef, and Rutjens (2024), suggesting some degree of cross-cultural robustness to how scientific disciplines are perceived.

The results also show that disciplines perceived as more precise and more consensual tend to be trusted more. We take this to support the rational impression account of trust in science, since this account predicts that science is trusted because it generates impressive knowledge, and that in turn precise and consensual knowledge tends to be more impressive. In this account, it would be (inter alia) because participants perceive a scientific discipline as precise and consensual that they trust it. However, the main issue with the current results is that they are only correlational. The causality could thus e.g. be reversed: people who, for other reasons, trust a scientific discipline might then say that it is more precise and more consensual, since they are both positive traits–a type of halo effect (for a review, see Forgas and Laham 2016).

Although it is impossible to rule out that the answers are partly driven by a halo effect, two lines of arguments suggest it would not account for all of the results. First, it has been shown that people have different impressions of the features of different disciplines, such that for instance they can deem rocket scientists (compared to other scientists) to be competent but less warm, while zoologists are thought to be warmer but less competent (Gligorić, Kleef, and Rutjens 2024). Similarly, in the present results for instance, participants deemed scientists to be slightly more precise than consensual, on average. Second, experiments demonstrating causal effects of consensus and precision (or impressiveness more generally) lend credence to a causal account of the present findings. Many experiments have shown that, as a rule, telling participants that a given scientific fact is consensual (e.g. the existence of anthropogenic climate change) increases acceptance of that belief (Van Stekelenburg et al. 2022; Većkalov et al. 2024). In turn, it is plausible that accepting that belief would lead to higher ratings of competence and trust of climate scientists. As for precision, experiments have shown that people tend to trust more a discipline after having been told of impressive results from that discipline–impressiveness being the broader trait relevant for the rational impression account, but it was also operationalized in part as precision in these experiments (Pfänder, Rouilhan, and Mercier 2025).

Besides the correlational nature of the data, the current study is also limited in only having recruited French participants, and not having questions about a broader range of disciplines, or more narrow subdisciplines.

The current results, and the rational impression account more generally, suggest ways to communicate about science that could affect trust in science in the public. Many studies have already attempted to increase the perception of consensus in specific scientific facts (Van Stekelenburg et al. 2022). However, communication could also stress how impressive some scientific disciplines are, for instance by stressing how precise some of their findings are. In the case of climate science, for examples, researchers have impressively precise findings, e.g. that over the past 485 million years, Earth's average temperature has varied between 11° and 36°C (Judd et al. 2024).

5.4.1 Data availability

Data is available on the OSF project page (https://osf.io/u8f3v/).

5.4.2 Code availability

The code used to create all results (including tables and figures) of this manuscript is also available on the OSF project page (https://osf.io/u8f3v/).

5.4.3 Competing interest

The authors declare having no competing interests.

5.5 Appendix A

5.5.1 Scale for measuring trust in scientists

Table 5.2: Items of the scale assessing general trust in scientists

Question	Answer Options	Mean (SD)
How open are most scientists to feedback?	Not open = 1, Somewhat open = 2, Neither open nor not open = 3, Somewhat open = 4, Very open = 5	3.103 (0.796)
How willing or unwilling are most scientists to be transparent?	Very unwilling = 1, Somewhat unwilling = 2, Neither willing nor unwilling = 3, Somewhat willing = 4, Very willing = 5	3.103 (0.979)
How much or little attention do scientists pay to others' views?	Very little attention = 1, Somewhat little attention = 2, Neither much nor little attention = 3, Somewhat much attention = 4, Very much attention = 5	2.967 (0.942)
How honest or dishonest are most scientists?	Very dishonest = 1, Somewhat dishonest = 2, Neither honest nor dishonest = 3, Somewhat honest = 4, Very honest = 5	3.441 (0.81)
How ethical or unethical are most scientists?	Very unethical = 1, Somewhat unethical = 2, Neither ethical nor unethical = 3, Somewhat ethical = 4, Very ethical = 5	3.359 (0.803)

Table 5.2: Items of the scale assessing general trust in scientists

Question	Answer Options	Mean (SD)
How sincere or insincere are most scientists?	Very insincere = 1, Somewhat insincere = 2, Neither sincere nor insincere = 3, Somewhat sincere = 4, Very sincere = 5	3.492 (0.828)
How expert or inexpert are most scientists?	Very inexpert = 1, Somewhat inexpert = 2, Neither expert nor inexpert = 3, Somewhat expert = 4, Very expert = 5	3.735 (0.726)
How intelligent or unintelligent are most scientists?	Very unintelligent = 1, Somewhat unintelligent = 2, Neither intelligent nor unintelligent = 3, Somewhat intelligent = 4, Very intelligent = 5	4.003 (0.717)
How qualified or unqualified are most scientists when it comes to conducting high-quality research?	Very unqualified = 1, Somewhat unqualified = 2, Neither qualified nor unqualified = 3, Somewhat qualified = 4, Very qualified = 5	3.971 (0.691)
How concerned or not concerned are most scientists about people's wellbeing?	Not concerned = 1, Somewhat not concerned = 2, Neither concerned nor not concerned = 3, Somewhat concerned = 4, Very concerned = 5	3.342 (0.882)
How eager or uneager are most scientists to improve others' lives?	Very uneager = 1, Somewhat uneager = 2, Neither eager nor uneager = 3, Somewhat eager = 4, Very eager = 5	3.632 (0.803)
How considerate or inconsiderate are most scientists of others' interests?	Very inconsiderate = 1, Somewhat inconsiderate = 2, Neither considerate nor inconsiderate = 3, Somewhat considerate = 4, Very considerate = 5	3.334 (0.893)

5.6 Appendix B

5.6.1 Regression tables main article

Table 5.3 and Table 5.4 show the regression tables for the results presented in the main article.

Table 5.3: Summary of mixed models with trust as outcome and consensus as predictor

	across_disciplines	within_disciplines	science_in_general_only
(Intercept)	1.423***	1.723***	2.298***
	(0.060)	(0.100)	(0.066)
consensus	0.605***	0.514***	0.334***
	(0.017)	(0.017)	(0.018)
SD (Intercept id)	1.156	1.138	
SD (consensus id)	0.308	0.292	
Cor (Intercept~consensus id)	−0.921	−0.911	
SD (Observations)	0.618	0.587	
SD (Intercept discipline)		0.179	
SD (consensus discipline)		0.008	
Cor (Intercept~consensus discipline)		1.000	
Num.Obs.	5052	5052	1012
R2			0.247
R2 Adj.			0.247
R2 Marg.	0.330	0.258	
R2 Cond.	0.620	0.629	
AIC	11 260.0	10 902.7	1521.6
BIC	11 299.1	10 961.5	1536.3
ICC	0.4	0.5	
Log.Lik.			−757.781
RMSE	0.55	0.52	0.51

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.4: Summary of mixed models with trust as outcome and precision as predictor

	across_disciplines	within_disciplines	science_in_general_only
(Intercept)	1.215***	1.457***	2.110***
	(0.055)	(0.107)	(0.070)
precision	0.659***	0.588***	0.383***
	(0.015)	(0.022)	(0.019)
SD (Intercept id)	1.096	1.104	
SD (precision id)	0.284	0.279	
Cor (Intercept~precision id)	−0.920	−0.915	
SD (Observations)	0.562	0.539	
SD (Intercept discipline)		0.204	
SD (precision discipline)		0.033	
Cor (Intercept~precision discipline)		−0.604	
Num.Obs.	5054	5054	1012
R2			0.276
R2 Adj.			0.276
R2 Marg.	0.413	0.353	
R2 Cond.	0.680	0.685	
AIC	10 421.7	10 163.3	1481.7
BIC	10 460.8	10 222.1	1496.4
ICC	0.5	0.5	
Log.Lik.			−737.836
RMSE	0.50	0.48	0.50

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 5.5: Summary of mixed models with trust as outcome, as preregistered using benevolence and competence only, and consensus as predictor

	across_disciplines	within_disciplines	science_in_general_only
(Intercept)	1.423***	1.723***	2.543***
	(0.060)	(0.100)	(0.067)
consensus	0.605***	0.514***	0.325***
	(0.017)	(0.017)	(0.019)
SD (Intercept id)	1.156	1.138	
SD (consensus id)	0.308	0.292	
Cor (Intercept~consensus id)	−0.921	−0.911	
SD (Observations)	0.618	0.587	
SD (Intercept discipline)		0.179	
SD (consensus discipline)		0.008	
Cor (Intercept~consensus discipline)		1.000	
Num.Obs.	5052	5052	1012
R2			0.230
R2 Adj.			0.230
R2 Marg.	0.330	0.258	
R2 Cond.	0.620	0.629	
AIC	11 260.0	10 902.7	1557.6
BIC	11 299.1	10 961.5	1572.4
ICC	0.4	0.5	
Log.Lik.			−775.818
RMSE	0.55	0.52	0.52

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.7 Appendix C

5.7.1 Regression tables preregistration

As stated in the main article, we had preregistered using a composite measure of only the competence and the benevolence dimensions (see Table 5.2). We report the results of our models using this measure in Table 5.5 and Table 5.6.

Table 5.6: Summary of mixed models with trust as outcome, as preregistered using benevolence and competence only, and precision as predictor

	across_disciplines	within_disciplines	science_in_general_only
(Intercept)	1.215***	1.457***	2.311***
	(0.055)	(0.107)	(0.071)
precision	0.659***	0.588***	0.386***
	(0.015)	(0.022)	(0.020)
SD (Intercept id)	1.096	1.104	
SD (precision id)	0.284	0.279	
Cor (Intercept~precision id)	−0.920	−0.915	
SD (Observations)	0.562	0.539	
SD (Intercept discipline)		0.204	
SD (precision discipline)		0.033	
Cor (Intercept~precision discipline)		−0.604	
Num.Obs.	5054	5054	1012
R2			0.278
R2 Adj.			0.277
R2 Marg.	0.413	0.353	
R2 Cond.	0.680	0.685	
AIC	10 421.7	10 163.3	1493.5
BIC	10 460.8	10 222.1	1508.3
ICC	0.5	0.5	
Log.Lik.			−743.752
RMSE	0.50	0.48	0.50

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.7: Summary of mixed models with trust as outcome, using the single-item measure, and consensus as predictor

	across_disciplines	within_disciplines	science_in_general_only
(Intercept)	1.423***	1.723***	1.612***
	(0.060)	(0.100)	(0.096)
consensus	0.605***	0.514***	0.588***
	(0.017)	(0.017)	(0.027)
SD (Intercept id)	1.156	1.138	
SD (consensus id)	0.308	0.292	
Cor (Intercept~consensus id)	−0.921	−0.911	
SD (Observations)	0.618	0.587	
SD (Intercept discipline)		0.179	
SD (consensus discipline)		0.008	
Cor (Intercept~consensus discipline)		1.000	
Num.Obs.	5052	5052	1012
R2			0.321
R2 Adj.			0.321
R2 Marg.	0.330	0.258	
R2 Cond.	0.620	0.629	
AIC	11 260.0	10 902.7	2294.4
BIC	11 299.1	10 961.5	2309.2
ICC	0.4	0.5	
Log.Lik.			−1144.206
RMSE	0.55	0.52	0.75

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

5.8 Appendix D

5.8.1 Regression tables single-item general trust

We had also assessed trust in scientists in general via a single item ("How much do you trust the researchers working in these disciplines? Science in general [1 = Do not trust at all, 5 = Trust very much]"), along with the measures for the different scientific disciplines. In Table 5.7 and Table 5.8 we report the results of our models when relying on this single-item trust question.

Table 5.8: Summary of mixed models with trust as outcome, using the single-item measure, and precision as predictor

	across_disciplines	within_disciplines	science_in_general_only
(Intercept)	1.215***	1.457***	1.154***
	(0.055)	(0.107)	(0.099)
precision	0.659***	0.588***	0.709***
	(0.015)	(0.022)	(0.027)
SD (Intercept id)	1.096	1.104	
SD (precision id)	0.284	0.279	
Cor (Intercept~precision id)	−0.920	−0.915	
SD (Observations)	0.562	0.539	
SD (Intercept discipline)		0.204	
SD (precision discipline)		0.033	
Cor (Intercept~precision discipline)		−0.604	
Num.Obs.	5054	5054	1012
R2			0.399
R2 Adj.			0.399
R2 Marg.	0.413	0.353	
R2 Cond.	0.680	0.685	
AIC	10 421.7	10 163.3	2171.3
BIC	10 460.8	10 222.1	2186.0
ICC	0.5	0.5	
Log.Lik.			−1082.626
RMSE	0.50	0.48	0.71

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6 Quasi-universal acceptance of basic science in the US

Many people report a low degree of trust in science, or endorse conspiracy theories that violate basic scientific knowledge. This might indicate a wholesale rejection of science in substantial segments of the population. In four studies, we asked 782 US participants questions about trust in science, conspiracy beliefs, and basic science (e.g. the relative size of electrons and atoms). Participants were provided with the scientifically consensual answer to the basic science questions, and asked whether they accept it. Acceptance of the scientific consensus was very high in the sample as a whole (95.1%), but also in every sub-sample (e.g., no trust in science: 87.3%; complete endorsement of flat Earth theory: 87.2%). This quasi-universal acceptance of basic science suggests that people are motivated to reject specific scientific beliefs, and not science as a whole. This could be leveraged in science communication.

available as a preprint here:

Pfänder, J., Kerzreho, L., & Mercier, H. (2025). *Quasi-universal acceptance of basic science in the US.* https://doi.org/10.31219/osf.io/qc43v_v2
For supplementary materials, please refer to the preprint.

6.1 Introduction

Trust in science is related to many desirable outcomes, from acceptance of anthropogenic climate change (Cologna and Siegrist 2020) or vaccination (Sturgis, Brunton-Smith, and Jackson 2021; Lindholt et al. 2021) to following recommendations during COVID (Algan et al. 2021, which suggests that trust in science was the most important predictor of these behaviors).

Although recent global evidence shows that trust in science is moderately high (Cologna et al. 2024), it is far from being at ceiling. Large-scale polls have shown that people who report a high degree of trust in science are a minority in most countries, and they are outnumbered by people who have low trust in science in many areas, e.g., most of Africa and significant parts of Asia (Wellcome Global Monitor 2018, 2020). Moreover, trust in science has recently been declining in some countries (Algan et al. 2021; Brian and Tyson 2023; although see Wellcome Global Monitor 2021; and Funk and Kennedy 2020) and in the US it is increasingly polarizing (Gauchat 2012; Krause et al. 2019; Li and Qian 2022).

Besides low answers on general trust in science questions, another indicator of distrust in science is the belief in conspiracy theories that question the scientific consensus on issues such as vaccination, climate change, and even the shape of the Earth. Conspiracy theories–in the realm of science or elsewhere–typically accuse a small group of powerful people to pursue nefarious goals in secrecy (Douglas et al. 2019; Mede and Schäfer 2020). Some of these conspiracy theories are widespread (Rutjens and Većkalov 2022). In 2023, a survey in eight different countries found that up to 24% of respondents agreed that "climate change is a hoax and scientists touting its existence are lying" (Stockemer and Bordeleau 2024). Similarly, in 2021, 40% of Americans believed that "the dangers of genetically-modified foods are being hidden from the public" (down from 45% in 2020, Uscinski et al. 2022). People who hold such views not only reject the relevant scientifically consensual facts, but also tend to believe in other conspiracy theories (Lewandowsky, Gignac, and Oberauer 2013; Hornsey, Harris, and Fielding 2018a, 2018b), and tend to say that they distrust science more generally (Vranic, Hromatko, and Tonković 2022; Stockemer and Bordeleau 2024). In spite of these correlations, the causal relationship between declaring general distrust in science and believing in one or several anti-science conspiracy theories is not clear. Although conspiracy thinking has been identified as a "root cause" of anti-science attitudes (Hornsey and Fielding 2017) in the past, these claims rest largely on observational data.

What does this apparent lack of trust in science actually entail? Do people who say they do not trust science, or who believe in conspiracy theories at odds with well-established science, reject most of science? Or, on the contrary, do they object to a few specific facets of science, while still accepting the overwhelming majority of basic science?

A common conception of trust is a willingness to be vulnerable to another party, whether an individual, a group, or an institution (Mayer, Davis, and Schoorman 1995; Rousseau et al. 1998). Accordingly, trust in science has been defined as "one's willingness to rely on science and scientists (as representatives of the system) despite having a bounded understanding of science" (Wintterlin et al. 2022, 2). Past research has disentangled this general concept of trust in science in various ways. Some research has identified different components of trust, the number of which varies, but which generally cover an epistemological and ethical dimension (Wilholt 2013; Intemann 2023). For example, Hendriks, Kienhues, and Bromme (2015) suggest distinguishing between expertise/competence, integrity, and benevolence, while Besley, Lee, and Pressgrove (2021a) add openness. Other research has highlighted differences in trust between scientific disciplines (Altenmüller, Wingen, and Schulte 2024; Gligorić, Kleef, and Rutjens 2024; Gauchat and Andrews 2018). However, little research has assessed trust in specific scientific findings, besides contentious topics such as vaccines (e.g. Hornsey, Harris, and Fielding 2018a), climate change (e.g. Stockemer and Bordeleau 2024), evolution (e.g. Nadelson and Hardy 2015), genetically modified organisms (GMOs) [e.g. ; Fernbach et al. (2019)], or a combination of such topics (see e.g. Lewandowsky, Gignac, and Oberauer 2013). To the best of our knowledge, no research has investigated the extent to which people trust basic science facts (e.g. electrons are smaller than atoms). An extensive literature on science literacy has assessed whether people know such facts (National Academies of Sciences, Engineering, and Medicine 2016), but not whether they accept the facts once presented to them.

Why does it matter whether trust in science is or is not related to trust in specific scientific findings? First, this question has theoretical implications. According to the most prominent explanation of (dis)trust in science–the deficit model–science knowledge is the main driver of attitudes towards science in general. A prediction of this model is that average trust in science on specific facts should be strongly associated with general trust in science. By contrast, a disconnect between the two would be in line with motivated reasoning accounts of trust in science. According to these accounts, science rejection serves to maintain coherence with other beliefs or behaviors (Lewandowsky and Oberauer 2016; Hornsey 2020). For example, someone might say they do not trust science in general because they are skeptical towards vaccines, not because they actually distrust most of science. A prediction of the motivated reasoning account is that general trust in science should be strongly correlated with conspiracy beliefs. Second, the present question has practical implications: Many communication attempts leverage the scientific consensus (e.g. on vaccination, climate change, etc., for review, see Van Stekelenburg et al. 2022; see also Većkalov et al. 2024). These attempts are more likely to be successful if everyone trusts basic science than if some people reject science wholesale.

6.1.1 The present studies

In a series of four pre-registered online studies (total n = 782), we asked US participants questions about well-established, consensual scientific facts. For each question, we asked participants what they thought the correct answer was (testing their knowledge of science), we informed them of the scientifically accepted answer, and asked them whether they accepted it (measuring their trust in basic science). We also measured participants' trust in science using standard measures, as well as their beliefs in various conspiracy theories and their tendency to engage in conspiratorial thinking. The four studies, including materials, hypotheses, and analyses, were pre-registered and all materials and data are accessible via the Open Science Framework (https://osf.io/8utsj/). The differences between the four studies are summarized presently, and the methods are detailed below.

6.1.1.1 Materials

In Study 1 (data collected on March 5, 2024), we used questions drawn from questionnaires of scientific knowledge (e.g. "Are electrons smaller, larger, or the same size as atoms? [Smaller; Same size; Larger]"), supplemented by a 'trick' question ("Where do trees mainly draw the materials with which they create their mass? [Earth; Water; Air]"; correct answer: Air). In Studies 2 (data collected on April 3, 2024) and 3 (data collected on April 22, 2024), this last question was removed. The scientific facts used in Studies 1 to 3 represent long-established and basic knowledge. In Study 4 (data collected on August 13, 2024) we used more recent, much less basic scientific discoveries (e.g. "What is the electric charge of the Higgs Boson, as established in 2012? [1.602176634 × 10-19; 0; 3.2×10-19C]"; correct answer: 0, i.e. electrically neutral).

6.1.1.2 Presentation of the scientific consensus

In Study 1, we simply told participants that they would be provided with the scientifically consensual answer. However, for participants to accept this answer, they must not only trust science, but also trust that we are presenting them with the actual scientifically consensual answer. To remove this issue, in Studies 2 to 3, we presented participants with a short explanation of the correct answer, as well as links to three sources per answer (e.g. Wikipedia, National Geographic or NASA). In Study 4, as the topics were more complex, we did not provide an explanation, but still provided two sources per answer.

6.1.1.3 Measure of acceptance of the scientific consensus

In Study 1, we simply looked at whether participants accept the scientifically consensual answer or not. In the subsequent studies, we asked participants to explain cases in which they disagreed with the scientific consensus. This revealed that a number of participants had made a mistake (misunderstanding, selecting the wrong answer). As a result, in Studies 3 and 4, participants who have indicated that they rejected the scientifically consensual answer were offered the option to revise their answer, or to keep rejecting it.

6.1.1.4 Additional questions

In Studies 3 and 4, we attempted to understand how some people who say they do not trust science still accept scientifically consensual answers, by asking them whether they accepted the answers on the basis of trust in science or because they had independently verified them.

6.1.1.5 Samples

Studies 1 and 2 were conducted on the standard sample of US participants recruited on the platform Prolific Academic. In order to increase the share of participants with low trust in science, and who endorse conspiracy theories, Studies 3 and 4 used the same platform, but only recruited participants who had declared previously being skeptical of vaccination.

6.1.1.6 Hypotheses

The main goal of the present studies is descriptive: to find out whether participants who report not trusting science, or who believe in conspiracy theories, still accept most well-established scientific facts. However, based on our literature review, we also tested two directional hypotheses (pre-registered as research questions in the Study 1):

H1: Higher trust in science is associated with more science knowledge and more acceptance of the scientific consensus

H2: Higher conspiracy thinking/belief is associated with less science knowledge and less acceptance of the scientific consensus

6.2 Methods

6.2.1 Deviations from preregistration

For Study 2, we restricted our main hypotheses about acceptance to cases in which participants initially provided a wrong answer. However, this meant the more participants had initially provided correct answers, the fewer opportunities they had for accepting correct answers. We provide results on these conditional correlations–for Study 2 and for all other studies–in the ESM [1]. However, for the analysis presented here, we proceeded as preregistered for all other studies, by reporting unconditional correlations between acceptance and trust in science, or, respectively, conspiracy belief.

6.2.2 Procedure

After providing their consent to participate in the study, participants were given an attention check "While watching the television, have you ever had a fatal heart attack?" [1-6; 1 = Never, 6 = Often]. All participants who did not answer "1 = Never" were excluded. Participants then read the following instructions:"We will ask you 10 questions about science. After each question, we will provide you with the scientifically consensual answer and ask whether you accept it." Next, participants answered a set of 10 basic science questions in random order. After each question, participants were presented with an answer reflecting the scientific consensus, and asked whether they accepted it. In Studies 2 and 3, participants additionally saw a short explanation, partly based on explanations generated by ChatGPT, and three links to authoritative sources supporting the answer. In Study 4, we provided only two links and no

[1] Only in Study 4 do we find evidence that changing one's mind towards the scientific consensus is associated with (more) trust in science (Studies 1: $r = 0.064$, $p = 0.387$; 2: $r = 0.161$, $p = 0.051$; 3: $r = 0.044$, $p = 0.619$; 4: $r = 0.148$, $p = 0.037$) and only in Study 2 evidence that it is associated with (less) conspiracy beliefs and (less) conspiracy thinking (Studies 1: $r = -0.139$, $p = 0.061$; 2: $r = -0.225$, $p = 0.006$; 3: $r = -0.044$, $p = 0.631$; 4: $r = -0.053$, $p = 0.455$).

explanation. Participants then answered questions on conspiracy thinking, conspiracy beliefs, and trust in science.

In Studies 2, 3, and 4, we presented participants with open-ended questions so they could explain their rejection of the scientific consensus. In Studies 3 and 4, we additionally gave participants the option to change their answer and accept the scientific consensus. Finally, at the end of Studies 3 and 4, we asked participants: "For the questions in which you agreed with the scientific consensus, would you say that…?" The answer options were: (i) "You mostly agree with the consensus because, on that question, you trust scientists", (ii) "You mostly agree with the consensus because you have been able to independently verify it", and (iii) "Other", with a text box for participants to explain. Participants who selected "You mostly agree with the consensus because you have been able to independently verify it", were asked the open-ended follow-up question: "Could you please tell us how you independently verified the information?".

6.2.3 Participants

After removing failed attention checks, the total sample size was 782 (194 in Study 1, six failed attention checks; 190 in Study 2, 11 failed attention checks; 200 in Study 3, no failed attention checks; 198 in Study 4, two failed attention checks) participants in the US, recruited through Prolific. Details and demographics can be found in the online supplemental material. While samples for Studies 1 and 2 were convenience samples, Studies 3 and 4 were conducted on a sample holding vaccine-skeptic beliefs. Prolific allows selecting participants based on their answers to a range of questions. We picked three of these questions and only recruited participants who met our criteria for each of them:

1. "Please describe your attitudes towards the COVID-19 (Coronavirus) vaccines: [For (I feel positively about the vaccines); Against (I feel negatively about the vaccines); Neutral (I don't have strong opinions either way); Prefer not to say]". We selected participants who answered "Against".
2. "Have you received a coronavirus (COVID-19) vaccination? [Yes (at least one dose); No; Prefer not to answer]". We select only people who answered "No".
3. "On a scale from 1-7, please rate to what extent you agree with the following statement: I believe that scheduled immunizations are safe for children. [1 (totally disagree); 2 (disagree); 3 (somewhat disagree); 4 (neither agree nor disagree); 5 (somewhat agree); 6 (agree); 7 (totally agree); rather not say]". We select only people who answered "1", "2", or "3".

6.2.4 Materials

6.2.4.1 Scientific facts

Studies 1 to 3 used 10 facts drawn from widely used questionnaires about science knowledge (Allum et al. 2008; Durant, Evans, and Thomas 1989; Miller 1998) sometimes referred to as the "Oxford scale" (Gauchat 2011). A 'trick' question was added in Study 1 and removed as its wording proved unclear. Study 4 used 10 more recent scientific discoveries. Table 6.1 shows all questions and their answer options.

Table 6.1: Science knowledge items

	Study 1-3	Study 4
1	Do antibiotics kill viruses as well as bacteria? [Yes, both; No, only viruses; No, only bacteria]	For which disease is the drug bedaquiline, developed in 2007, a treatment? [Tetanus; Tuberculosis; Malaria]
2	Are electrons smaller, larger, or the same size as atoms? [Smaller; Same size; Larger]	What is the maximum speed a proton can attain in the largest particle collider as to 2015? [90% of the speed of light; 99% of the speed of light; the speed of light]
3	Have the continents on Earth been moving for millions of years or have they always been where they are now? [They have been moving; They have always been where they are now]	Kepler-452b is an exoplanet revolving around the star Kepler-452. How far away from the star is it, as established by astronomers in 2015? [97 million mi; 1,2 million mi; 1254 million mi]
4	What decides whether a baby is a boy or a girl ? Is it the father's genes, the mother's genes, or both? [The mother's genes; the father's genes; both]	Using bomb-pusle dating with carbon 14, what is the age of the oldest known vertebrate, as established in 2016? [138 years; 205 years; 392 years]
5	Do lasers work by focusing sound waves? [Yes; No]	How many more glial cells are there in the brain in comparison with neurons, as established in 2016? [The same amount; Twice as many; Ten times as many]
6	How long does it take for Earth to go around the sun: one day, one month, or one year? [One day; One month; One year]	As predicted by the general theory of relativity, how many times would the Earth keep orbiting if the Sun disappeared, as established in 2012? [47 seconds; 8 minutes; 2 hours]
7	Are diamonds made of carbon? [Yes; No]	What is the electric charge of the Higgs Boson, as established in 2012? [1.602176634 x 10-19; 0; 3.2x10-19C]
8	Which travels faster : light or sound? [Light; Sound]	What is the age of the oldest materials formed on Earth, as established in 2020? [Less than 4.6 Ga; Around 4.6 Ga; More than 4.6 Ga]
9	Is common table salt made of calcium carbonate? [Yes; No]	With the best current cloning techniques, what is the average success rate when operated on mice, as of 2010? [2,7%; 9,4%; 17,2%]
10	Is water made of molecules containing one oxygen and two hydrogen atoms? [Yes; No]	What was the strength of the Earth magnetic field 3.7 billion years ago, as discovered this year? [15 microtesla; 30 microtesla; 45 microtesla]

11	*Where do trees mainly draw the materials with which they create their mass? [Earth; Water; Air]

* Only used in Study 1

6.2.4.2 Conspiracy beliefs

We selected 10 science/health related conspiracy theories from the Belief in Conspiracy Theory Inventory (BCTI) (Pennycook, Binnendyk, and Rand 2022) (Table 6.2). Participants were asked: "Below is a list of events for which the official version has been disputed. For each event, we would like you to indicate to what extent you believe the cover-up version of events is true or false. [1-9; labels: 1 - completely false, 5 - unsure, 9 - completely true]".

6.2.4.3 Conspiracy thinking

For all results presented here, we used the four-item conspiracy mentality questionnaire (CMQ) (Bruder et al. 2013). We also assessed the single item conspiracy beliefs scale (SICBS) (Lantian et al. 2016). Details and comparisons between the scales can be found in the ESM.

Table 6.2: Conspiracy items.

1	The Apollo moon landings never happened and were staged in a Hollywood film studio.
2	A cure for cancer was discovered years ago, but this has been suppressed by the pharmaceutical industry and the U.S. Food and Drug Administration (FDA).
3	The spread of certain viruses and/or diseases is the result of the deliberate, concealed efforts of vested interests.
4	The claim that the climate is changing due to emissions from fossil fuels is a hoax perpetrated by corrupt scientists who want to spend more taxpayer money on climate research.
5	The Earth is flat (not spherical) and this fact has been covered up by scientists and vested interests.
6	There is a causal link between vaccination and autism that has been covered up by the pharmaceutical industry.
7	In the 1950s and 1960s more than 100 million Americans received a polio vaccine contaminated with a potentially cancer-causing virus.
8	Proof of alien contact is being concealed from the public.
9	Hydroxychloroquine has been demonstrated to be a safe and effective treatment of COVID and this information is being suppressed.
10	Dinosaurs never existed, evolution is not real, and scientists have been faking the fossil record.

6.2.4.4 Trust in science

In all analyses reported in the main paper, we measure trust in science via a question selected from the Wellcome Global Monitor surveys (Wellcome Global Monitor 2018, 2020): "In general, would you say that you trust science a lot, some, not much, or not at all? [1 = Not at all, 2 = Not much, 3 = Some, 4 = A lot]". We chose this question as it seemed to be the most general one. In the ESM, we additionally report results for two alternative measures of trust included in our studies: Another from the WGM surveys ("How much do you trust scientists in this country? Do you trust them a lot, some, not much, or not at all? [1 = Not at all, 2 = Not much, 3 = Some, 4 = A lot]"), and one from the Pew Research Center (e.g. Funk, Johnson, and Hefferon 2019) ("How much confidence do you have in scientists to act in the best interests of the public? [1-5; 1 = No confidence at all, 5 = A great deal of confidence]"), the latter having been used in a recent international study on trust in science (Cologna et al. 2024). We selected these items so that we could compare the answers in our sample to global survey results. We find that all three items are highly correlated throughout all studies, and that our results reported here generally replicate when using either of the alternatives measures[2] (see ESM).

6.3 Results

The main outcome of interest is acceptance of the scientifically consensual facts presented. Overall, acceptance was very high (aggregating across all studies: 95.1 %; Studies 1: 93 %; 2: 98 %; 3: 98 %; 4: 91 %). Note that this includes both participants who had previously correctly answered the knowledge question, and participants who changed their mind when presented with the scientific consensus. In Studies 3 and 4, we gave participants a second chance in case they had initially rejected the consensus, which slightly increased acceptance rates in those studies (initial acceptance in Studies 3: 96 %; 4: 86 %).

As shown in Figure 6.1, these very high rates of acceptance hold for: participants who do not trust science at all (4.2 % of participants, acceptance rate of 87.3 %), participants who rank in the top two deciles of the conspiracy thinking scale (34.3 % of participants, acceptance rate of 93.2 %), participants who consider as "completely true" (the maximum of the 9-point scale) conspiracy theories stating that the earth is flat (3.7 % of participants, acceptance rate of 87.2 %), or that climate change due to fossil emissions is a hoax (11.4 % of participants, acceptance rate of 91.7 %).

Participants in the lowest decile of acceptance still had an average acceptance rate of 67.4 %. Even the three participants who considered as "completely true" that the earth is flat and who said they do "not trust science at all" had an average acceptance rate of 86.7 %.

These high acceptance rates do not merely reflect science knowledge: participants only correctly answered 65.8 % of the questions (Studies 1: 74 %; 2: 79 %; 3: 75 %; 4: 36 %) before they were provided with the scientifically consensual answer. Even the lowest decile in science

[2] With two exceptions: In Study 3, we find no correlation between acceptance and the Pew question; In Study 4, we find a correlation between knowledge and both alternative trust measures, but not with our main measure (see ESM)

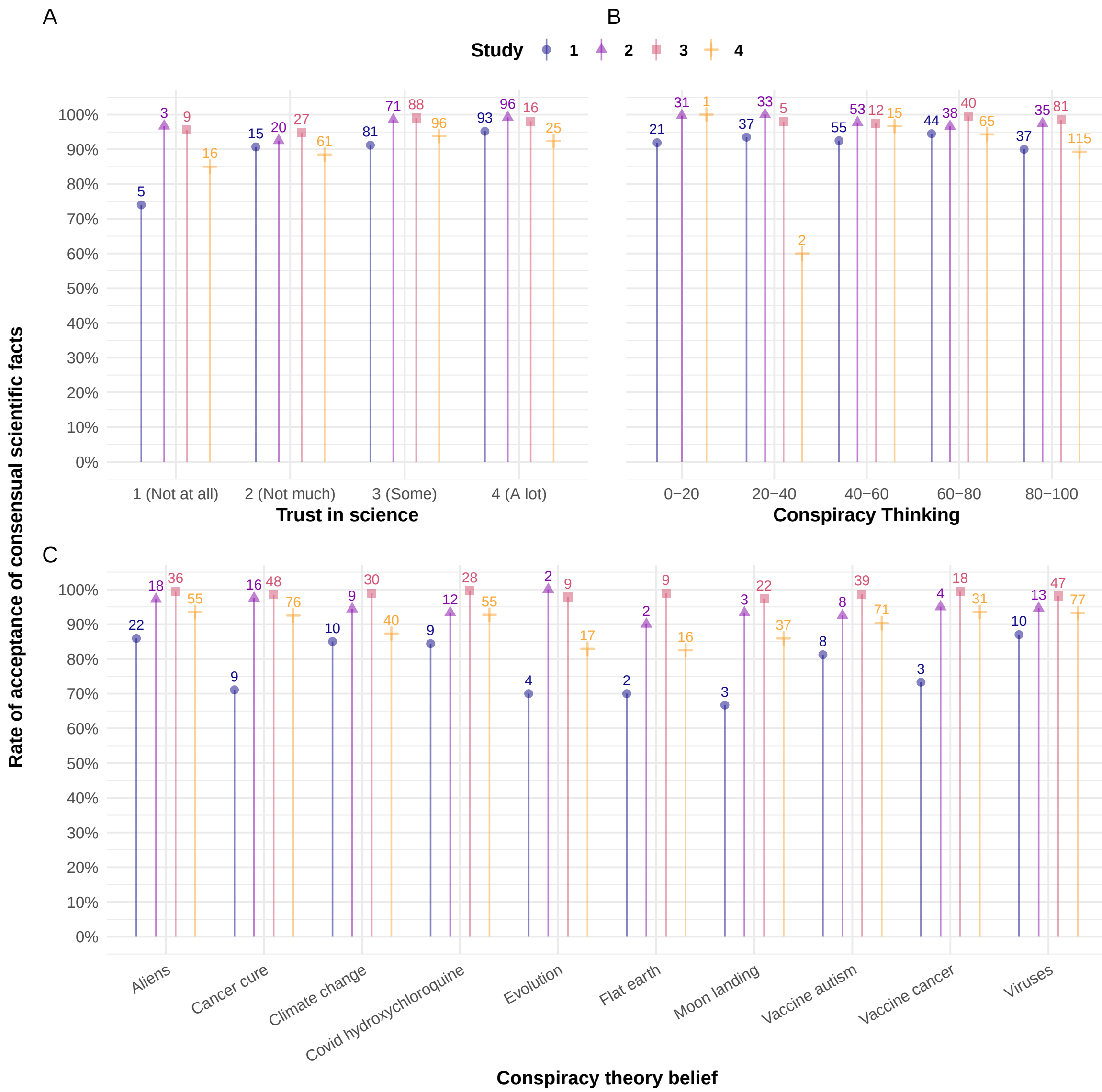

Figure 6.1: Points represent the average share of acceptance and numbers the absolute count of participants as a function of: **A** the level of trust in science ("In general, would you say that you trust science a lot, some, not much, or not at all? [1 = Not at all, 2 = Not much, 3 = Some, 4 = A lot]"); **B** the average conspiracy thinking (CMQ, five items on a scale from 0 to 100); **C** the belief in specific conspiracy theories (i.e. participants who answered 9, "completely true", for a given theory, see Table 6.2 for the list of the theories).

knowledge, which on average answered correctly only on 20 % of the questions, had an average acceptance rate of 91.7 %.

Did participants who had initially provided a wrong answer change their minds towards the scientific consensus? Yes. In most cases (Studies 1: 76.3 %; 2: 92.9 %; 3: 95.5 %; 4: 89.5 %), participants readily accepted the scientific consensus after having initially given the wrong answer to a question.

How do knowledge and acceptance relate to declared general trust in science and conspiracy belief (two strongly correlated variables, pooled $r = -0.636$, $p < .001$)?

Regarding H1, we find a consistent association between trust in science and acceptance of the scientific consensus (Studies 1: $r = 0.272$, $p < .001$; 2: $r = 0.299$, $p < .001$; 3: $r = 0.198$, $p = 0.019$; 4: $r = 0.155$, $p = 0.029$), but a less consistent relation between trust in science and science knowledge (Studies 1: $r = 0.29$, $p < .001$; 2: $r = 0.279$, $p < .001$; 3: $r = 0.14$, $p = 0.100$; 4: $r = 0.104$, $p = 0.144$).

Regarding H2, the results are mixed for the relation of conspiracy beliefs (measured as the average acceptance of all the conspiracy beliefs) with both acceptance of the scientific consensus (Studies 1: $r = -0.334$, $p < .001$; 2: $r = -0.371$, $p < .001$; 3: $r = -0.023$, $p = 0.788$; 4: $r = -0.048$, $p = 0.498$) and science knowledge (Studies 1: $r = -0.385$, $p < .001$; 2: $r = -0.402$, $p < .001$; 3: $r = -0.164$, $p = 0.055$; 4: $r = -0.023$, $p = 0.742$).

Why did participants reject the scientific consensus? We collected a total of 364 answers (Studies 2: 35; 3: 74; 4: 255) from 167 (Studies 2: 25; 3: 47; 4: 95) participants to the open-ended questions on why they had rejected the scientific consensus on a particular question. Based on the answers, we created five categories (Table 6.3). All individual answers can be accessed in cleaned data sheets via the OSF project page.

Table 6.3: Justifications for rejecting the scientific consensus by category, Studies 2, 3 and 4 combined.

Category	N (instances)	Share (instances)	N (participans)*
Not convinced	158	43.4%	70
No justification	103	28.3%	51
Personal convictions	43	11.8%	31
Mistake	42	11.5%	34
Religious Beliefs	18	4.9%	13

Note: *Participans with at least one answer in that category

Why did participants say they accept the scientific consensus? In Studies 3 and 4–the vaccine hesitant samples–we had asked participants about cases in which they agreed with the scientific consensus. A total of 320 (Studies 3: 122; 4: 198) participants answered this question. There were more participants saying they accepted the scientific consensus because they independently verified the fact (Studies 3: 47.5%; 4: 47%), than participants saying it was because they trust scientists (Studies 3: 41.8%; 4: 36.4%)[3]. Answers to a question about how they had done so can be found in the ESM.

In an exploratory analysis, we ran linear regressions to test whether there are differences between participants who said they had trusted science and those who said they had verified the information independently. Participants who said they accepted the consensus because of trust in scientists reported trusting science more (Studies 3: mean = 3; $\hat{\beta}_{\text{Trust}} = 0.192$, p = 0.330 on a scale from 1 to 4; 4: mean = 2.92; $\hat{\beta}_{\text{Trust}} = 0.465$, $p < .001$) than those who said they verified independently (Studies 3 mean = 2.81; 4: mean = 2.45). We did not find a difference regarding acceptance (Studies 3: $\hat{\beta}_{\text{Acceptance}} = 0.007$, p = 0.523 on a scale from 0 to 1; 4: $\hat{\beta}_{\text{Acceptance}} = 0.039$, p = 0.116) or regarding beliefs in conspiracy theories (Studies 3: $\hat{\beta}_{\text{BCTI}} = -0.163$, p = 0.694 on a scale from 1 to 9; 4: $\hat{\beta}_{\text{BCTI}} = -0.21$, p = 0.322). We also did not find a difference in time spent on the survey in Study 3 ($\hat{\beta}_{\text{Time}} = -0.02$, p = 0.985; median = 7.6416667 mins), but in Study 4, people who said they had accepted the consensus because they they trust scientists tended to spend on average two minutes less on the survey ($\hat{\beta}_{\text{Time}} = -2.159$, p = 0.018; median = 9.375 mins).[4] In Study 4, in which we used facts that participants were unlikely to have encountered before, we tracked whether people clicked on the source links provided–a behavior that you would expect from people who report verifying facts independently. On average, participants clicked only on 1.36 links (out of 20 possible clicks) and there was no difference between the two groups ($\hat{\beta}_{\text{Clicks}} = -1.091$, p = 0.062).

More detailed results addressing all our pre-registered research questions can be found in the ESM.

6.4 Discussion

In four studies, we asked US participants whether they accepted scientifically consensual answers on basic science questions. We found quasi-universal acceptance of basic science, with an overall rate of acceptance of 95.1 %, which remained very high for participants who declared not trusting science at all (87.3 % of acceptance), or who endorsed theories blatantly violating scientific knowledge, such as flat earth (87.2 % of acceptance).

This disconnect between declared general trust in science and average trust in (basic) scientific findings goes against predictions from the knowledge–attitudes model of trust in science (see, e.g., Bauer, Allum, and Miller 2007), which posits that knowledge of scientific results is the main cause of trust in science. In line with past research (Allum et al. 2008), we find that both knowledge and acceptance of science are only weakly correlated with general trust in science.

[3] 10.7% in Study 3 and 16.7% in Study 4 answered with “other” and gave an open-ended explanation.

[4] For these analyses, we excluded outliers who took over 30 mins for the survey, which was estimated to take around 10 mins, and for which the median time was seven minutes. As a result, we excluded one participant in Study 3 and four in Study 4. Significance levels are not affected by these exclusions.

This supports recent definitions of science literacy, which suggest that for science literacy to be a meaningful concept for science attitudes, it needs to go beyond measuring knowledge of isolated science facts (National Academies of Sciences, Engineering, and Medicine 2016).

By contrast, our findings align with a motivated reasoning account of trust in science (Lewandowsky and Oberauer 2016), in which "people tend to reject findings that threaten their core beliefs or worldview" (p. 217). A number of participants in our studies endorsed specific conspiracy theories questioning basic tenets of science (on evolution, the shape of the earth, etc.), while still accepting the vast majority of basic scientific facts presented to them. This suggests that these participants had reasons to reject only specific scientific knowledge, and that this rejection prompted them to express a lower trust in science when asked general questions on the topic. Consistent with this explanation, we found a strong association between belief in anti-science conspiracy theories and general trust in science.

Some of the present results also speak to the alienation model (Gauchat 2011), and more specifically to the need for epistemic autonomy (Fricker 2006). Declaring not trusting science, or endorsing conspiracy theories (Harris 2023) might reflect a desire to maintain epistemic autonomy and not appear to 'blindly' accept epistemic authority. Such a need to appear epistemically autonomous could be reconciled with the acceptance of basic science facts if it was thought to stem from independent evaluation instead of trust. The majority of participants in Studies 3 and 4, two samples of vaccine-skeptical participants who scored high across a range of different conspiracy beliefs (see ESM), tended to claim that they had accepted the scientific consensus based on their own evaluation, however implausible that might be: it is not clear how participants could independently verify, say, the ratio of glial cells to neurons. These participants did also not engage (more than others) in simple forms of verification, i.e. clicking on sources provided in the answers. This result aligns with other results on conspiracy thinking: Recent studies have shown that, although they claim not to be, conspiracy theorists are in fact just as susceptible to social influence as others (Pummerer et al. 2024; Altay et al. 2023).

In applied terms, the present results have implications for science communication. Across various domains of science knowledge such as climate change (Većkalov et al. 2024) or vaccination (Salmon et al. 2015), researchers have observed a consensus gap: a gap between the scientific consensus and public opinion. Some of the most worrying consensus gaps relate to climate change, as substantial segments of the population disagree with scientists on what is happening and what to do about it (e.g., Egan and Mullin 2017). Yet, we found that for much of basic, non-contentious science, such consensus gaps are absent: The vast majority of our participants–including two vaccine-skeptical samples (Studies 3 and 4)–did not appear to have general grounds for distrusting science, which should have led them to reject most or all of the science knowledge presented to them. Since even people who say they distrust science, or who reject specific scientific facts appear to accept most of basic science, stressing the basic science components of publicly controversial fields, from GMOs to climate change, might help reduce the consensus gaps observed in these domains (on climate change, see, e.g. Ranney and Clark 2016).

The present studies have a number of limitations, in particular the lack of representative samples, and the focus on a single country. While our results support motivated reasoning accounts of trust in science, they leave important questions unaddressed, in particular: Why do people reject specific scientific findings but not others? Suggestions have already been made for a number of issues such as vaccination (Miton and Mercier 2015), GMOs (Blancke et al. 2015), or nuclear energy (Hacquin et al. 2021). However, research is still needed to better understand what motivates these rejections (see e.g., Hornsey 2020). Moreover, rejection of scientific facts can manifest in various ways, for example calling into question the integrity of scientists (Mede 2023; Mede et al. 2022), or denying the very possibility of scientifically investigating certain issues (Munro 2010). The classification of justifications we present here does not address these processes in detail. Finally, by design, most of the basic science knowledge presented in this study didn't directly relate to anything controversial for most people. Future studies could look at basic science that does relate to scientific topics which are the object of public and policy-relevant debates (e.g. GM food, climate science, vaccines), or scientific findings for which there is a more obvious potential conflict of interest (e.g. non-publicly funded medical research).

6.4.1 Data availability

6.4.2 Code availability

The code used to create all results (including tables and figures) of this manuscript is also available on the OSF project page (https://osf.io/8utsj/).

6.4.3 Competing interest

The authors declare having no competing interests.

7 Spotting False News and Doubting True News, A Meta-Analysis of News Judgments

How good are people at judging the veracity of news? We conducted a systematic literature review and pre-registered meta-analysis of 303 effect sizes from 67 experimental articles evaluating accuracy ratings of true and fact-checked false news ($N_{participants}$ = 194'438 from 40 countries across 6 continents). We found that people rated true news as more accurate than false news (Cohen's d = 1.12 [1.01, 1.22], p < .001) and were better at rating false news as false than at rating true news as true (Cohen's d = 0.32 [0.24, 0.39], p < .001). In other words, participants were able to discern true from false news, and erred on the side of skepticism rather than credulity. We found no evidence that the political concordance of the news had an effect on discernment, but participants were more skeptical of politically discordant news (Cohen's d = 0.78 [0.62, 0.94], p < .001). These findings lend support to crowdsourced fact-checking initiatives, and suggest that, to improve discernment, there is more room to increase the acceptance of true news than to reduce the acceptance of fact-checked false news.

published as:

Pfänder, J., & Altay, S. (2025). *Spotting false news and doubting true news: A systematic review and meta-analysis of news judgements.* **Nature Human Behaviour**, 1–12. https://doi.org/10.1038/s41562-024-02086-1
For supplementary materials, please refer either to the open-access published version, or the preprint via the OSF.

7.1 Introduction

Many have expressed concerns that we live in a "post-truth" era and that people cannot tell the truth from falsehoods anymore. In parallel, populist leaders around the world have tried to erode trust in the news by delegitimizing journalists and the news media more broadly (Egelhofer et al. 2022). Since the 2016 US presidential election, our systematic literature review shows that over 4000 scientific articles have been published on the topic of false news. Across the world, numerous experiments evaluating the effect of interventions against misinformation or susceptibility to misinformation have relied on a similar design feature: having participants rate the accuracy of true and fact-checked false headlines–typically in a Facebook-like format, with an image, title, lede, and source, or as an isolated title/claim. Taken together, these studies allow us to shed some light on the most common fears voiced about false news, namely that people may fall for false news, distrust true news, or may be unable to discern between true and false news. In particular, we investigated whether people rate true news as more accurate than fact-checked false news (discernment) and whether they were better at rating false news as inaccurate than at rating true news as accurate (skepticism bias). We also investigated various moderators of discernment and skepticism bias such as political congruence, the topic of the news, or the presence of a source.

Establishing whether people can spot false news is important to design interventions against misinformation: if people lack the skills to spot false news, interventions should be targeted at improving skills to detect false news, whereas if people have the ability to spot false news but nonetheless engage with it, the problem lies elsewhere and may be one of motivation or (in)attention that educational interventions may struggle to address.

Past work has reliably shown that people do not fare better than chance at detecting lies because most verbal and non-verbal cues people use to detect lies are unreliable (Brennen and Magnussen 2023). Why would this be any different for detecting false news? People make snap judgments to evaluate the quality of the news they come across (Mont'Alverne et al. 2022), and rely on seemingly imperfect proxies such as the source of information, police and fonts, the presence of hyperlinks, the quality of visuals, ads, or the tone of the text (Metzger 2007; Ross Arguedas et al. 2022). In experimental settings, participants report relying on intuitions and tacit knowledge to judge the accuracy of news headlines (Altay, Lyons, and Modirrousta-Galian, n.d.). Yet, a scoping review of the literature on belief in false news (including a total of 26 articles) has shown that, in experiments, participants "can detect deceitful messages reasonably well" (Bryanov and Vziatysheva 2021, 19). Similarly, a survey on 150 misinformation experts has shown that 53% of experts agree that "people can tell the truth from falsehoods" – while only 25% of experts disagreed with the statement (Altay, Lyons, and Modirrousta-Galian, n.d.). Unlike the unreliable proxies people rely on to detect lies in interpersonal contexts, there are reasons to believe that some of the cues people use to detect false news may, on average, be reliable. For instance, the news outlets people trust the least do publish lower quality news and more false news, as people's trust ratings of news outlets correlate strongly with fact-checkers' ratings in the US and Europe (Pennycook and Rand 2019; Schulz, Fletcher, and Popescu 2020). Moreover, false news has some distinctive properties, such as being more politically slanted (Mourão and Robertson 2019), being more novel, surprising, or disgusting, being more sensationalist, funnier, less boring, and less negative (Vosoughi, Roy, and Aral 2018; X. Chen, Pennycook, and Rand 2023), or being more interesting-if-true (Altay, Araujo, and Mercier 2022). These features aim at increasing engagement, but they do so at the expense of accuracy, and in many cases, people may pick up on it. This led us to preregister the hypothesis that people would rate true news as more accurate than false news. Yet, legitimate concerns have been raised about the lack of data outside of the US, especially in some Global South countries where the misinformation problem is arguably worse. Our meta-analysis covers 40 countries across 6 continents and directly addresses concerns about the over-representation of US-data.

H1: People rate true news as more accurate than false news.

While many fear that people are exposed to too much misinformation, too easily fall for it, and are overly influenced by it, a growing body of researchers is worried that people are exposed to too little reliable information, commonly reject it, and are excessively resistant to it (Acerbi, Altay, and Mercier 2022; Mercier 2020). Establishing whether true news skepticism (excessively rejecting true news) is of similar magnitude to false news gullibility (excessively accepting false news) is important for future studies on misinformation: if people are excessively gullible, interventions should primarily aim at fostering skepticism, whereas if people are excessively skeptical, interventions should focus on increasing trust in reliable information. For these reasons, in addition to investigating discernment (H1), we also looked at skepticism bias by comparing the magnitude of true news skepticism to false news gullibility. Research in psychology has shown that people exhibit a "truth bias" (Brashier and Marsh 2020; Street and Masip 2015), such that they tend to accept incoming statements rather than reject them. Similarly, work on interpersonal communication has shown that, by default, people tend to accept communicated information (Levine 2014). However, there are reasons to think that the truth-default-theory may not apply to news judgments. It has been hypothesized that people display a truth bias in interpersonal contexts because information in these contexts is, in fact, often true (Brashier and Marsh 2020). When it comes to news judgments, it is not clear that people by default expect news stories to be true. Trust in the news and journalists is low worldwide (Newman et al. 2022), and a significant part of the population holds cynical views of the news (Mihailidis and Foster 2021). Similarly, populist leaders across the world have attacked the credibility of the news media and instrumentalized the concept of fake news to discredit quality journalism (Egelhofer and Lecheler 2019; Van Duyn and Collier 2019). Disinformation strategies such as "flooding the zone" with false information (Paul and Matthews 2016; Ulusoy et al. 2021) have been shown to increase skepticism in news judgments (Altay, Lyons, and Modirrousta-Galian, n.d.). Moreover, in many studies included in our meta-analysis, the news stories were presented in a social media format (most often Facebook), which could fuel skepticism in news judgments. Indeed, people trust news (Mont'Alverne et al. 2022)–and information more generally (Fletcher and Nielsen 2017)–less on social media than on news websites. In line with these observations, some empirical evidence suggests that for news judgments, people display the opposite of a truth bias (Luo, Hancock, and Markowitz 2022), namely a skepticism bias, whereby people tend to rate all news as more false than they are (Altay, Lyons, and Modirrousta-Galian, n.d.; Batailler et al. 2022; Modirrousta-Galian and Higham 2023). We thus predicted that when judging the accuracy of news, participants will err on the side of skepticism more than on the side of gullibility.

H2: People are better at rating false news as false than true news as true.

Finally, we investigated potential moderators of H1 and H2, such as the country where the experiment was conducted, the format of the news headlines, the topic, whether the source of the news was displayed, and the political concordance of the news. Past work has suggested that displaying the source of the news has a small effect at best on accuracy ratings (Dias, Pennycook, and Rand 2020), whereas little work has investigated differences in news judgments across countries, topics, and formats. The effect of political concordance on news judgments is debated. Participants may be motivated to believe politically congruent (true and false) news, motivated to disbelieve politically incongruent news, or not be politically motivated at all but still display such biases (Tappin, Pennycook, and Rand 2020). We formulated research questions instead of hypotheses for our moderator analyses because of a lack of strong theoretical expectations.

7.2 Results

7.2.1 Descriptives

We conducted a systematic literature review and pre-registered meta-analysis based on 67 publications, providing data on 195 samples (194438 participants) and 303 effects (i.e. **k**, the meta-analytic observations). Our meta-analysis includes publications from 40 countries across 6 continents. However, 34% of all participants were recruited in the United States alone, and 54% in Europe. Only 6% of participants were recruited in Asia, and even less in Africa (2%; see Figure 7.1 for the number of effect sizes per country). The average sample size was 997.12 (min = 19, max = 32134, median = 482).

In total, participants rated the accuracy of 2167 unique news items. On average, a participant rated 19.76 news items per study (min = 2, max = 240, median = 18). For 71 samples, news items were sampled from a pool of news (the pool size ranged from 12 to 255, with an average pool size of 57.46 items). The vast majority of studies (294 out of 303 effects) used a within participant design for manipulating news veracity, with each participant rating both true and false news items. Almost all effect sizes are from online studies (286 out of 294).

(ref:map) A map of the number of effect sizes per country.

7.2.2 Analytic procedures

All analyses were pre-registered unless explicitly stated otherwise (for deviations see methods section). The choice of models was informed by simulations we conducted before having the data. To test H1, we calculated a discernment score by subtracting the mean accuracy ratings of false news from the mean accuracy ratings of true news, such that higher scores indicate better discernment. This differential measure of discernment is common in the literature on misinformation (Guay et al. 2023). To test H2, we first calculated a judgment error for true and false news respectively. Error is defined as the distance between optimal accuracy ratings and actual accuracy ratings (see Figure 7.2). We then calculate the skepticism bias as the difference between the two errors, subtracting the false news error score from the true news error score.

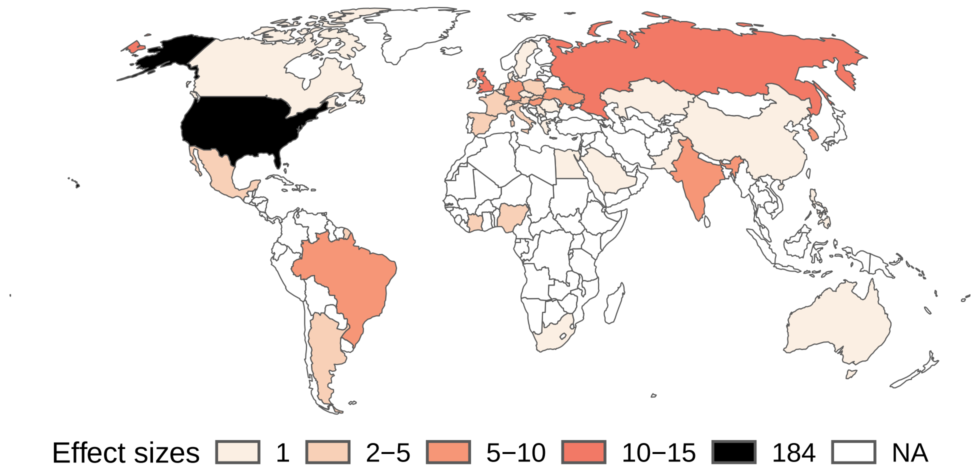

Figure 7.1: A map of the number of effect sizes per country.

Note that we cannot use more established Signal Detection Theory (SDT) measures, because we rely on mean ratings and not individual ratings. However, in the appendix, we show that for the studies we have raw data on, our main findings hold when relying on d' (sensitivity) and c (response bias) from SDT.

To be able to compare effect sizes across different scales, we calculated Cohen's d, a common standardized mean difference. To account for statistical dependence between true and false news ratings arising from the within-participant design used by most studies (294 out of 303 effect sizes), we calculated the standard error following the Cochrane recommendations for crossover trials (Higgins et al. 2019). For the remaining 9 effect sizes from studies that used a between-participant design, we calculated the standard error assuming independence between true and false news ratings (see methods). In the appendix, we show that our results hold across alternative standardized effect measures, among which the one we had originally pre-registered, a standardized mean change using change score standardization (SMCC). We chose to deviate from the pre-registration and use Cohen's d instead, because it is easier to interpret and corresponds to the standards for crossover trials recommended by the Cochrane manual (Higgins et al. 2019). In the appendix, we also provide effect estimates in units of the original scales separately for each scale.

We used multilevel meta models with clustered standard errors at the sample level to account for cases in which the same sample contributed various effect sizes (i.e. the meta-analytic units of observation). All confidence intervals reported in this paper are 95% confidence intervals.

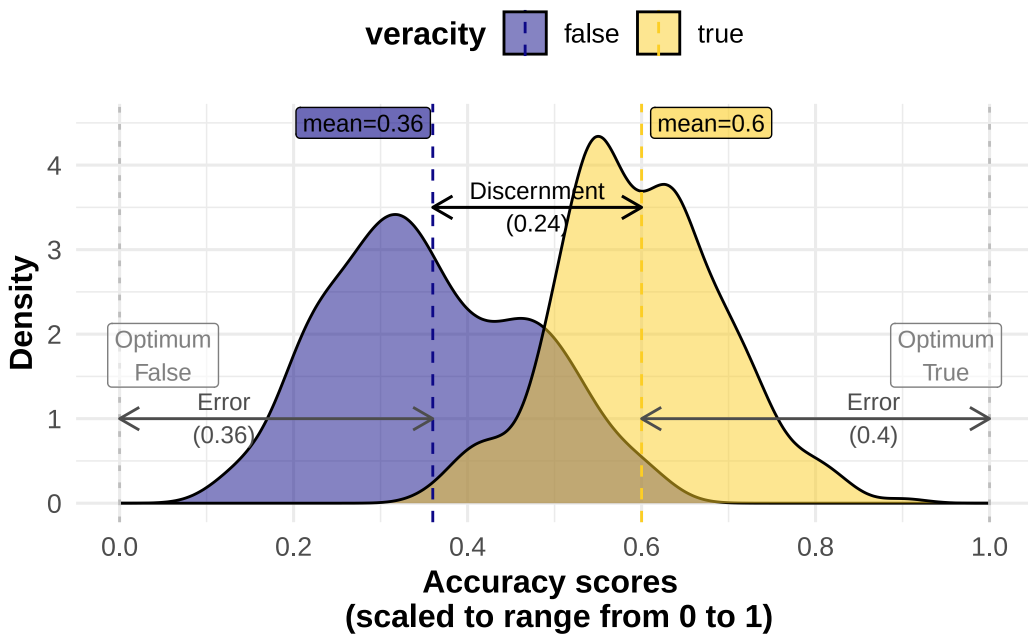

Figure 7.2: *Illustration of outcome measures.* The figure shows the distributions of accuracy ratings for true and fact-checked false news, scaled to range from 0 to 1. The figure illustrates discernment (the distance between the mean for true news and the mean for false news) and the errors (distance to the right end for true news and to the left end for false news) from which the skepticism bias is computed. A larger error for true news compared to false news yields a positive skepticism bias. In this descriptive figure, unlike in the meta-analysis, ratings and outcomes sizes are not weighted by sample size.

All statistical tests are two-tailed.

7.2.3 Main results

7.2.3.1 Discernment (H1)

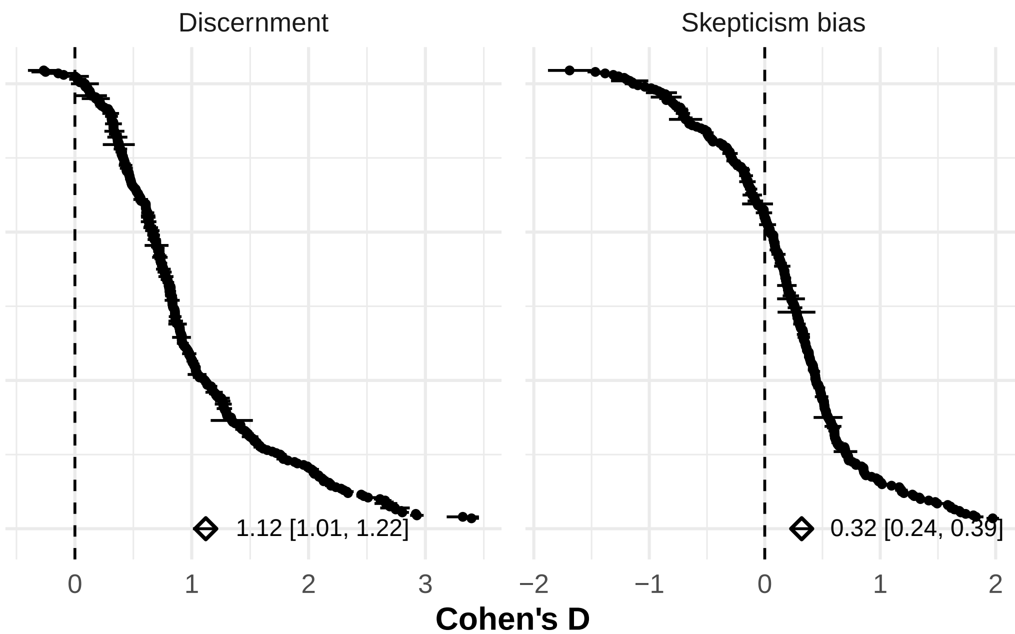

Figure 7.3: *Forest plots for discernment and skepticism bias.* The figure displays all n = 303 effect sizes for both outcomes. Effects are weighed by their sample size. Effect sizes are calculated as Cohen's d. Horizontal bars represent 95% confidence intervals. The average estimate is the result of a multilevel meta model with clustered standard errors at the sample level.

Supporting H1, participants rated true news as more accurate than false news on average. Pooled across all studies, the average discernment estimate is large (d = 1.12 [1.01, 1.22], z = 20.79, $p < .001$). As shown in Figure 7.3, 298 of 303 estimates are positive. Of the positive estimates, 3 have a confidence interval that includes 0, as does 1 of the negative estimates. Most of the variance in the effect sizes observed above is explained by between-sample heterogeneity ($I2_{between} = 92.04\%$). Within-sample heterogeneity is comparatively small ($I2_{within} = 7.93\%$), indicating that when the same participants were observed on several occasions (i.e. the same sample contributed several effect sizes), on average, discernment performance was similar across those observations. The share of the variance attributed to sampling error is very small (0.03%), which is indicative of the large sample sizes and thus precise estimates.

7.2.3.2 Skepticism bias (H2)

We found support for H2, with participants being better at rating false news as inaccurate than at rating true news as accurate (i.e. false news discrimination was on average higher than true news discrimination). However, the average skepticism bias estimate is small (d = 0.32 [0.24, 0.39], z = 8.11, $p < .001$). As shown in Fig Figure 7.3), 203 of 303 estimates are positive. Of the positive estimates, 6 have a confidence interval that includes 0, as do 7 of the negative estimates. By contrast with discernment, most of the variance in skepticism bias is explained by within-sample heterogeneity ($I2_{within}$ = 60.96%; $I2_{between}$ = 38.99%; sampling error = 0.05%). Whenever we observe within sample variation in our data, it is because several effects were available for the same sample. This is mostly the case for studies with multiple survey waves, or when effects were split by different news topics, suggesting that these factors may account for some of that variation. In the moderator analyses below, most variables vary between samples, thereby glossing over much of that within-variation. An exception is political concordance.

7.2.4 Moderators

Following the pre-registered analysis plan, we ran a separate meta regression for each moderator by adding the respective moderator variable as a fixed effect to the multilevel meta models. We report regression tables and visualizations in the appendix. Here, we report the regression coefficients as "Delta"s, since they designate differences between categories. For example, in the moderator analysis of political concordance on skepticism bias, "concordant" marks the baseline category. The predicted value for this category can be read from the intercept (-.2). The "Delta" is the predicted difference between concordant and discordant (.78). To obtain the predicted value for discordant news, one needs to add the "Delta" to the intercept (-.2 + .78 = .58).

7.2.4.0.1 Cross-cultural variability

For samples based in the United States (184/303 effect sizes), discernment was higher than for samples based in other countries, on average (Δ Discernment = 0.23 [0.02, 0.44], z = 2.14 , p = 0.033 ; baseline discernment other countries pooled = 0.99 [0.84, 1.14], z = 12.82, $p < .001$). However, we did not find a statistically significant difference regarding skepticism bias (Δ Skepticism bias = 0.04 [-0.12, 0.19], z = 0.47 , p = 0.638). A visualization of discernment and skepticism bias across countries can be found in the appendix.

7.2.4.0.2 Scales

The studies in our meta analysis used a variety of accuracy scales, including both binary (e.g. "Do you think the above headline is accurate? - Yes, No") and continuous ones (e.g. "To the best of your knowledge, how accurate is the claim in the above headline" 1 = Not at all accurate, 4 = Very accurate).

Regarding discernment, two scale types differed from the most common 4-point scale (Baseline discernment 4-point-scale = 1.28 [1.07, 1.49], z = 11.96, $p < .001$): Both 6-point scales (Δ Discernment = -0.41 [-0.7, -0.12], z = -2.8, p = 0.006) and binary scales (Δ Discernment = -0.37 [-0.66, -0.08], z = -2.5, p = 0.013) yielded lower discernment. Regarding skepticism bias, studies using a 4-point scale (Baseline skepticism bias 4-point scale = 0.51 [0.3, 0.72], z = 4.75, $p < .001$) reported a larger skepticism bias compared to studies using a binary and a 7-point scale (Δ Skepticism bias = -0.29 [-0.51, -0.06], z = -2.47, p = 0.014 for binary scales; -0.5 [-0.76, -0.23], z = -3.67, $p < .001$ for 7-point scales). Interpreting these observed differences is not straightforward. We attempt a more detailed discussion of differences between binary and Likert-scale studies in the appendix.

7.2.4.0.3 Format

Studies using headlines with pictures as stimuli (Δ Skepticism bias = 0.22 [0.04, 0.39], z = 2.45, p = 0.015; 65 effects), or headlines with pictures and a lede (Δ Skepticism bias = 0.33 [0.14, 0.52], z = 3.4, $p < .001$; 56 effects), displayed a stronger skepticism bias compared to studies relying on headlines with no picture/lede (Baseline skepticism bias headlines only = 0.23 [0.13, 0.33], z = 4.45, $p < .001$; 163 effects). We do not find differences related to format for discernment, neither for headlines with pictures (Δ Discernment = -0.01 [-0.28, 0.27], z = -0.04, p = 0.969), nor for headlines with pictures and a lede (Δ Discernment = 0.11 [-0.12, 0.33], z = 0.93, p = 0.353).

7.2.4.0.4 Topic

We did not find statistically significant differences in discernment and skepticism bias across news topics, when distinguishing between the categories "political" (Δ Skepticism bias = 0.03 [-0.13, 0.19], z = 0.43, p = 0.671; Δ Discernment = -0.26 [-0.51, 0], z = -1.98, p = 0.049; 196 effects; 43 articles), "covid" (baseline; 54 effects; 13 articles) and "other" (Δ Skepticism bias = -0.02 [-0.2, 0.16], z = -0.22, p = 0.825; Δ Discernment = -0.01 [-0.35, 0.34], z = -0.03, p = 0.976; 53 effects; 20 articles), a category which regroups all not explicitly as "covid"or "political" labeled news topics by the authors for the respective papers, and which includes news topics reaching from health, cancer and science, to economics, history and military matters.

7.2.4.0.5 Sources

In line with past findings, we did not observe a statistically significant difference in discernment between studies displaying the source of the news items (Δ Discernment = -0.22 [-0.47, 0.03], z = -1.75, p = 0.082; 112 effects) and studies that did not (147 effects; for 44 this information was not explicitly provided). We do not find a difference regarding skepticism bias either (Δ Skepticism bias = 0.11 [-0.06, 0.29], z = 1.3, p = 0.194).

7.2.4.0.6 Political Concordance

The moderators investigated above were (mostly) not experimentally manipulated within studies, but instead varied between studies, which impedes causal inference. Political concordance is an exception in this regard. It was manipulated within 31 different samples, across 14 different papers. In those experiments, typically, a pre-test establishes the political slant of news headlines (e.g. pro-republican vs. pro-democrat). In the main study, participants then rate the accuracy for news items of both political slants, and provide information about their own political stance. The ratings of items are then grouped into concordant or discordant (e.g. pro-republican news rated by Republicans will be coded as concordant while pro-republican news rated by Democrats will be coded as discordant).

Political concordance had no statistically significant effect on discernment (Δ Discernment = 0.08 [-0.01, 0.17], z = 1.72, p = 0.097). It did, however, make a difference regarding skepticism bias (see Figure 7.4): When rating concordant items, there was no evidence that participants showed a skepticism bias (Baseline skepticism bias concordant items = -0.2 [-0.42, 0.01], z = -1.93, p = 0.064), while for discordant news items, participants displayed a positive skepticism bias (Δ Skepticism bias = 0.78 [0.62, 0.94], z = 10.04, p < .001). In other words, participants were not gullible when facing concordant news headlines (as would have suggested a negative skepticism bias), but were skeptical when facing discordant ones.

7.2.5 Individual level data

In the results above, accuracy ratings were averaged across participants. It is unclear how these average results generalize to the individual level. Do they hold for most participants? Or are they driven by a relatively small group of participants with excellent discernment skills, or, respectively, extreme skepticism? For 22 articles ($N_{Participants}$ = 42074, $N_{Observations}$ = 813517), we have the raw data for all ratings that individual participants made on each news headline they saw. On this data, we ran a descriptive, non-preregistered analysis: We calculated a discernment and skepticism bias score for each participant based on all the news items they were rating. To compare across different scales, we transposed all accuracy scores on a scale from 0 to 1, resulting in a range of possible values from -1 to 1 for both discernment and skepticism bias.

As shown in Figure 7.5, 79.92 % of individual participants had a positive discernment score, and 59.06 % of participants had a positive skepticism bias score. Therefore, our main results based on mean ratings across participants seem to be representative of individual participants (see appendix for further discussion).

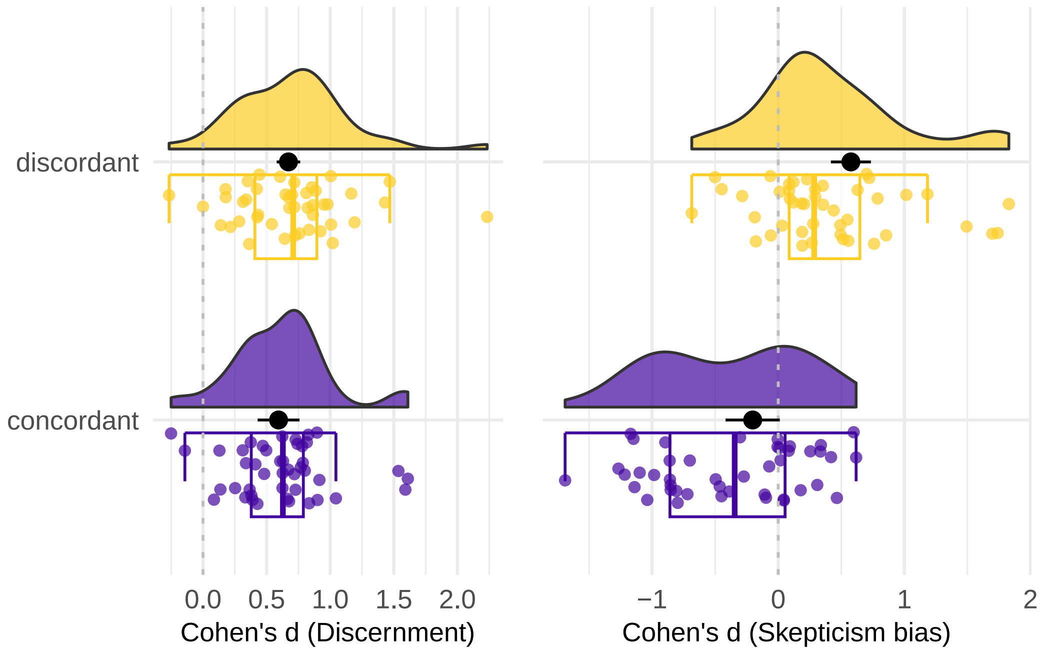

Figure 7.4: *Effect of political concordance on discernment and skepticism bias.* The figure shows the distribution of the n = `r descriptives$concordance$n_effect$value` effect sizes for politically concordant and discordant items. The black dots represent the predicted average of the meta-regression, the black horizontal bars the 95% confidence intervals. Note that the figure does not represent the different weights (i.e. the varying sample sizes) of the data points, but that these weights are taken into account in the meta-regression.

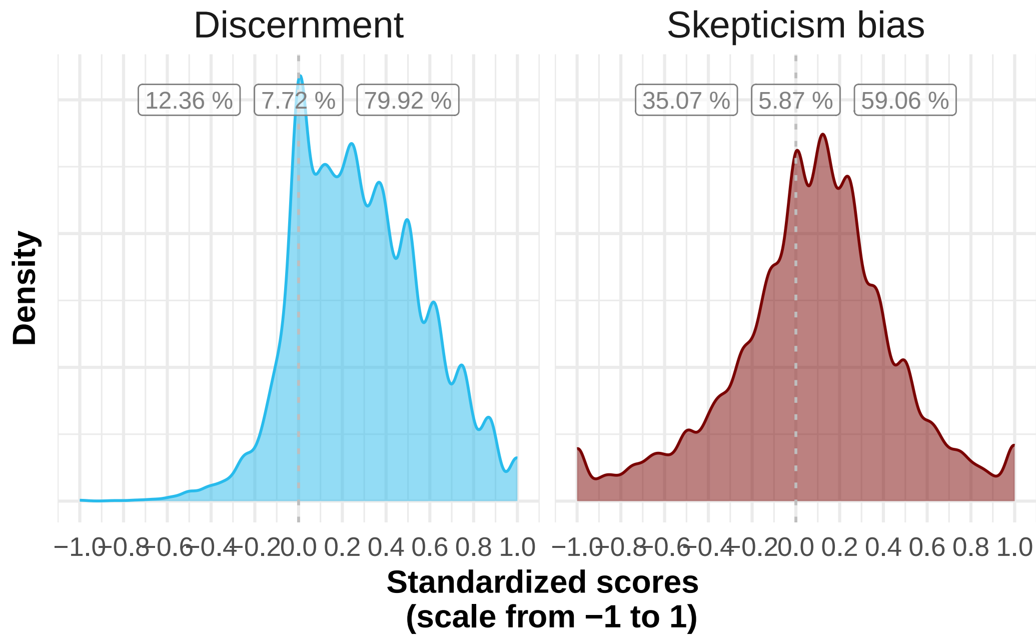

Figure 7.5: *Outcomes on the participant-level.* The figure shows the distribution of average discernment and skepticism bias scores of individual participants in the subset of studies that we have raw data on. We standardized original accuracy ratings to range from 0 to 1, to be able to compare across scales. Therefore, the worst possible score is -1 where, for discernment, an individual classified all news wrongly, and for skepticism bias, an individual classified all true news correctly (as true) and all false news incorrectly (as true). The best possible score is 1 where, for discernment, an individual classified all news correctly, and for skepticism bias, an individual classified all true news incorrectly (as false) and all false news correctly (as false). The percentage labels (from left to right) represent the share of participants with a negative score, a score of exactly 0, and a positive score, for both measures respectively.

7.3 Discussion

This meta-analysis sheds light on some of the most common fears voiced about false news. In particular, we investigated whether people are able to discern true from false news, and whether they are better at judging the veracity of true news or false news (skepticism bias). Across 303 effect sizes ($N_{participants}$ = 194438) from 40 countries across 6 continents, we found that people rated true news as much more accurate than fact-checked false news ($d_{discernment}$ = 1.12 [1.01, 1.22], z = 20.79, p < .001) and are slightly better at rating fact-checked false news as inaccurate than at rating true news as accurate ($d_{\text{skepticism bias}}$ = 0.32 [0.24, 0.39], z = 8.11, p < .001).

The finding that people can discern true from false news when prompted to do so has important implications for interventions against misinformation. First, it suggests that most people do not lack the skills to spot false news–at least the kind of fact-checked false news used in the studies included in our meta-analysis. If people don't lack the skills to spot false news, why do they sometimes fall for false news? In some contexts, people may lack the motivation to use their discernment skills or may only apply them selectively (Pennycook, Epstein, et al. 2021; Rathje et al. 2023). Thus, instead of teaching people how to spot false news, it may be more fruitful to target motivations, either by manipulating features of the environment in which people encounter news (Capraro and Celadin, n.d.; Globig, Holtz, and Sharot 2023), or by intrinsically motivating people to use their skills and pay more attention to accuracy (Pennycook, Epstein, et al. 2021). For instance, it has been shown that design features of current social media environments sometimes impede discernment (Epstein et al. 2023).

Second, the fact that people can, on average, discern true from false news lends support to crowdsourced fact-checking initiatives. While fact-checkers cannot keep up with the pace of false news production, the crowd can, and it has been shown that even small groups of participants perform as well as professional fact-checkers (Allen et al. 2021; Martel et al. 2022). The cross-cultural scope of our findings suggests that these initiatives may be fruitful in many countries across the world. In every country included in the meta-analysis, participants on average rated true news as more accurate than false news (see appendix). In line with past work (Allen et al. 2021), we have shown that this was not only true on average, but for a large majority (79.92 %) of participants for which we had individual level data. Our results are also informative for the work of fact-checkers. Since people appear to be quite good at discerning true from false news, fact-checkers may want to focus on headlines that are less clearly false or true. However, we cannot rule out that people's current discernment skills stem in part from the current and past work of fact-checking organizations.

The fact that people disbelieve true news slightly more than they believe fact-checked false news speaks to the nature of the misinformation problem and how to fight it: the problem may be less that people are gullible, and fall for falsehoods too easily, but instead that people are excessively skeptical, and do not believe reliable information enough (Altay, Berriche, and Acerbi, n.d.; Mercier 2020). Even assuming that the rejection of true news and the acceptance of false news are of similar magnitude (and that both can be improved), given that true news are much more prevalent in people's news diet than false news (Allen et al. 2020), true news skepticism may be more detrimental to the accuracy of people's beliefs than false news acceptance (Acerbi, Altay, and Mercier 2022). This skepticism is concerning in the context of the low and declining trust and interest in news across the world (Altay, Fletcher, and Nielsen 2024), as well as the attacks of populist leaders on the news media (Van Duyn and Collier 2019) and growing news avoidance (Newman et al. 2023). Interventions aimed at reducing misperceptions should therefore consider increasing the acceptance of true news in addition to reducing the acceptance of false news (Acerbi, Altay, and Mercier 2022; Altay, De Angelis, and Hoes, n.d.). At the very least, when testing interventions, researchers should evaluate their effect on both true and false news, not just false news (Guay et al., n.d.). At best, interventions should use methods that allow to estimate discrimination while accounting for response bias, such as Signal Detection Theory, and make sure that apparent increases in discernment are not due to a more conservative response bias (Higham, Modirrousta-Galian, and Seabrooke 2024; Modirrousta-Galian and Higham 2023). This is all the more important given that recent evidence suggests that many interventions against misinformation, such as media literacy tips (Hoes et al. 2023), fact-checking (Bachmann and Valenzuela 2023), or educational games aimed at inoculating people against misinformation (Modirrousta-Galian and Higham 2023), may reduce belief in false news at the expense of fostering skepticism towards true news.

We also investigated various moderators of discernment and skepticism bias. We found that discernment was greater in studies conducted in the United States compared to the rest of the world. This could be due to the inclusion of many countries from the Global South, where belief in misinformation and conspiracy theories has been documented to be higher (Alper, n.d.). In line with past work (Dias, Pennycook, and Rand 2020), the presence of a source had no statistically significant effects on discernment or skepticism bias. Neither did the topic of the news. Participants showed greater skepticism in studies that presented headlines in a social media format (with an image and lede) or along with an image compared to studies that used plain headlines. This suggests that the skepticism towards true news documented in this meta-analysis may be partially due to the social media format of the news headlines. Past work has shown that people report trusting news on social media less (Mont'Alverne et al. 2022; Newman et al. 2022), and experimental manipulations have shown that the Facebook news format reduces belief in news (Besalú and Pont-Sorribes 2021; Karlsen and Aalberg 2023)–although the causal effects documented in these experiments are much smaller than observational differences in reported trust levels between news on social media and on news outlets (Agadjanian et al. 2023). Low trust in news on social media may be a good thing, given that on average news on social media may be less accurate than news on news websites, but it is also worrying given that most of news consumption worldwide is shifting online and on social media in particular (Newman et al. 2023).

The political concordance of the news had no effect on discernment, but participants were excessively skeptical of politically discordant news. That is, participants were equally skilled at discerning true from false news for concordant and discordant items, but they rated news generally (true and false) as more false when politically discordant. This finding is in line with recent evidence on partisan biases in news judgments (Gawronski, Ng, and Luke 2023), and supports the idea that people are not excessively gullible of news they agree with, but are instead excessively skeptical of news they disagree with (Mercier 2020; Trouche et al. 2018). It suggests that interventions aimed at reducing partisan motivated reasoning, or at improving political reasoning in general, should focus more on increasing openness to opposing viewpoints than on increasing skepticism towards concordant viewpoints. Future studies should investigate whether the effect of congruence is specific to politics or if it holds across other topics, and compare it to a baseline of neutral items.

Our meta-analysis has two main conceptual limitations. First, participants evaluated the news stories in artificial settings that do not mimic the real-world. For instance, the mere fact of asking participants to rate the accuracy of the news stories may have increased discernment by increasing attention to accuracy (Pennycook, Epstein, et al. 2021). When browsing on social media, people may be less discerning (and perhaps less skeptical) than in experimental settings because they would pay less attention to accuracy (Epstein et al. 2023). However, given people's low exposure to misinformation online (Altay, Kleis Nielsen, and Fletcher 2022), people may mostly protect themselves from misinformation not by detecting misinformation on the spot, but by relying on the reputation of the sources and avoiding unreliable sources (Altay, Hacquin, and Mercier 2022). Second, our results reflect choices made by researchers about news selection. The vast majority of studies in our meta-analysis relied on fact-checked false news, determined by fact-checking websites (e.g. Snopes, PolitiFact). By contrast, three papers (Garrett and Bond 2021; Aslett et al. 2024; Allen et al. 2021) automated their news selection by scraping headlines from media outlets in real-time, and had both participants and fact-checkers (or the researchers themselves, in the case of Garrett and Bond (2021)) rating the veracity of the headlines shortly after. The three studies (53 effect sizes; 10170 participants; all in the United States) find (i) lower discernment than our meta-analytic average, and (ii) a negative skepticism (i.e. a credulity) bias (see appendix for a detailed discussion). This highlights the importance of news selection in misinformation research: Researchers need to think carefully about what population of news they sample from, and be clear about the generalizability of their findings (Pennycook, Binnendyk, et al. 2021; Altay, Berriche, and Acerbi, n.d.).

Our meta-analysis further has methodological limitations which we address in a series of robustness checks in the appendix. We show that our results hold across alternative effect size estimators. We also show that we obtain similar results when running a participant-level analysis on a subset of studies for which we have raw data and when relying on d' (sensitivity) and c (response bias) from Signal Detection Theory for that subset. A comparison of binary and Likert-scale ratings suggests that skepticism bias stems partly from mis-classifications, partly from degrees of confidence.

In conclusion, we found that in experimental settings, people are able to discern mainstream true news from fact-checked false news, but when they err, they tend to do so on the side of skepticism more than on the side of gullibility (although the effect is small and likely contingent on false news selection). These findings lend support to crowdsourced fact-checking initiatives, and suggest that, to improve discernment, there may be more room to increase the acceptance of true news than to reduce the acceptance of false news.

7.4 Methods

7.4.1 Data

We undertook a systematic review and meta-analysis of the experimental literature on accuracy judgments of news, following the PRISMA guidelines (Page et al. 2021). All records resulting from our literature searches can be found on the OSF project page (https://osf.io/96zbp/). We documented rejection decisions for all retrieved papers. They, too, can be found on the OSF project page.

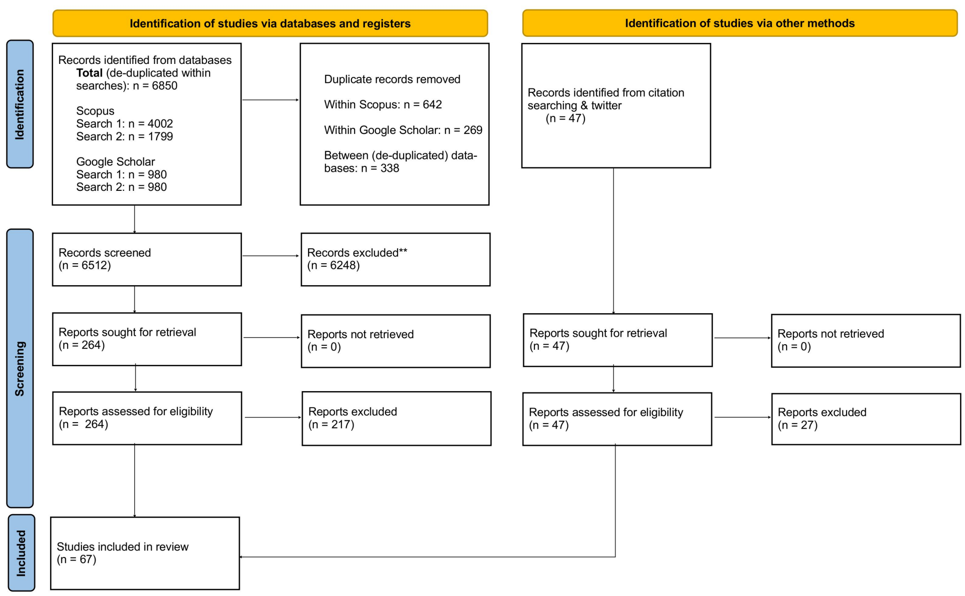

Figure 7.6: *PRISMA flow diagram.* A flow diagram for the systematic literature review, based on the 2020 PRISMA template.

7.4.1.1 Eligibility criteria

For a publication to be included in our meta-analysis, we set six eligibility criteria: (1) We considered as relevant all document types with original data (not only published ones, but also reports, pre-prints and working papers). When different publications were using the same data, a scenario we encountered several times, we included only one publication (which we picked arbitrarily). (2) We only included articles that measured perceived accuracy (including "accuracy", "credibility", "trustworthiness", "reliability" or "manipulativeness"), and (3) did so for both true and false news. (4) We only included studies relying on real-world news items. Accordingly, we excluded studies in which researchers made up the false news items, or manipulated the properties of the true news items. (5) We could only include articles that provided us with the relevant summary statistics (means and standard deviations for both false and true news), or publicly available data that allowed us to calculate those. In cases where we were not able to retrieve the relevant summary statistics either way, we contacted the authors. (6) Finally, to ensure comparability, we only included studies that provided a neutral control condition. For example, Calvillo and Smelter (2020), among other things, test the effect of an interest prime vs. an accuracy prime. A neutral control condition–one that is comparable to those of other studies–would have been no prime at all. We therefore excluded the paper. Rejection decisions for all retrieved papers are documented and can be accessed on the OSF project page (https://osf.io/96zbp/). We provide a list of all included articles in the appendix.

7.4.1.2 Deviations from eligibility criteria

We followed our eligibility criteria with 4 exceptions. We rejected one paper based on a criterion that we had not previously set: scale asymmetry. Baptista et al. (2021) asked participants: "According to your knowledge, how do you rate the following headline?", providing a very asymmetrical set of answer options ("1—not credible; 2—somehow credible; 3—quite credible; 4—credible; 5—very credible"). The paper provides 6 effect sizes, all of which strongly favor our second hypothesis (one effect being as large as d = 2.54). We decided to exclude this paper from our analysis because of its very asymmetric scale (no clear scale midpoint, and labels not symmetrically mapping onto a false/true dichotomy, by contrast to all other response scales included here). Further, we stretched our criterion for real-world news on three instances. Maertens et al. (2021) and Roozenbeek et al. (2020) used artificial intelligence trained on real-world news to generate false news. Bryanov et al. (2023) had journalists create the false news items. We reasoned that asking journalists to write news should be similar enough to real-wolrd news, and that LLMs already produce news headlines that are indistinguishable from real news, so it should not make a big difference.

7.4.1.3 Literature search

Our literature review is based on two systematic searches. We conducted our first search on March 2, 2023 using Scopus (search string: ' "false news" OR "fake news" OR "false stor*" AND "accuracy" OR "discernment" OR "credibilit*" OR "belief" OR "susceptib*" ') and google scholar (search string: ' "Fake news" | "False news"|"False stor*" "Accuracy" | "Discernment"|"Credibility"|"Belief"|"Suceptib*", no citations, no patents'). On Scopus, given the initially high volume of papers (12425), we excluded papers not written in English, that were not articles or conference papers, and that were from disciplines that are likely irrelevant for the present search (e.g., Dentistry, Veterinary, Chemical Engineering, Chemistry, Nursing, Pharmacology, Microbiology, Materials Science, Medicine) or unlikely to use an experimental design (e.g. Computer Science, Engineering, Mathematics, see appendix for detailed search string). After these filters were applied, we ended up with 4002 results. The Google Scholar search was intended to identify important pre-prints or working papers that the Scopus search would have missed. We only considered the first 980 results of that search–a limit imposed by the "Publish or Perish" software we used to store Google Scholar search results in a data frame.

After submitting a manuscript version, reviewers remarked that not including the terms "misinformation" or "disinformation" in our search string might have omitted relevant results. On March 22nd, 2024, we therefor conducted a second, pre-registered (https://doi.org/10.17605/OSF.IO/YN6R2, registered on March 12, 2024) search using an extended query string (search string for both Scopus and Google Scholar: ' "false news" OR "fake news" OR "false stor*" OR "misinformation" OR "disinformation") AND ("accuracy" OR "discernment" OR "credibilit*" OR "belief" OR "suceptib*" OR "reliab*" OR "vulnerabi*" '; see appendix for detailed search string). After removing duplicates–642 between the first and the second Scopus search and 269 between the first and the second Google Scholar search–the second search yielded an additional 1157 results for Scopus and 711 results for Google Scholar. In total, the Scopus searches yielded 5159, the Google Scholar searches 1691 unique results.

We identified and removed 338 duplicates between the Google Scholar and the Scopus searches and ended up with 6512 documents for screening. We had two screening phases: first titles, second abstracts. For the results from the second literature search, both authors screened the results independently. In case of conflicting decisions, an article passed onto the next stage (i.e. received abstract screening or full text assessment). For the results from the second literature search, screening was done based on titles and abstracts only, so that the screeners would not be influenced by information on the authors or the publishing journal. The vast majority of documents (6248) had irrelevant titles and were removed during that phase. Most irrelevant titles were not about false news or misinformation (e.g. "Formation of a tourist destination image: Co-occurrence analysis of destination promotion videos"), and some were about false news or misinformation but were not about belief or accuracy (e.g. "Freedom of Expression and Misinformation Laws During the COVID-19 Pandemic and the European Court of Human Rights"). We stored the remaining 264 records in the reference management system Zotero for retrieval. Of those, we rejected a total of 217 papers that did not meet our inclusion criteria. We rejected 87 papers based on their abstract and 130 after assessment of the full text. We documented all rejection decisions, available on the OSF project page (https://osf.io/96zbp/). We included the remaining 47 papers from the systematic literature search. To complement the systematic search results, we conducted forward and backward citation search through Google Scholar. We also reviewed additional studies that we had on our computers and papers we found scrolling through twitter (mostly unpublished manuscripts). Taken together, we identified an additional 47 papers via those methods. Of these, we excluded 27 papers after full text assessment because they did not meet our inclusion criteria. For these papers, too, we documented our exclusion decisions. They can be found together with the ones of the systematic search on the OSF project page (https://osf.io/96zbp/). We included the remaining 20 papers. In total, we included 67 papers in our meta analysis, 47 of which were peer-reviewed and 20 grey literature (reports and working papers). We retrieved the relevant summary statistics directly from the paper for 21 papers, calculated them ourselves based on publicly available raw data for 31 papers, and got them from the authors after request for 15 papers.

7.4.2 Statistical methods

Unless explicitly stated otherwise, we pre-registered (https://doi.org/10.17605/OSF.IO/SVC7U, registered on April 28, 2023) all reported analyses. Our choice of statistical models was informed by simulations, which can also be found on the OSF project page. We conducted all analyses in R version 4.4.1 (2024-06-14) (R Core Team 2022) using Rstudio version 2024.9.0.375 (Posit team 2023) and the `tidyverse` package version 2.0.0 (Wickham et al. 2019). For effect size calculations, we rely on the `escalc()`, for models on the `rma.mv()`, for clustered standard errors on the `robust()` function, all from the `metafor` package version 4.6.0 (Viechtbauer 2010).

7.4.2.1 Deviations from pre-registration

We pre-registered standardized mean changes using change score standardization (SMCC) as an estimator for our effect sizes (Gibbons, Hedeker, and Davis 1993). However, in line with Cochrane guidelines (Higgins et al. 2019), we chose to rely on the more common Cohen's d for the main analysis. We report results from the pre-registered SMCC (along with other alternative estimators) in the appendix. All estimators yield similar results. We did not pre-register considering scale symmetry, proportion of true news and false news selection (taken from fact checking sites vs. verified by researchers) as moderator variables. We report the results regarding these variables in the appendix.

7.4.2.2 Outcomes

We have two complementary measures of assessing the quality of people's news judgment. The first measure is discernment. It measures the overall quality of news judgment across true and false news. We calculate discernment by subtracting the mean accuracy ratings of false news from the mean accuracy ratings of true news, such that more positive scores indicate better discernment. However, discernment is a limited diagnostic of the quality of people's news judgment. Imagine a study A in which participants rate 50% of true news and 20% of false news as accurate, and a study B finding 80% of true news and 50% of false news rated as accurate. In both cases, the discernment is the same: Participants rated true news as more accurate by 30 percentage points than false news. However, the performance by news type is very different. In study A, people do well for false news–they only mistakenly classify 20% as accurate–but are at chance for true news. In study B, it's the opposite. We therefore use a second measure: skepticism bias. For any given level of discernment, it indicates whether people's judgments were better on true news or on false news, and to what extent. First, we calculate an error for false and true news separately, which we define as the distance of participants' actual ratings to the best possible ratings. For example, for study A, the mean error for true news is 50% (100%-50%), because in the best possible scenario, participants would have classified 100% of true news as true. The error for false news in Study A is 20% (20%-0%), because the best possible performance for participants would have been to classify 0% of false news as accurate. We calculate skepticism bias by subtracting the mean error for false news from the mean error for true news. For example, for Study A, the skepticism bias is 30% (50%-20%). A positive skepticism bias indicates that people doubt true news more than they believe false news.

Skepticism bias can only be (meaningfully) interpreted on scales using symmetrical labels, i.e. the intensity of the labels to qualify true and false news are equivalent (e.g., "True" vs "False" or "Definitely fake" [1] to "Definitely real" [7]). 69% of effects included in the meta-analysis used scales with perfectly symmetrical labels, while 26% used imperfectly symmetrical scale labels, i.e., the intensity of the labels to qualify true and false news are similar but not equivalent (e.g., [1] not at all accurate, [2] not very accurate, [3] somewhat accurate, [4] very accurate; here for instance 'not all accurate' is stronger than 'very accurate'). We could only compute this variable for scales that explicitly labeled scale points, resulting in missing values for 5% of effects. In the appendix, we show that scale symmetry has no statistically significant effect on skepticism bias.

7.4.2.3 Effect sizes

The studies in our meta analysis used a variety of response scales, including both binary (e.g. "Do you think the above headline is accurate? - Yes, No") and continuous ones (e.g. "To the best of your knowledge, how accurate is the claim in the above headline" 1 = Not at all accurate, 4 = Very accurate). To be able to compare across the different scales, we calculated standardized effects, i.e. effects expressed in units of standard deviations. Precisely, we calculated Cohen's d as

$$\text{Cohen's d} = \frac{\bar{x}_{\text{true}} - \bar{x}_{\text{false}}}{SD_{\text{pooled}}}$$

with

$$SD_{\text{pooled}} = \sqrt{\frac{SD_{\text{true}}^2 + SD_{\text{false}}^2}{2}}$$

The vast majority of experiments (294 out of 303 effects) in our meta analysis manipulated news veracity within participants, i.e. having participants rate both false and true news. Following the Cochrane manual, we account for the dependency between ratings that this design generates when calculating the standard error for Cohen's d. Precisely, we calculate the standard error for within participant designs as

$$SE_{\text{Cohen's d (within)}} = \sqrt{\frac{2(1 - r_{\text{true,false}})}{n} + \frac{\text{Cohen's d}^2}{2n}}$$

where r is the correlation between true and false news. Ideally, for each effect size (i.e. the meta-analytic units of observation) in our data, we need the estimate of r. However, this correlation is generally not reported in the original papers. We could only obtain it for a subset of samples for which we collected the summary statistics ourselves, based on the raw data. Based on this subset of correlations, we calculated an average correlation, which we then imputed for all effect size calculations. This approach is in line with the Cochrane recommendations for crossover trials (Higgins et al. 2019). In our case, this average correlation is 0.26.

For the 9 (out of 303) effects from studies that used a between participant design, we calculated the standard error as

$$SE_{\text{Cohen's d (between)}} = \sqrt{\frac{n_{\text{true}} + n_{\text{false}}}{n_{\text{true}} n_{\text{false}}} + \frac{\text{Cohen's d}^2}{2(n_{\text{true}} + n_{\text{false}})}}$$

For all effect size calculations, we defined the sample size n as the number of instances of news ratings. That is, we multiplied the number of participants with the number of news items rated per participant.

7.4.2.4 Models

In our models for the meta analysis, each effect size was weighted by the inverse of its standard error, thereby giving more weight to studies with larger sample sizes. We used random effects models, which assume that there is not only one true effect size but a distribution of true effect sizes (Harrer et al. 2021). These models assume that variation in effect sizes is not only due to sampling error alone, and thereby allow to model other sources of variance. We estimated the overall effect of our outcome variables using a three-level meta-analytic model with random effects on the sample and the publication level. This approach allowed us to account for the hierarchical structure of our data, in which samples (level three) contribute multiple effects (level two), (level one being the participant level of the original studies, see Harrer et al. (2021)). A common case where a sample provides several effect sizes occurs when participants rated both politically concordant and discordant news. In this case, if possible, we entered summary statistics separately for the concordant and discordant items, yielding two effect sizes (i.e. two different rows in our data frame). Another case where multiple effects per sample occurred was when follow-up studies were conducted on the same participants (but different news items). While our multi-level models account for this hierarchical structure of the data, they do not account for dependencies in sampling error. When one same sample contributes several effect sizes, one should expect their respective sampling errors to be correlated (Harrer et al. 2021). To account for dependency in sampling errors, we computed cluster-robust standard errors, confidence intervals, and statistical tests for all meta-analytic estimates.

To assess the effect of moderator variables, we calculated meta regressions. We calculated a separate regression for each moderator, by adding the moderator variable as a fixed effect to the multilevel meta models presented above. We pre-registered a list of six moderator variables to test. Those included the *country* of studies (levels: United States vs. all other countries), *political concordance* (levels: politically concordant vs. politically discordant), *news family* (levels: political, including both concordant and discordant vs. covid related vs. other, including categories as diverse as history, environment, health, science and military related news items), the *format* in which the news were presented (levels: headline only vs. headline and picture vs. headline, picture and lede), whether news items were accompanied by a *source* or not, and the *response scale* used (levels: 4-point vs. binary vs. 6-point vs. 7-point vs. other, for all other numeric scales that were not frequent). We ran an additional regression for two non-preregistered variables, namely the *symmetry of scales* (levels: perfectly symmetrical vs. imperfectly symmetrical) and *false news selection* (levels: taken from fact check sites vs. verified by researchers). We further descriptively checked whether the *proportion of true news* among all news would yield differences.

7.4.2.5 Publication bias

We ran some standard procedures for detecting publication bias. However, a priori we did not expect publication bias to be present because our variables of interest were not those of interest to the researchers of the original studies: Researchers generally set out to test factors that alter discernment, and not the state of discernment in the control group. No study measured skepticism bias in the way we define it here.

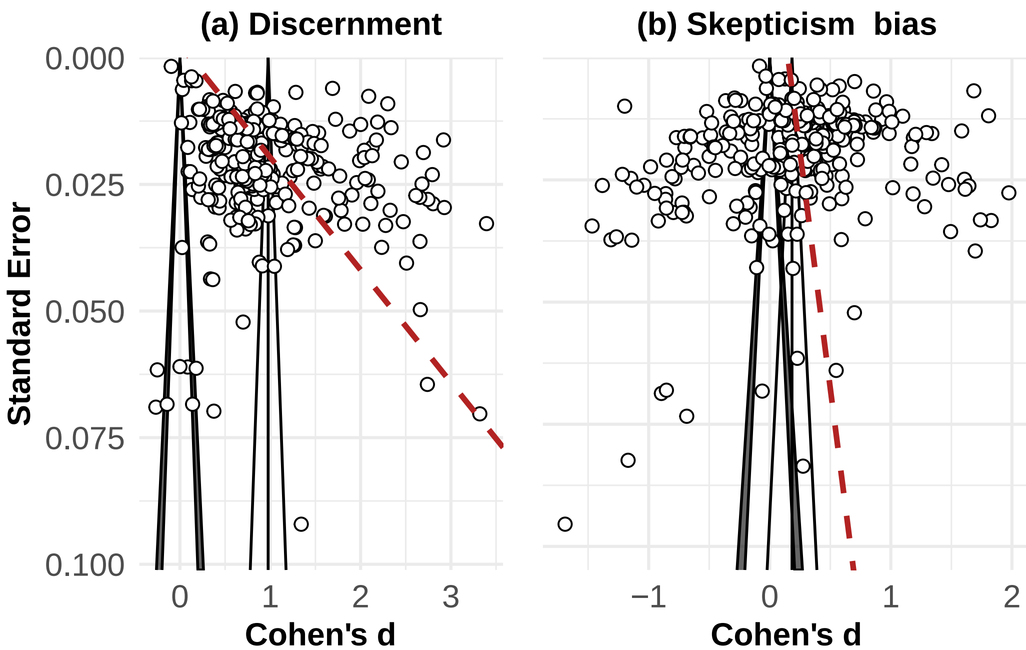

Figure 7.7: *Funnel plots for discernment and skepticism bias.* Dots represent effect sizes. In the absence of publication bias and heterogeneity, one would then expect to see the points forming a funnel shape, with the majority of the points falling inside of the pseudo-confidence region centered around the average effect estimate, with bounds of ±1.96 SE (the standard error value from the y-axis). The dashed red regression line illustrates the estimate of the Egger's regression test. For both outcomes, the slope differs significantly from zero, see Appendix.

Regarding discernment, we find evidence that smaller studies tend to report larger effect sizes, according to Egger's regression test (see Figure 7.7); see also the appendix). We do not find evidence for asymmetry regarding skepticism bias. However, it is unclear how meaningful these results are. As illustrated by the funnel plot, there is generally high between-effect size heterogeneity: Even when focusing only on the most precise effect sizes (top of the funnel), the estimates vary substantially. It thus seems reasonable to assume that most of the dispersion of effect sizes does not arise from studies' sampling error, but from studies estimating different true effects. Further, even the small studies are relatively high powered, suggesting that they would have yielded significant, publishable results even with smaller effect sizes. Lastly, Egger's regression test can lead to an inflation of false positive results when applied to standardized mean differences (Pustejovsky 2019; Harrer et al. 2021).

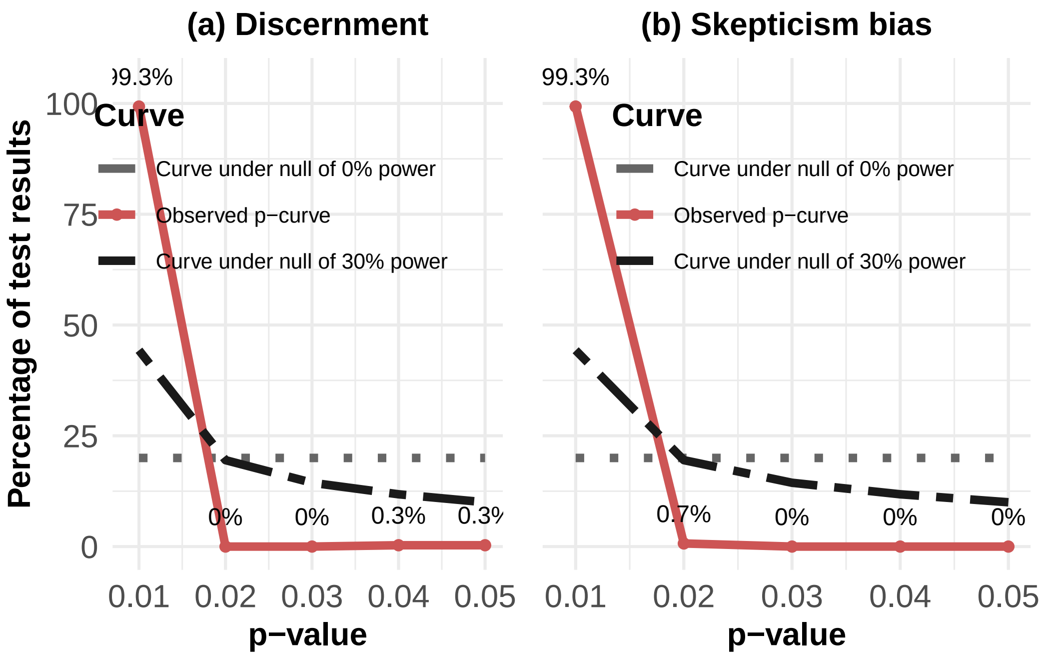

Figure 7.8: *P-curves for discernment and skepticism bias.* The p-curve shows the percentage of effect sizes for a given p value within the range of 0.1 and 0.5. All values smaller than 0.01 are rounded to that value. The reference lines indicate the expected percentage of studies for a given p value, assuming that there is a true effect and certain statistical power to detect it (either 0% or 30% power). The observed p-curve is negatively sloped and heavily right skewed (the tail points to the right) for both outcomes, which suggests no widespread p-hacking.

We do not find any evidence to suspect p-hacking for either discernment or skepticism bias from visually inspecting p-curves for both outcomes (see Figure 7.8).

7.5 Data availability

The extracted data used to produce our results are available on the OSF project page (https://osf.io/96zbp/).

7.6 Code availability

The code used to create all results (including tables and figures) of this manuscript is also available on the OSF project page (https://osf.io/96zbp/).

7.7 Acknowledgements

The authors thank Aurélien Allard, Hugo Mercier, Gordon Pennycook, Ariana Modirrousta-Galian and Ben Tappin for their valuable feedback on earlier versions of the manuscript. JP received funding from the SCALUP ANR grant ANR-21-CE28-0016-01. SA received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement nr. 883121). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

7.8 Author Contributions Statement

JP: Conceptualization, Systematic literature search, Methodology, Software, Formal Analysis, Data curation, Visualization, Writing - Original draft, Writing - Review & Editing. SA: Conceptualization, Systematic literature search, Writing - Original draft, Writing - Review & Editing.

7.9 Competing interest

The authors declare having no competing interests.

8 Conclusion

It has long been a puzzle to the deficit model–which suggests that trust in science is primarily driven by science knowledge–that knowledge of science is at best weakly associated with science attitudes (Allum et al. 2008; National Academies of Sciences, Engineering, and Medicine 2016). The rational impression account can make sense of this: it lays out how trusting science without recalling specific knowledge can be the result of a sound inference process, rooted in basic cognitive mechanisms of information evaluation.

The account is compatible with the finding that education, and in particular science education, has been repeatedly identified as one of the strongest correlates of trust in science (Bak 2001; Noy and O'Brien 2019; Wellcome Global Monitor 2018, 2020; but see Cologna et al. 2025). By contrast with the deficit model, it suggests that the main causal role of education for public trust in science is not transmission of knowledge and understanding, but impression generation.

The rational impression account aligns with recent normative accounts which shift the focus from listing particular key institutional features that make science trustworthy-certain methods, norms, or processes–to the diversity of science: Cartwright et al. (2022) make the case that scientific knowledge emerges from a tangle of results, relying on diverse research methods. Oreskes (2019) makes a similar case: She argues that scientific practice takes place in different scientific communities who rely on a variety of different research methods. Through some shared practices, in particular peer-review, these communities engage in critical dialogue. What makes scientific knowledge trustworthy, according to Oreskes, is when from this diversity of actors and methods, a consensus emerges. According to this view, to infer trustworthiness, people should have a representation of the diversity of science. The rational impression account is, in a way, less strict: it does not require a representation of diversity. However, for inferences from convergence to be sound, people do need to have a representation of science as an institution of independent thinkers.

The studies presented in this thesis face several limitations which are detailed in the individual chapters. Here, I will point out some more general, theoretical limitations of the rational impression account.

The rational impression account fits with a sociological literature investigating how "individual cognition and practice establish and maintain institutional fields and status hierarchies, especially in the face of imperfect knowledge" (Gauchat and Andrews 2018, 569). It proposes a possible micro-level model of trust in science and should be seen as complementing, not competing with, macro-level processes that shape public trust in science. Sociological macro-level accounts have made the case that trust in science is entangled with broader cultural and political dynamics. These accounts, like the individual-level accounts reviewed above, tend to focus on explaining distrust in science. For example, Gauchat (2011) describes the 'alienation model', according to which the "public disassociation with science is a symptom of a general disenchantment with late modernity, mainly, the limitations associated with codified expertise, rational bureaucracy, and institutional authority" (Gauchat 2011, 2). This explanation builds on the work of social theorists (Habermas 1989; Beck 1992; Giddens 1991; see Gauchat 2011 for an overview) who suggested that a modern, complex world increasingly requires expertise, and thus shapes institutions of knowledge elites. People who are not part of these institutions experience a lack of agency, resulting in a feeling of alienation. Similarly, Gauchat (2023) argues that politicization of science in the US needs to be seen in its broader cultural context. Precisely, according to Gauchat, science has enabled the authority of the modern regulatory state. Consequently, conservative distrust of science reflects deeper structural tensions with the institutions and rational–legal authority of modern governance. At the micro-level, this is consistent with research showing that right-wing authoritarian ideology is associated with distrust towards science and scientists (Kerr and Wilson 2021).

Another limitation of the rational impression account is that it assumes people have a representation of science as consensual. However, in practice–with perhaps some exceptions, such as during the Covid-19 pandemic–most people do not literally compare the opinions of different scientists for themselves and come to the conclusion that something is largely consensual. Where, then, could the representation of consensus possibly emerge? A plausible explanation, I believe, is that education fosters a representation of consensus: During education, in particular during early education, knowledge is typically presented as simply the result of science–a seemingly unanimous enterprise that produces knowledge. School books hardly teach about historical science controversies, suggest uncertainty around scientific findings, or cover cutting-edge research where disagreements are the norm. This could induce a default consensus assumption in people's perceptions of science. However, this argument is of course only speculative.

The rational impression account is also limited in its implications. First, I do not believe that flooding people with impressive consensual science knowledge is the key to overcoming all distrust in science. In the context of trust in political institutions, research has shown that trust and distrust are not necessarily symmetrical: what causes the former might not help alleviate the latter (Bertsou 2019). I believe this is at least to some degree true for science, too. Especially in a context of the global north, where essentially everyone has been exposed to science through a basic science education, trust in science via the mechanisms of the rational impression account is likely to be the default state. This is in line with our findings on quasi-universal trust in basic science in the US (Chapter 6). Consensus messaging has been shown to help convince people to trust science on particular issues, such as climate change or vaccines, but it is less clear whether consensus messaging could also enhance perceptions of trustworthiness. It might be the case that the people convinced by consensus messages might have already generally trusted science, but have not held strong opinions on the specific matter, as has been argued to be the case for a large segment of the public on most matters (Bourdieu 1979; Zaller 1992). For people who do not only lack trust, but who actively distrust, motivated reasoning accounts are likely better suited as a theoretical framework. Addressing relevant underlying motivations directly might be more fruitful to mitigate distrust in science than exposing people to consensual science more generally.

Second, and related, just because I propose an account by which trust in science can be rational, this does not mean that, conversely, all distrust in science is irrational. Some groups of people do in fact have good reasons not to trust science. For example, some science has historically contributed to fostering racism (see e.g. Fuentes 2023; Nobles et al. 2022), via instances such as the tragically famous Tuskegee syphilis study (Brandt 1978; Scharff et al. 2010).

Third, I do not think that science communication should stress consensus at all costs. In the rational impression account, consensus plays a central role for generating trust. However, this should not incentivize science communicators to neglect transparency about uncertainty. Acknowledging uncertainty in science communication has been argued to be crucial for fostering long term trust in science (Druckman 2015). For example, in the context of Covid-19 vaccines, Petersen et al. (2021) have shown that communicating uncertainty is crucial for building long term trust in health authorities.

Fourth, science communication should not aim for impressiveness at all costs either. Research has shown that intellectual humility can increase trust in scientists (Koetke et al. 2024). Trying to oversell scientific results might therefore backfire. People appear to value transparency via open data practices in science (Song, Markowitz, and Taylor 2022), and trust science that replicates more (Hendriks, Kienhues, and Bromme 2020). I would therefore expect that simply doing better, more transparent science and being humble about it is likely to be the most effective strategy to impress the public and elicit perceptions of trustworthiness.

Fifth, educators should not stop aiming at fostering a proper understanding of science. Most students might not understand all of the content, or recall much specific knowledge later on. However, for some students at least, some of that knowledge will be remembered, and will prove important in their lives. Second, to be impressive, a piece of information does not need to be confusingly complex. In fact, a proper understanding of research findings and their methods might even help in appreciating their complexity–even if, once again, that understanding is forgotten later.

Despite these caveats, I believe that the rational impressions account offers optimism for studies of science-society interfaces, and the field of science communication in particular: Exposure to science, especially one that leaves an impression, might be the foundation of public trust in science. Low scientific literacy levels should not discourage education and communication efforts, as they are not necessarily a good indicator of the value added in terms of fostering trust in science.

Taking a broader perspective, our account fits into a picture of humans as not gullible (Mercier 2017, 2020). The "failure" of the deficit model, i.e. the fact that science knowledge appears to not be strongly associated with trust in science, might suggest that public trust in science is, to a large extent, irrational. The notion that trust in science is irrational or easily granted may amplify concerns about the impact of misinformation: if trust lacks a solid, rational foundation, then we would expect misinformation to easily lead people astray. There is much work to be done still to understand how misinformation impacts people's beliefs, and in particular elite-driven misinformation and more subtle forms of misinformation, such as one-sided reporting. But in Chapter 7, we show that people are generally able to distinguish between true and false news and, if anything, tend to be generally skeptical of news. As a consequence, for a better informed public, fighting for (true) information seems at least as relevant as fighting against misinformation. Misinformation researchers increasingly acknowledge this: A recent report on science misinformation by the National Science Foundation (National Academies of Sciences 2024) dedicates considerable space on developing strategies to produce better information, for example by promoting high-quality science, health, and medical journalism.

The rational impression account stresses the role of fighting for information, when it comes to fostering trust in science. Well-placed trust in science does not require profound understanding or recall of specific knowledge; but it does require exposure to good science.

References

> **Note**
>
> References preceeded by a '*' are studies included in the meta-analysis

Acerbi, Alberto, Sacha Altay, and Hugo Mercier. 2022. "Research Note: Fighting Misinformation or Fighting for Information?" *Harvard Kennedy School Misinformation Review*, January. https://doi.org/10.37016/mr-2020-87.

Achterberg, Peter, Willem De Koster, and Jeroen Van Der Waal. 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): 704–20. https://doi.org/10.1177/0963662515617367.

Agadjanian, Alexander, Jacob Cruger, Sydney House, Annie Huang, Noah Kanter, Celeste Kearney, Junghye Kim, Isabelle Leonaitis, Sarah Petroni, and Leonardo Placeres. 2023. "A Platform Penalty for News? How Social Media Context Can Alter Information Credibility Online." *Journal of Information Technology & Politics* 20 (3): 338–48.

Algan, Yann, Daniel Cohen, Eva Davoine, Martial Foucault, and Stefanie Stantcheva. 2021. "Trust in Scientists in Times of Pandemic: Panel Evidence from 12 Countries." *Proceedings of the National Academy of Sciences* 118 (40): e2108576118. https://doi.org/10.1073/pnas.2108576118.

*Ali, Khudejah, Cong Li, Khawaja Zain-ul-abdin, and Muhammad Adeel Zaffar. 2022. "Fake News on Facebook: Examining the Impact of Heuristic Cues on Perceived Credibility and Sharing Intention." *Internet Research* 32 (1): 379–97.

Allen, Jennifer, Antonio A. Arechar, Gordon Pennycook, and David G. Rand. 2021. "Scaling up Fact-Checking Using the Wisdom of Crowds." *Science Advances* 7 (36): eabf4393.

Allen, Jennifer, Baird Howland, Markus Mobius, David Rothschild, and Duncan J. Watts. 2020. "Evaluating the Fake News Problem at the Scale of the Information Ecosystem." *Science Advances* 6 (14): eaay3539.

Allum, Nick, Patrick Sturgis, Dimitra Tabourazi, and Ian Brunton-Smith. 2008. "Science Knowledge and Attitudes Across Cultures: A Meta-Analysis." *Public Understanding of Science* 17 (1): 35–54. https://doi.org/10.1177/0963662506070159.

Alper, Sinan. n.d. "When Conspiracy Theories Make Sense: The Role of Social Inclusiveness."

*Altay, Sacha, Andrea De Angelis, and Emma Hoes. 2024. "Media Literacy Tips Promoting Reliable News Improve Discernment and Enhance Trust in Traditional Media." *Communications Psychology* 2 (1): 1–9. https://doi.org/10.1038/s44271-024-00121-5.

*Altay, Sacha, and Fabrizio Gilardi. n.d. "People Are Skeptical of Headlines Labeled as AI-Generated, Even If True or Human-Made, Because They Assume Full AI Automation." https://doi.org/10.31234/osf.io/83k9r.

*Altay, Sacha, Benjamin A. Lyons, and Ariana Modirrousta-Galian. 2024. "Exposure to Higher Rates of False News Erodes Media Trust and Fuels Overconfidence." *Mass Communication and Society*, August, 1–25. https://doi.org/10.1080/15205436.2024.2382776.

*Altay, Sacha, Rasmus Kleis Nielsen, and Richard Fletcher. 2022. "The Impact of News Media and Digital Platform Use on Awareness of and Belief in COVID-19 Misinformation." https://doi.org/10.31234/osf.io/7tm3s.

Altay, Sacha, Emma de Araujo, and Hugo Mercier. 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): 373–94. https://doi.org/10.1080/21670811.2021.1941163.

Altay, Sacha, Manon Berriche, and Alberto Acerbi. n.d. "Misinformation on Misinformation: Conceptual and Methodological Challenges." *Social Media*.

Altay, Sacha, Andrea De Angelis, and Emma Hoes. n.d. "Beyond Skepticism: Framing Media Literacy Tips to Promote Reliable Informatio." https://doi.org/10.31234/osf.io/5gckb.

Altay, Sacha, Richard Fletcher, and Rasmus Kleis Nielsen. 2024. "News Participation Is Declining: Evidence from 46 Countries Between 2015 and 2022." *New Media & Society*, May, 14614448241247822. https://doi.org/10.1177/14614448241247822.

Altay, Sacha, Anne-Sophie Hacquin, and Hugo Mercier. 2022. "Why Do so Few People Share Fake News? It Hurts Their Reputation." *New Media & Society* 24 (6): 1303–24. https://doi.org/10.1177/1461444820969893.

Altay, Sacha, Rasmus Kleis Nielsen, and Richard Fletcher. 2022. "Quantifying the "Infodemic": People Turned to Trustworthy News Outlets During the 2020 Coronavirus Pandemic." *Journal of Quantitative Description: Digital Media* 2 (August). https://doi.org/10.51685/jqd.2022.020.

Altay, Sacha, Benjamin Lyons, and Ariana Modirrousta-Galian. n.d. "Exposure to Higher Rates of False News Erodes Media Trust and Fuels Skepticism in News Judgment." https://doi.org/10.31234/osf.io/t9r43.

Altay, Sacha, Kenzo Nera, Waqas Ejaz, Céline Schöpfer, and Frédéric Tomas. 2023. "Conspiracy Believers Claim to Be Free Thinkers but (Under)Use Advice Like Everyone Else." *British Journal of Social Psychology* 62 (4): 1782–97. https://doi.org/10.1111/bjso.12655.

Altenmüller, Marlene S., Tobias Wingen, and Anna Schulte. 2024. "Explaining Polarized Trust in Scientists: A Political Stereotype-Approach." *Science Communication* 46 (1): 92–115. https://doi.org/10.1177/10755470231221770.

*Arechar, Antonio A., Jennifer Allen, Adam J. Berinsky, Rocky Cole, Ziv Epstein, Kiran Garimella, Andrew Gully, et al. 2023. "Understanding and Combatting Misinformation Across 16 Countries on Six Continents." *Nature Human Behaviour* 7 (9): 1502–13. https://doi.org/10.1038/s41562-023-01641-6.

*Arin, K. Peren, Deni Mazrekaj, and Marcel Thum. 2023. "Ability of Detecting and Willingness to Share Fake News." *Scientific Reports* 13 (1): 7298. https://doi.org/10.1038/s41598-023-34402-6.

Asch, Solomon E. 1956. "Studies of Independence and Conformity: I. A Minority of One

Against a Unanimous Majority." *Psychological Monographs: General and Applied* 70 (9): 1–70. https://doi.org/10.1037/h0093718.

Aslett, Kevin, Zeve Sanderson, William Godel, Nathaniel Persily, Jonathan Nagler, and Joshua A. Tucker. 2024. "Online Searches to Evaluate Misinformation Can Increase Its Perceived Veracity." *Nature* 625 (7995): 548–56. https://doi.org/10.1038/s41586-023-06883-y.

Bachmann, Ingrid, and Sebastián Valenzuela. 2023. "Studying the Downstream Effects of Fact-Checking on Social Media: Experiments on Correction Formats, Belief Accuracy, and Media Trust." *Social Media + Society* 9 (2): 20563051231179694. https://doi.org/10.1177/20563051231179694.

*Badrinathan, Sumitra. 2021. "Educative Interventions to Combat Misinformation: Evidence from a Field Experiment in India." *American Political Science Review* 115 (4): 1325–41. https://doi.org/10.1017/S0003055421000459.

*Bago, Bence, David G. Rand, and Gordon Pennycook. 2020. "Fake News, Fast and Slow: Deliberation Reduces Belief in False (but Not True) News Headlines." *Journal of Experimental Psychology: General* 149 (8): 1608–13. https://doi.org/10.1037/xge0000729.

*Bago, Bence, Leah R. Rosenzweig, Adam J. Berinsky, and David G. Rand. 2022. "Emotion May Predict Susceptibility to Fake News but Emotion Regulation Does Not Seem to Help." *Cognition and Emotion*, June, 1–15. https://doi.org/10.1080/02699931.2022.2090318.

Bak, Hee-Je. 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): 779–95. https://doi.org/10.1111/0038-4941.00059.

Baptista, João Pedro, Elisete Correia, Anabela Gradim, and Valeriano Piñeiro-Naval. 2021. "The Influence of Political Ideology on Fake News Belief: The Portuguese Case." *Publications* 9 (2): 23. https://doi.org/10.3390/publications9020023.

*Basol, Melisa, Jon Roozenbeek, Manon Berriche, Fatih Uenal, William P. McClanahan, and Sander van der Linden. 2021. "Towards Psychological Herd Immunity: Cross-Cultural Evidence for Two Prebunking Interventions Against COVID-19 Misinformation." *Big Data & Society* 8 (1): 205395172110138. https://doi.org/10.1177/20539517211013868.

Batailler, Cédric, Skylar M. Brannon, Paul E. Teas, and Bertram Gawronski. 2022. "A Signal Detection Approach to Understanding the Identification of Fake News." *Perspectives on Psychological Science* 17 (1): 78–98. https://doi.org/10.1177/1745691620986135.

Bauer, Martin W., Nick Allum, and Steve Miller. 2007. "What Can We Learn from 25 Years of PUS Survey Research? Liberating and Expanding the Agenda." *Public Understanding of Science* 16 (1): 79–95. https://doi.org/10.1177/0963662506071287.

Bayes, Robin, and James N Druckman. 2021. "Motivated Reasoning and Climate Change." *Current Opinion in Behavioral Sciences* 42 (December): 27–35. https://doi.org/10.1016/j.cobeha.2021.02.009.

Beck, Ulrich. 1992. *Risk society: towards a new modernity.* Repr. Theory, culture & society. London: Sage.

Bednarik, Peter, and Thomas Schultze. 2015. "The Effectiveness of Imperfect Weighting in Advice Taking." *Judgment and Decision Making* 10 (3): 265–76. https://doi.org/10.1017/S1930297500004666.

Bernard, Stéphane, Paul Harris, Nathalie Terrier, and Fabrice Clément. 2015. "Children

Weigh the Number of Informants and Perceptual Uncertainty When Identifying Objects." *Journal of Experimental Child Psychology* 136 (August): 70–81. https://doi.org/10.1016/j.jecp.2015.03.009.

Bernard, Stéphane, Joëlle Proust, and Fabrice Clément. 2015. "Four- to Six-Year-Old Children's Sensitivity to Reliability Versus Consensus in the Endorsement of Object Labels." *Child Development* 86 (4): 1112–24. https://doi.org/10.1111/cdev.12366.

Bertsou, Eri. 2019. "Rethinking Political Distrust." *European Political Science Review* 11 (2): 213–30. https://doi.org/10.1017/S1755773919000080.

Besalú, Reinald, and Carles Pont-Sorribes. 2021. "Credibility of Digital Political News in Spain: Comparison Between Traditional Media and Social Media." *Social Sciences* 10 (5): 170. https://doi.org/10.3390/socsci10050170.

Besley, John C., Nicole M. Lee, and Geah Pressgrove. 2021a. "Reassessing the Variables Used to Measure Public Perceptions of Scientists." *Science Communication* 43 (1): 3–32. https://doi.org/10.1177/1075547020949547.

———. 2021b. "Reassessing the Variables Used to Measure Public Perceptions of Scientists." *Science Communication* 43 (1): 3–32. https://doi.org/10.1177/1075547020949547.

Besley, John C., and Leigh Anne Tiffany. 2023. "What Are You Assessing When You Measure "Trust" in Scientists with a Direct Measure?" *Public Understanding of Science* 32 (6): 709–26. https://doi.org/10.1177/09636625231161302.

Blancke, Stefaan, Frank Van Breusegem, Geert De Jaeger, Johan Braeckman, and Marc Van Montagu. 2015. "Fatal Attraction: The Intuitive Appeal of GMO Opposition." *Trends in Plant Science* 20 (7): 414–18. https://doi.org/10.1016/j.tplants.2015.03.011.

Bogert, J. M., Buczny ,J., Harvey ,J. A., and J. and Ellers. 2024. "The Effect of Trust in Science and Media Use on Public Belief in Anthropogenic Climate Change: A Meta-Analysis." *Environmental Communication* 18 (4): 484–509. https://doi.org/10.1080/17524032.2023.2280749.

Bourdieu, Pierre. 1979. "Public Opinion Does Not Exist." *Communication and Class Struggle* 1: 124–30.

Brandt, Allan M. 1978. "Racism and Research: The Case of the Tuskegee Syphilis Study." *The Hastings Center Report* 8 (6): 21. https://doi.org/10.2307/3561468.

*Brashier, Nadia M., Gordon Pennycook, Adam J. Berinsky, and David G. Rand. 2021. "Timing Matters When Correcting Fake News." *Proceedings of the National Academy of Sciences* 118 (5): e2020043118. https://doi.org/10.1073/pnas.2020043118.

Brashier, Nadia M., and Elizabeth J. Marsh. 2020. "Judging Truth." *Annual Review of Psychology* 71 (1): 499–515. https://doi.org/10.1146/annurev-psych-010419-050807.

Brennen, Tim, and Svein Magnussen. 2023. "Lie Detection: What Works?" *Current Directions in Psychological Science*, May, 096372142311730. https://doi.org/10.1177/09637214231173095.

Brian, Kennedy, and Alec Tyson. 2023. "Americans' Trust in Scientists, Positive Views of Science Continue to Decline." https://www.pewresearch.org/science/2023/11/14/americans-trust-in-scientists-positive-views-of-science-continue-to-decline/.

*Bronstein, Michael V., Gordon Pennycook, Adam Bear, David G. Rand, and Tyrone D. Cannon. 2019. "Belief in Fake News Is Associated with Delusionality, Dogmatism, Religious Fundamentalism, and Reduced Analytic Thinking." *Journal of Applied Research in Memory and Cognition* 8 (1): 108–17. https://doi.org/10.1016/j.jarmac.2018.09.005.

Bruder, Martin, Peter Haffke, Nick Neave, Nina Nouripanah, and Roland Imhoff. 2013. "Measuring Individual Differences in Generic Beliefs in Conspiracy Theories Across Cultures: Conspiracy Mentality Questionnaire." *Frontiers in Psychology* 4. https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00225.

Bryanov, Kirill, Reinhold Kliegl, Olessia Koltsova, Alex Miltsov, Sergei Pashakhin, Alexander Porshnev, Yadviga Sinyavskaya, Maksim Terpilovskii, and Victoria Vziatysheva. 2023. "What Drives Perceptions of Foreign News Coverage Credibility? A Cross-National Experiment Including Kazakhstan, Russia, and Ukraine." *Political Communication* 40 (2): 115–46. https://doi.org/10.1080/10584609.2023.2172492.

Bryanov, Kirill, and Victoria Vziatysheva. 2021. "Determinants of Individuals' Belief in Fake News: A Scoping Review Determinants of Belief in Fake News." Edited by Stefano Triberti. *PLOS ONE* 16 (6): e0253717. https://doi.org/10.1371/journal.pone.0253717.

Budescu, David V, and Adrian K Rantilla. 2000. "Confidence in Aggregation of Expert Opinions." *Acta Psychologica* 104 (3): 371–98. https://doi.org/10.1016/S0001-6918(00)00037-8.

Budescu, David V, Adrian K Rantilla, Hsiu-Ting Yu, and Tzur M Karelitz. 2003. "The Effects of Asymmetry Among Advisors on the Aggregation of Their Opinions." *Organizational Behavior and Human Decision Processes* 90 (1): 178–94. https://doi.org/10.1016/S0749-5978(02)00516-2.

Budescu, David V, and Hsiu-Ting Yu. 2007. "Aggregation of Opinions Based on Correlated Cues and Advisors." *Journal of Behavioral Decision Making* 20 (2): 153–77. https://doi.org/10.1002/bdm.547.

Calvillo, Dustin P., and Thomas J. Smelter. 2020. "An Initial Accuracy Focus Reduces the Effect of Prior Exposure on Perceived Accuracy of News Headlines." *Cognitive Research: Principles and Implications* 5 (1): 55. https://doi.org/10.1186/s41235-020-00257-y.

Capraro, Valerio, and Tatiana Celadin. n.d. ""I Think This News Is Accurate": Endorsing Accuracy Decreases the Sharing of Fake News and Increases the Sharing of Real News." *Personality and Social Psychology Bulletin*.

Cartwright, Nancy, Jeremy Hardie, Eleonora Montuschi, Matthew Soleiman, and Ann C. Thresher. 2022. *The Tangle of Science: Reliability Beyond Method, Rigour, and Objectivity*. New York: Oxford University Press.

Chen, Eva E., Kathleen H. Corriveau, and Paul L. Harris. 2013. "Children Trust a Consensus Composed of Outgroup Members-But Do Not Retain That Trust." *Child Development* 84 (1): 269–82. https://doi.org/10.1111/j.1467-8624.2012.01850.x.

Chen, Xi, Gordon Pennycook, and David Rand. 2023. "What Makes News Sharable on Social Media?" *Journal of Quantitative Description: Digital Media* 3.

*Clayton, Katherine, Spencer Blair, Jonathan A. Busam, Samuel Forstner, John Glance, Guy Green, Anna Kawata, et al. 2020. "Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media." *Political Behavior* 42 (4): 1073–95. https://doi.org/10.1007/s11109-019-09533-0.

Clemm Von Hohenberg, Bernhard. 2023. "Truth and Bias, Left and Right: Testing Ideological

Asymmetries with a Realistic News Supply." *Public Opinion Quarterly* 87 (2): 267–92. https://doi.org/10.1093/poq/nfad013.

Collins, Randall. 2002. *The sociology of philosophies: a global theory of intellectual change*. 4. print., 1. Harvard Univ. Pr. paperback ed., 2000. Cambridge, Mass. London: Belknap Press of Harvard Univ. Press.

Cologna, Viktoria, Reto Knutti, Naomi Oreskes, and Michael Siegrist. 2021. "Majority of German Citizens, US Citizens and Climate Scientists Support Policy Advocacy by Climate Researchers and Expect Greater Political Engagement." *Environmental Research Letters* 16 (2): 024011. https://doi.org/10.1088/1748-9326/abd4ac.

Cologna, Viktoria, Niels G. Mede, Sebastian Berger, John Besley, Cameron Brick, Marina Joubert, Edward Maibach, et al. 2024. "Trust in Scientists and Their Role in Society Across 67 Countries," February. https://doi.org/10.31219/osf.io/6ay7s.

Cologna, Viktoria, Niels G. Mede, Sebastian Berger, John Besley, Cameron Brick, Marina Joubert, Edward W. Maibach, et al. 2025. "Trust in Scientists and Their Role in Society Across 68 Countries." *Nature Human Behaviour*, January, 1–18. https://doi.org/10.1038/s41562-024-02090-5.

Cologna, Viktoria, and Michael Siegrist. 2020. "The Role of Trust for Climate Change Mitigation and Adaptation Behaviour: A Meta-Analysis." *Journal of Environmental Psychology* 69 (June): 101428. https://doi.org/10.1016/j.jenvp.2020.101428.

Coppock, Alexander. 2019. "Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach." *Political Science Research and Methods* 7 (3): 613–28. https://doi.org/10.1017/psrm.2018.10.

Corriveau, Kathleen H., Maria Fusaro, and Paul L. Harris. 2009. "Going With the Flow: Preschoolers Prefer Nondissenters as Informants." *Psychological Science* 20 (3): 372–77. https://doi.org/10.1111/j.1467-9280.2009.02291.x.

Crutchfield, Richard S. 1955. "Conformity and Character." *American Psychologist* 10 (5): 191–98. https://doi.org/10.1037/h0040237.

Cuddy, Amy J. C., Susan T. Fiske, and Peter Glick. 2008. "Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map." In, 40:61–149. Elsevier. https://doi.org/10.1016/S0065-2601(07)00002-0.

De Condorcet, Nicolas. 2014. *Essai Sur l'application de l'analyse à La Probabilité Des Décisions Rendues à La Pluralité Des Voix*. Cambridge University Press.

Desai, Saoirse Connor, Belinda Xie, and Brett K. Hayes. 2022. "Getting to the Source of the Illusion of Consensus." *Cognition* 223 (June): 105023. https://doi.org/10.1016/j.cognition.2022.105023.

Dias, Nicholas, Gordon Pennycook, and David G. Rand. 2020. "Emphasizing Publishers Does Not Effectively Reduce Susceptibility to Misinformation on Social Media." *Harvard Kennedy School Misinformation Review*, January. https://doi.org/10.37016/mr-2020-001.

Dietrich, Franz, and Kai Spiekermann. 2013. "Epistemic Democracy with Defensible Premises." *Economics and Philosophy* 29 (1): 87–120. https://doi.org/10.1017/S0266267113000096.

———. 2024. "Deliberation and the Wisdom of Crowds." *Economic Theory*, July. https://doi.org/10.1007/s00199-024-01595-4.

Douglas, Karen M., Joseph E. Uscinski, Robbie M. Sutton, Aleksandra Cichocka, Turkay Nefes, Chee Siang Ang, and Farzin Deravi. 2019. "Understanding Conspiracy Theories." *Political Psychology* 40 (S1): 3–35. https://doi.org/10.1111/pops.12568.

Druckman, James N. 2015. "Communicating Policy-Relevant Science." *PS: Political Science & Politics* 48 (S1): 58–69. https://doi.org/10.1017/S1049096515000438.

———. 2022. "Threats to Science: Politicization, Misinformation, and Inequalities." *The ANNALS of the American Academy of Political and Social Science* 700 (1): 8–24. https://doi.org/10.1177/00027162221095431.

Druckman, James N., and Mary C. McGrath. 2019. "The Evidence for Motivated Reasoning in Climate Change Preference Formation." *Nature Climate Change* 9 (2): 111–19. https://doi.org/10.1038/s41558-018-0360-1.

Drummond, Caitlin, and Baruch Fischhoff. 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): 9587–92. https://doi.org/10.1073/pnas.1704882114.

Durant, John R., Geoffrey A. Evans, and Geoffrey P. Thomas. 1989. "The Public Understanding of Science." *Nature* 340 (6228): 11–14. https://doi.org/10.1038/340011a0.

Egan, Patrick J., and Megan Mullin. 2017. "Climate Change: US Public Opinion." *Annual Review of Political Science* 20: 209–27.

Egelhofer, Jana Laura, Ming Boyer, Sophie Lecheler, and Loes Aaldering. 2022. "Populist Attitudes and Politicians' Disinformation Accusations: Effects on Perceptions of Media and Politicians." *Journal of Communication* 72 (6): 619–32. https://doi.org/10.1093/joc/jqac031.

Egelhofer, Jana Laura, and Sophie Lecheler. 2019. "Fake News as a Two-Dimensional Phenomenon: A Framework and Research Agenda." *Annals of the International Communication Association* 43 (2): 97–116. https://doi.org/10.1080/23808985.2019.1602782.

Einav, Shiri. 2018. "Thinking for Themselves? The Effect of Informant Independence on Children's Endorsement of Testimony from a Consensus." *Social Development* 27 (1): 73–86. https://doi.org/10.1111/sode.12264.

Epstein, Ziv, Nathaniel Sirlin, Antonio Arechar, Gordon Pennycook, and David Rand. 2023. "The Social Media Context Interferes with Truth Discernment." *Science Advances* 9 (9): eabo6169.

*Erlich, Aaron, and Calvin Garner. 2023. "Is Pro-Kremlin Disinformation Effective? Evidence from Ukraine." *The International Journal of Press/Politics* 28 (1): 5–28. https://doi.org/10.1177/19401612211045221.

*Faragó, Laura, Péter Krekó, and Gábor Orosz. 2023. "Hungarian, Lazy, and Biased: The Role of Analytic Thinking and Partisanship in Fake News Discernment on a Hungarian Representative Sample." *Scientific Reports* 13 (1): 178. https://doi.org/10.1038/s41598-022-26724-8.

*Fazio, Lisa, David Rand, Stephan Lewandowsky, Mark Susmann, Adam J. Berinsky, Andrew Guess, Panayiota Kendeou, et al. n.d. "Combating Misinformation: A Megastudy of Nine Interventions Designed to Reduce the Sharing of and Belief in False and Misleading Headlines." https://doi.org/10.31234/osf.io/uyjha.

Feinstein, Justin S., Melissa C. Duff, and Daniel Tranel. 2010. "Sustained Experience of Emotion After Loss of Memory in Patients with Amnesia." *Proceedings of the National Academy of Sciences* 107 (17): 7674–79. https://doi.org/10.1073/pnas.0914054107.

Fernbach, Philip M., Nicholas Light, Sydney E. Scott, Yoel Inbar, and Paul Rozin. 2019. "Extreme Opponents of Genetically Modified Foods Know the Least but Think They Know the Most." *Nature Human Behaviour* 3 (3): 251–56. https://doi.org/10.1038/s41562-018-0520-3.

Fiske, S. T., and C. Dupree. 2014. "Gaining Trust as Well as Respect in Communicating to Motivated Audiences about Science Topics." *Proceedings of the National Academy of Sciences* 111 (Supplement_4): 13593–97. https://doi.org/10.1073/pnas.1317505111.

Fletcher, Richard, and Rasmus-Kleis Nielsen. 2017. "People Dont Trust News Media–and This Is Key to the Global Misinformation Debate." *AA. VV., Understanding and Addressing the Disinformation Ecosystem*, 13–17.

Forgas, Joseph P., and Simon M. Laham. 2016. "Halo Effects." In, 276–90. Psychology Press.

Fricker, Elizabeth. 2006. "Testimony and Epistemic Autonomy." In, edited by Jennifer Lackey and Ernest Sosa, 1st ed., 225–50. Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780199276011.003.0011.

Fuentes, Agustín. 2023. "Systemic Racism in Science: Reactions Matter." *Science* 381 (6655): eadj7675. https://doi.org/10.1126/science.adj7675.

Funk, Cary, Courtney Johnson, and Meg Hefferon. 2019. "5 Key Findings about Public Trust in Scientists in the u.s." https://www.pewresearch.org/fact-tank/2019/08/05/5-key-findings-about-public-trust-in-scientists-in-the-u-s/.

Funk, Cary, and Brian Kennedy. 2020. "Public Confidence in Scientists Has Remained Stable for Decades."

Funk, Cary, Alec Tyson, Brian Kennedy, and Courtney Johnson. 2020. "Science and Scientists Held in High Esteem Across Global Publics." https://www.pewresearch.org/science/2020/09/29/science-and-scientists-held-in-high-esteem-across-global-publics/.

Fusaro, Maria, and Paul L. Harris. 2008. "Children Assess Informant Reliability Using Bystanders' Non-Verbal Cues." *Developmental Science* 11 (5): 771–77. https://doi.org/10.1111/j.1467-7687.2008.00728.x.

Galton, Francis. 1907. "Vox Populi." *Nature* 75 (1949): 450–51. https://doi.org/10.1038/075450a0.

Garrett, R. Kelly, and Robert M. Bond. 2021. "Conservatives' Susceptibility to Political Misperceptions." *Science Advances* 7 (23): eabf1234. https://doi.org/10.1126/sciadv.abf1234.

Gauchat, Gordon. 2011. "The Cultural Authority of Science: Public Trust and Acceptance of Organized Science." *Public Understanding of Science* 20 (6): 751–70. https://doi.org/10.1177/0963662510365246.

———. 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): 167–87. https://doi.org/10.1177/0003122412438225.

———. 2023. "The Legitimacy of Science." *Annual Review of Sociology* 49 (1): 263–79. https://doi.org/10.1146/annurev-soc-030320-035037.

Gauchat, Gordon, and Kenneth T. Andrews. 2018. "The Cultural-Cognitive Mapping of

Scientific Professions.” *American Sociological Review* 83 (3): 567–95. https://doi.org/10.1177/0003122418773353.

Gawronski, Bertram, Nyx L. Ng, and Dillon M. Luke. 2023. “Truth Sensitivity and Partisan Bias in Responses to Misinformation.” *Journal of Experimental Psychology: General* 152 (8): 2205–36. https://doi.org/10.1037/xge0001381.

Gibbons, Robert D., Donald R. Hedeker, and John M. Davis. 1993. “Estimation of Effect Size from a Series of Experiments Involving Paired Comparisons.” *Journal of Educational Statistics* 18 (3): 271–79. https://doi.org/10.3102/10769986018003271.

Giddens, Anthony. 1991. *Modernity and Self-Identity: Self and Society in the Late Modern Age*. Stanford, Calif: Stanford University Press.

Gligorić, Vukašin, Gerben A. van Kleef, and Bastiaan T. Rutjens. 2024. “How Social Evaluations Shape Trust in 45 Types of Scientists.” *PLOS ONE* 19 (4): e0299621. https://doi.org/10.1371/journal.pone.0299621.

Globig, Laura K, Nora Holtz, and Tali Sharot. 2023. “Changing the Incentive Structure of Social Media Platforms to Halt the Spread of Misinformation.” *eLife* 12 (June): e85767. https://doi.org/10.7554/eLife.85767.

*Gottlieb, Jessica, Claire Adida, and Richard Moussa. n.d. “Reducing Misinformation in a Polarized Context: Experimental Evidence from Côte d’ivoire.” https://doi.org/10.31219/osf.io/6x4wy.

Guay, Brian, Adam J. Berinsky, Gordon Pennycook, and David Rand. 2023. “How to Think about Whether Misinformation Interventions Work.” *Nature Human Behaviour* 7 (8): 1231–33. https://doi.org/10.1038/s41562-023-01667-w.

Guay, Brian, Adam Berinsky, Gordon Pennycook, and David Rand. n.d. “How to Think about Whether Misinformation Interventions Work.” https://doi.org/10.31234/osf.io/gv8qx.

*Guess, Andrew M., Michael Lerner, Benjamin Lyons, Jacob M. Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjan Sircar. 2020. “A Digital Media Literacy Intervention Increases Discernment Between Mainstream and False News in the United States and India.” *Proceedings of the National Academy of Sciences* 117 (27): 15536–45. https://doi.org/10.1073/pnas.1920498117.

*Guess, Andrew, Shannon McGregor, Gordon Pennycook, and David Rand. n.d. “Unbundling Digital Media Literacy Tips: Results from Two Experiments.” https://doi.org/10.31234/osf.io/u34fp.

Gunaydin, Gul, Emre Selcuk, and Vivian Zayas. 2017. “Impressions Based on a Portrait Predict, 1-Month Later, Impressions Following a Live Interaction.” *Social Psychological and Personality Science* 8 (1): 36–44. https://doi.org/10.1177/1948550616662123.

Habermas, Jürgen. 1989. *Jürgen Habermas on society and politics: a reader*. Edited by Steven Seidman. 1. [print.]. Sociology, political science, philosophy. Boston: Beacon Press.

Hacquin, Anne-Sophie, Sacha Altay, Lene Aarøe, and Hugo Mercier. 2021. “Disgust Sensitivity and Public Opinion on Nuclear Energy.” *Journal of Environmental Psychology*, 101749.

Hahn, Ulrike, Christoph Merdes, and Momme von Sydow. 2018. “How Good Is Your Evidence and How Would You Know?” *Topics in Cognitive Science* 10 (4): 660–78. https://doi.org/10.1111/tops.12374.

*Hameleers, Michael, Marina Tulin, Claes De Vreese, Toril Aalberg, Peter Van Aelst, Ana Sofia Cardenal, Nicoleta Corbu, et al. 2023. "Mistakenly Misinformed or Intentionally Deceived? Mis- and Disinformation Perceptions on the Russian War in Ukraine Among Citizens in 19 Countries." *European Journal of Political Research*, December, 1475–6765.12646. https://doi.org/10.1111/1475-6765.12646.

Hardwig, John. 1991. "The Role of Trust in Knowledge." *The Journal of Philosophy* 88 (12): 693. https://doi.org/10.2307/2027007.

Harkins, Stephen G., and Richard E. Petty. 1987. "Information Utility and the Multiple Source Effect." *Journal of Personality and Social Psychology* 52 (2): 260.

Harrer, Mathias, Pim Cuijpers, Furukawa Toshi A, and David D Ebert. 2021. *Doing Meta-Analysis with r: A Hands-on Guide*. 1st ed. Boca Raton, FL; London: Chapman & Hall/CRC Press.

Harries, Clare, Ilan Yaniv, and Nigel Harvey. 2004. "Combining Advice: The Weight of a Dissenting Opinion in the Consensus." *Journal of Behavioral Decision Making* 17 (5): 333–48. https://doi.org/10.1002/bdm.474.

Harris, Keith Raymond. 2023. "Conspiracy Theories, Populism, and Epistemic Autonomy." *Journal of the American Philosophical Association* 9 (1): 21–36. https://doi.org/10.1017/apa.2021.44.

Harvey, Nigel, and Ilan Fischer. 1997. "Taking Advice: Accepting Help, Improving Judgment, and Sharing Responsibility." *Organizational Behavior and Human Decision Processes* 70 (2): 117–33. https://doi.org/10.1006/obhd.1997.2697.

Hastie, Reid, and Tatsuya Kameda. 2005. "The Robust Beauty of Majority Rules in Group Decisions." *Psychological Review* 112 (2): 494–508. https://doi.org/10.1037/0033-295X.112.2.494.

Hendriks, Friederike, Dorothe Kienhues, and Rainer Bromme. 2015. "Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI)." Edited by Jelte M. Wicherts. *PLOS ONE* 10 (10): e0139309. https://doi.org/10.1371/journal.pone.0139309.

———. 2016. "Trust in Science and the Science of Trust." In, edited by Bernd Blöbaum, 143–59. Cham: Springer International Publishing. http://link.springer.com/10.1007/978-3-319-28059-2_8.

———. 2020. "Replication Crisis = Trust Crisis? The Effect of Successful Vs Failed Replications on Laypeople's Trust in Researchers and Research." *Public Understanding of Science* 29 (3): 270–88. https://doi.org/10.1177/0963662520902383.

Herrmann, Patricia A., Cristine H. Legare, Paul L. Harris, and Harvey Whitehouse. 2013. "Stick to the Script: The Effect of Witnessing Multiple Actors on Children's Imitation." *Cognition* 129 (3): 536–43. https://doi.org/10.1016/j.cognition.2013.08.010.

Hess, Nicole H., and Edward H. Hagen. 2006. "Psychological Adaptations for Assessing Gossip Veracity." *Human Nature* 17 (3): 337–54. https://doi.org/10.1007/s12110-006-1013-z.

Heyman, Richard E., Michael F. Lorber, J. Mark Eddy, and Tessa V. West. 2014. "Behavioral Observation and Coding." In, edited by Charles M. Judd and Harry T. Reis, 2nd ed., 345–72. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511996481.018.

Higgins, Julian PT, James Thomas, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew J. Page, and Vivian A. Welch. 2019. *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons.

Higham, Philip A., Ariana Modirrousta-Galian, and Tina Seabrooke. 2024. "Mean Rating Difference Scores Are Poor Measures of Discernment: The Role of Response Criteria." *Current Opinion in Psychology* 56 (April): 101785. https://doi.org/10.1016/j.copsyc.2023.101785.

*Hlatky, Roman. n.d. "Unintended Consequences? Russian Disinformation and Public Opinion." https://doi.org/10.31219/osf.io/85vmt.

Hoes, Emma, Brian Aitken, Jingwen Zhang, Tomasz Gackowski, and Magdalena Wojcieszak. 2023. "Prominent Misinformation Interventions Reduce Misperceptions but Increase Skepticism." https://doi.org/10.31234/osf.io/zmpdu.

Hornsey, Matthew J. 2020. "Why Facts Are Not Enough: Understanding and Managing the Motivated Rejection of Science." *Current Directions in Psychological Science* 29 (6): 583–91. https://doi.org/10.1177/0963721420969364.

Hornsey, Matthew J., and Kelly S. Fielding. 2017. "Attitude Roots and Jiu Jitsu Persuasion: Understanding and Overcoming the Motivated Rejection of Science." *American Psychologist* 72 (5): 459–73. https://doi.org/10.1037/a0040437.

Hornsey, Matthew J., Emily A. Harris, Paul G. Bain, and Kelly S. Fielding. 2016. "Meta-Analyses of the Determinants and Outcomes of Belief in Climate Change." *Nature Climate Change* 6 (6): 622–26. https://doi.org/10.1038/nclimate2943.

Hornsey, Matthew J., Emily A. Harris, and Kelly S. Fielding. 2018a. "The Psychological Roots of Anti-Vaccination Attitudes: A 24-Nation Investigation." *Health Psychology* 37 (4): 307–15. https://doi.org/10.1037/hea0000586.

———. 2018b. "Relationships Among Conspiratorial Beliefs, Conservatism and Climate Scepticism Across Nations." *Nature Climate Change* 8 (7): 614–20. https://doi.org/10.1038/s41558-018-0157-2.

Huang, Jacob Poushter, Moira Fagan and Christine. 2021. "Americans Are Less Concerned – but More Divided – on Climate Change Than People Elsewhere." https://www.pewresearch.org/short-reads/2021/09/14/americans-are-less-concerned-but-more-divided-on-climate-change-than-people-elsewhere/.

Hutmacher, Fabian, Regina Reichardt, and Markus Appel. 2024. "Motivated Reasoning about Climate Change and the Influence of Numeracy, Need for Cognition, and the Dark Factor of Personality." *Scientific Reports* 14 (1): 5615. https://doi.org/10.1038/s41598-024-55930-9.

Intemann, Kristen. 2023. "Science Communication and Public Trust in Science." *Interdisciplinary Science Reviews* 48 (2): 350–65. https://doi.org/10.1080/03080188.2022.2152244.

Jayles, Bertrand, Hye-rin Kim, Ramón Escobedo, Stéphane Cezera, Adrien Blanchet, Tatsuya Kameda, Clément Sire, and Guy Theraulaz. 2017. "How Social Information Can Improve Estimation Accuracy in Human Groups." *Proceedings of the National Academy of Sciences* 114 (47): 12620–25. https://doi.org/10.1073/pnas.1703695114.

Judd, Emily J., Jessica E. Tierney, Daniel J. Lunt, Isabel P. Montañez, Brian T. Huber, Scott L. Wing, and Paul J. Valdes. 2024. "A 485-Million-Year History of Earth's Surface Temperature." *Science* 385 (6715): eadk3705. https://doi.org/10.1126/science.adk3705.

Kahan, Dan M. 2015. "Climate-Science Communication and the Measurement Problem." *Political Psychology* 36 (S1): 1–43. https://doi.org/10.1111/pops.12244.

Kahan, Dan M., Ellen Peters, Maggie Wittlin, Paul Slovic, Lisa Larrimore Ouellette, Donald Braman, and Gregory Mandel. 2012. "The Polarizing Impact of Science Literacy and Numeracy on Perceived Climate Change Risks." *Nature Climate Change* 2 (10): 732–35. https://doi.org/10.1038/nclimate1547.

Kämmer, Juliane E., Shoham Choshen-Hillel, Johannes Müller-Trede, Stephanie L. Black, and Jürgen Weibler. 2023. "A Systematic Review of Empirical Studies on Advice-Based Decisions in Behavioral and Organizational Research." *Decision* 10 (2): 107–37. https://doi.org/10.1037/dec0000199.

Karlsen, Rune, and Toril Aalberg. 2023. "Social Media and Trust in News: An Experimental Study of the Effect of Facebook on News Story Credibility." *Digital Journalism* 11 (1): 144–60. https://doi.org/10.1080/21670811.2021.1945938.

Kennedy, Alec Tyson, and Brian. 2024. "Public Trust in Scientists and Views on Their Role in Policymaking." https://www.pewresearch.org/science/2024/11/14/public-trust-in-scientists-and-views-on-their-role-in-policymaking/.

Kerr, John R., and Marc S. Wilson. 2021. "Right-Wing Authoritarianism and Social Dominance Orientation Predict Rejection of Science and Scientists." *Group Processes & Intergroup Relations* 24 (4): 550–67. https://doi.org/10.1177/1368430221992126.

Koenig, Melissa A., Fabrice Clément, and Paul L. Harris. 2004. "Trust in Testimony: Children's Use of True and False Statements." *Psychological Science* 15 (10): 694–98. https://doi.org/10.1111/j.0956-7976.2004.00742.x.

*Koetke, Jonah, Karina Schumann, Tenelle Porter, and Ilse Smilo-Morgan. 2023. "Fallibility Salience Increases Intellectual Humility: Implications for People's Willingness to Investigate Political Misinformation." *Personality and Social Psychology Bulletin* 49 (5): 806–20. https://doi.org/10.1177/01461672221080979.

Koetke, Jonah, Karina Schumann, Shauna M. Bowes, and Nina Vaupotič. 2024. "The Effect of Seeing Scientists as Intellectually Humble on Trust in Scientists and Their Research." *Nature Human Behaviour*, November, 1–14. https://doi.org/10.1038/s41562-024-02060-x.

Koetke, Jonah, Karina Schumann, and Tenelle Porter. 2021. "Trust in Science Increases Conservative Support for Social Distancing." *Group Processes & Intergroup Relations* 24 (4): 680–97. https://doi.org/10.1177/1368430220985918.

König, Laura M, Marlene Sophie Altenmüller, Julian Fick, Jan Crusius, Oliver Genschow, and Melanie Sauerland. 2023. "How to Communicate Science to the Public? Recommendations for Effective Written Communication Derived from a Systematic Review." https://doi.org/10.31234/osf.io/cwbrs.

Krause, Nicole M, Dominique Brossard, Dietram A Scheufele, Michael A Xenos, and Keith Franke. 2019. "Trends—americans' Trust in Science and Scientists." *Public Opinion Quarterly* 83 (4): 817–36. https://doi.org/10.1093/poq/nfz041.

*Kreps, Sarah E., and Douglas L. Kriner. 2023. "Assessing Misinformation Recall and Accuracy Perceptions: Evidence from the COVID-19 Pandemic." *Harvard Kennedy School Misinformation Review*, October. https://doi.org/10.37016/mr-2020-123.

Ladha, Krishna K. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes."

American Journal of Political Science 36 (3): 617. https://doi.org/10.2307/2111584.

Lantian, Anthony, Dominique Muller, Cécile Nurra, and Karen M. Douglas. 2016. "Measuring Belief in Conspiracy Theories: Validation of a French and English Single-Item Scale." *International Review of Social Psychology* 29 (1): 1. https://doi.org/10.5334/irsp.8.

Larrick, Richard P., and Jack B. Soll. 2006. "Intuitions About Combining Opinions: Misappreciation of the Averaging Principle." *Management Science* 52 (1): 111–27. https://doi.org/10.1287/mnsc.1050.0459.

*Lee, Eun-Ju, and Jeong-woo Jang. 2023. "How Political Identity and Misinformation Priming Affect Truth Judgments and Sharing Intention of Partisan News." *Digital Journalism* 0 (0): 1–20. https://doi.org/10.1080/21670811.2022.2163413.

Lee, John J. 2021. "Party Polarization and Trust in Science: What about Democrats?" *Socius* 7 (January): 237802312110101. https://doi.org/10.1177/23780231211010101.

Levine, Timothy R. 2014. "Truth-Default Theory (TDT): A Theory of Human Deception and Deception Detection." *Journal of Language and Social Psychology* 33 (4): 378–92. https://doi.org/10.1177/0261927X14535916.

Lewandowsky, Stephan, Gilles E. Gignac, and Klaus Oberauer. 2013. "The Role of Conspiracist Ideation and Worldviews in Predicting Rejection of Science." *PLOS ONE* 8 (10): e75637. https://doi.org/10.1371/journal.pone.0075637.

Lewandowsky, Stephan, and Klaus Oberauer. 2016. "Motivated Rejection of Science." *Current Directions in Psychological Science* 25 (4): 217–22. https://doi.org/10.1177/0963721416654436.

———. 2021. "Worldview-Motivated Rejection of Science and the Norms of Science." *Cognition* 215 (October): 104820. https://doi.org/10.1016/j.cognition.2021.104820.

Li, Nan, and Yachao Qian. 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. https://doi.org/10.1017/pls.2021.18.

Linden, Sander van der. 2021. "The Gateway Belief Model (GBM): A Review and Research Agenda for Communicating the Scientific Consensus on Climate Change." *Current Opinion in Psychology* 42 (December): 7–12. https://doi.org/10.1016/j.copsyc.2021.01.005.

Lindholt, Marie Fly, Frederik Jørgensen, Alexander Bor, and Michael Bang Petersen. 2021. "Public Acceptance of COVID-19 Vaccines: Cross-National Evidence on Levels and Individual-Level Predictors Using Observational Data." *BMJ Open* 11 (6): e048172. https://doi.org/10.1136/bmjopen-2020-048172.

Liquin, Emily G., and Tania Lombrozo. 2022. "Motivated to Learn: An Account of Explanatory Satisfaction." *Cognitive Psychology* 132 (February): 101453. https://doi.org/10.1016/j.cogpsych.2021.101453.

Lombrozo, Tania. 2006. "The Structure and Function of Explanations." *Trends in Cognitive Sciences* 10 (10): 464–70. https://doi.org/10.1016/j.tics.2006.08.004.

———. 2007. "Simplicity and Probability in Causal Explanation." *Cognitive Psychology* 55 (3): 232–57. https://doi.org/10.1016/j.cogpsych.2006.09.006.

Lopes, Diniz, Jorge Vala, and Leonel Garcia-Marques. 2007. "Social Validation of Everyday Knowledge: Heterogeneity and Consensus Functionality." *Group Dynamics: Theory, Research, and Practice* 11 (3): 223–39. https://doi.org/10.1037/1089-2699.11.3.223.

*Lühring, Jula, Apeksha Shetty, Corinna Koschmieder, David Garcia, Annie Waldherr, and Hannah Metzler. n.d. "Emotions in Misinformation Studies: Distinguishing Affective State from Emotional Response and Misinformation Recognition from Acceptance." https://doi.org/10.31234/osf.io/udqms.

Luo, Mufan, Jeffrey T. Hancock, and David M. Markowitz. 2022. "Credibility Perceptions and Detection Accuracy of Fake News Headlines on Social Media: Effects of Truth-Bias and Endorsement Cues." *Communication Research* 49 (2): 171–95. https://doi.org/10.1177/0093650220921321.

Lupia, Arthur, David B. Allison, Kathleen Hall Jamieson, Jennifer Heimberg, Magdalena Skipper, and Susan M. Wolf. 2024. "Trends in US Public Confidence in Science and Opportunities for Progress." *Proceedings of the National Academy of Sciences* 121 (11): e2319488121. https://doi.org/10.1073/pnas.2319488121.

*Lutzke, Lauren, Caitlin Drummond, Paul Slovic, and Joseph Árvai. 2019. "Priming Critical Thinking: Simple Interventions Limit the Influence of Fake News about Climate Change on Facebook." *Global Environmental Change* 58 (September): 101964. https://doi.org/10.1016/j.gloenvcha.2019.101964.

*Lyons, Benjamin, Andy J. King, and Kimberly Kaphingst. n.d. "A Health Media Literacy Intervention Increases Skepticism of Both Inaccurate and Accurate Cancer News Among U.S. Adults." https://doi.org/10.31219/osf.io/hm9ty.

*Lyons, Benjamin, Ariana Modirrousta-Galian, Sacha Altay, and Nikita Antonia Salovich. 2024. "Reduce Blind Spots to Improve News Discernment? Performance Feedback Reduces Overconfidence but Does Not Improve Subsequent Discernment," February. https://doi.org/10.31219/osf.io/kgfrb.

*Lyons, Benjamin, Jacob Montgomery, and Jason Reifler. n.d. "Partisanship and Older Americans' Engagement with Dubious Political News." https://doi.org/10.31219/osf.io/etb89.

*Maertens, Rakoen, Friedrich M. Götz, Hudson F. Golino, Jon Roozenbeek, Claudia R. Schneider, Yara Kyrychenko, John R. Kerr, et al. 2024. "The Misinformation Susceptibility Test (MIST): A Psychometrically Validated Measure of News Veracity Discernment." *Behavior Research Methods* 56 (3): 1863–99. https://doi.org/10.3758/s13428-023-02124-2.

Maertens, Rakoen, Friedrich Martin Götz, Claudia R. Schneider, Jon Roozenbeek, John R Kerr, Stefan Stieger, William Patrick McClanahan, Karly Drabot, and Sander van der Linden. 2021. "The Misinformation Susceptibility Test (MIST): A Psychometrically Validated Measure of News Veracity Discernment." https://doi.org/10.31234/osf.io/gk68h.

Mahr, Johannes B., and Gergely Csibra. 2018. "Why Do We Remember? The Communicative Function of Episodic Memory." *Behavioral and Brain Sciences* 41 (January): e1. https://doi.org/10.1017/S0140525X17000012.

Mairal, Santos *Espina, Florencia Bustos, Guillermo Solovey, and Joaquín Navajas. 2023. "Interactive Crowdsourcing to Fact-Check Politicians." *Journal of Experimental Psychology: Applied*, August. https://doi.org/10.1037/xap0000492.

Mannes, Albert E. 2009. "Are We Wise About the Wisdom of Crowds? The Use of Group Judgments in Belief Revision." *Management Science* 55 (8): 1267–79. https://doi.org/10.1287/mnsc.1090.1031.

*Martel, Cameron, Gordon Pennycook, and David G. Rand. 2020. "Reliance on Emotion Promotes Belief in Fake News." *Cognitive Research: Principles and Implications* 5 (1): 47. https://doi.org/10.1186/s41235-020-00252-3.

Martel, Cameron, Jennifer Nancy Lee Allen, Gordon Pennycook, and David Gertler Rand. 2022. "Crowds Can Effectively Identify Misinformation at Scale." https://doi.org/10.31234/osf.io/2tjk7.

Mayer, Roger C., James H. Davis, and F. David Schoorman. 1995. "An Integrative Model of Organizational Trust." *The Academy of Management Review* 20 (3): 709. https://doi.org/10.2307/258792.

Maynard-Smith, John, and David Harper. 2003. *Animal Signals*. Oxford University Press.

Mede, Niels G. 2023. "Science-Related Populism: Conceptualization, Empirical Investigation, and Implications for Science Communication (Dissertation Summary)." *Studies in Communication Sciences* 23 (3). https://doi.org/10.24434/j.scoms.2023.03.4403.

Mede, Niels G., Viktoria Cologna, Sebastian Berger, John Besley, Cameron Brick, Marina Joubert, Edward W. Maibach, et al. 2025. "Perceptions of Science, Science Communication, and Climate Change Attitudes in 68 Countries – the TISP Dataset." *Scientific Data* 12 (1): 114. https://doi.org/10.1038/s41597-024-04100-7.

Mede, Niels G., and Mike S. Schäfer. 2020. "Science-Related Populism: Conceptualizing Populist Demands Toward Science." *Public Understanding of Science* 29 (5): 473–91. https://doi.org/10.1177/0963662520924259.

Mede, Niels G., Mike S. Schäfer, Julia Metag, and Kira Klinger. 2022. "Who Supports Science-Related Populism? A Nationally Representative Survey on the Prevalence and Explanatory Factors of Populist Attitudes Toward Science in Switzerland." Edited by Ghaffar Ali. *PLOS ONE* 17 (8): e0271204. https://doi.org/10.1371/journal.pone.0271204.

Mercier, Hugo. 2016. "The Argumentative Theory: Predictions and Empirical Evidence." *Trends in Cognitive Sciences* 20 (9): 689–700. https://doi.org/10.1016/j.tics.2016.07.001.

———. 2017. "How Gullible Are We? A Review of the Evidence from Psychology and Social Science." *Review of General Psychology* 21 (2): 103–22. https://doi.org/10.1037/gpr0000111.

———. 2020. *Not Born Yesterday: The Science of Who We Trust and What We Believe*. https://doi.org/10.1515/9780691198842.

Mercier, Hugo, and Nicolas Claidière. 2022. "Does Discussion Make Crowds Any Wiser?" *Cognition* 222 (May): 104912. https://doi.org/10.1016/j.cognition.2021.104912.

Mercier, Hugo, and Helena Miton. 2019. "Utilizing Simple Cues to Informational Dependency." *Evolution and Human Behavior* 40 (3): 301–14. https://doi.org/10.1016/j.evolhumbehav.2019.01.001.

Mercier, Hugo, and Olivier Morin. 2019. "Majority Rules: How Good Are We at Aggregating Convergent Opinions?" *Evolutionary Human Sciences* 1: e6. https://doi.org/10.1017/ehs.2019.6.

Merton, Robert K. 1979. *The Sociology of Science: Theoretical and Empirical Investigations*. Edited by Norman W. Storer. Chicago, IL: University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/S/bo28451565.html.

Metzger, Miriam J. 2007. "Making Sense of Credibility on the Web: Models for Evaluating

Online Information and Recommendations for Future Research." *Journal of the American Society for Information Science and Technology* 58 (13): 2078–91. https://doi.org/10.1002/asi.20672.

Mihailidis, Paul, and Bobbie Foster. 2021. "The Cost of Disbelief: Fracturing News Ecosystems in an Age of Rampant Media Cynicism." *American Behavioral Scientist* 65 (4): 616–31. https://doi.org/10.1177/0002764220978470.

Miller, Jon D. 1998. "The Measurement of Civic Scientific Literacy." *Public Understanding of Science* 7 (3): 203–23. https://doi.org/10.1088/0963-6625/7/3/001.

———. 2004. "Public Understanding of, and Attitudes Toward, Scientific Research: What We Know and What We Need to Know." *Public Understanding of Science* 13 (3): 273–94. https://doi.org/10.1177/0963662504044908.

Miton, Helena, and Hugo Mercier. 2015. "Cognitive Obstacles to Pro-Vaccination Beliefs." *Trends in Cognitive Sciences* 19 (11): 633–36. https://doi.org/10.1016/j.tics.2015.08.007.

*Modirrousta-Galian, Ariana, Philip A. Higham, and Tina Seabrooke. 2023. "Effects of Inductive Learning and Gamification on News Veracity Discernment." *Journal of Experimental Psychology: Applied* 29 (3): 599–619. https://doi.org/10.1037/xap0000458.

———. 2024. "Wordless Wisdom: The Dominant Role of Tacit Knowledge in True and Fake News Discrimination." *Journal of Applied Research in Memory and Cognition*.

Modirrousta-Galian, Ariana, and Philip A. Higham. 2023. "Gamified Inoculation Interventions Do Not Improve Discrimination Between True and Fake News: Reanalyzing Existing Research with Receiver Operating Characteristic Analysis." *Journal of Experimental Psychology: General*.

Molleman, Lucas, Alan N. Tump, Andrea Gradassi, Stefan Herzog, Bertrand Jayles, Ralf H. J. M. Kurvers, and Wouter van den Bos. 2020. "Strategies for Integrating Disparate Social Information." *Proceedings of the Royal Society B: Biological Sciences* 287 (1939): 20202413. https://doi.org/10.1098/rspb.2020.2413.

Mont'Alverne, Camila, Sumitra Badrinathan, A. Ross Arguedas, Benjamin Toff, Richard Fletcher, and R. Nielsen. 2022. "The Trust Gap: How and Why News on Digital Platforms Is Viewed More Sceptically Versus News in General."

Morgan, Thomas J. H., Kevin N. Laland, and Paul L. Harris. 2015. "The Development of Adaptive Conformity in Young Children: Effects of Uncertainty and Consensus." *Developmental Science* 18 (4): 511–24. https://doi.org/10.1111/desc.12231.

Morgan, Thomas J. H., L. E. Rendell, M. Ehn, W. Hoppitt, and K. N. Laland. 2012. "The Evolutionary Basis of Human Social Learning." *Proceedings of the Royal Society B: Biological Sciences* 279 (1729): 653–62. https://doi.org/10.1098/rspb.2011.1172.

Mourão, Rachel R., and Craig T. Robertson. 2019. "Fake News as Discursive Integration: An Analysis of Sites That Publish False, Misleading, Hyperpartisan and Sensational Information." *Journalism Studies* 20 (14): 2077–95. https://doi.org/10.1080/1461670X.2019.1566871.

*Muda, Rafał, Gordon Pennycook, Damian Hamerski, and Michał Białek. 2023. "People Are Worse at Detecting Fake News in Their Foreign Language." *Journal of Experimental Psychology: Applied* 29 (4): 712–24. https://doi.org/10.1037/xap0000475.

Munro, Geoffrey D. 2010. "The Scientific Impotence Excuse: Discounting Belief-Threatening

Scientific Abstracts." *Journal of Applied Social Psychology* 40 (3): 579–600. https://doi.org/10.1111/j.1559-1816.2010.00588.x.

Nadelson, Louis S, and Kimberly K Hardy. 2015. "Trust in Science and Scientists and the Acceptance of Evolution." *Evolution: Education and Outreach* 8 (1): 9. https://doi.org/10.1186/s12052-015-0037-4.

National Academies of Sciences, Engineering. 2024. *Understanding and Addressing Misinformation about Science*.

National Academies of Sciences, Engineering, and Medicine. 2016. *Science Literacy: Concepts, Contexts, and Consequences*. Edited by Catherine E. Snow and Kenne A. Dibner. Washington, D.C.: National Academies Press. https://doi.org/10.17226/23595.

National Science Board, National Science Foundation. 2024. "Science and Technology: Public Perceptions, Awareness, and Information Sources." https://ncses.nsf.gov/pubs/nsb20244.

Newman, Nic, Richard Fletcher, Kirsten Eddy, Craig T. Robertson, and Rasmus Kleis Nielsen. 2023. "Digital News Report 2023."

Newman, Nic, Richard Fletcher, Craig T Robertson, Kirsten Eddy, and Rasmus Kleis Nielsen. 2022. "Reuters Institute Digital News Report 2022."

Nguyen, C. Thi. 2020. "Cognitive Islands and Runaway Echo Chambers: Problems for Epistemic Dependence on Experts." *Synthese* 197 (7): 2803–21. https://doi.org/10.1007/s11229-018-1692-0.

Nobles, Melissa, Chad Womack, Ambroise Wonkam, and Elizabeth Wathuti. 2022. "Science Must Overcome Its Racist Legacy: Nature's Guest Editors Speak." *Nature* 606 (7913): 225–27. https://doi.org/10.1038/d41586-022-01527-z.

Noy, Shiri, and Timothy L. O'Brien. 2019. "Science for Good? The Effects of Education and National Context on Perceptions of Science." *Public Understanding of Science* 28 (8): 897–916. https://doi.org/10.1177/0963662519863575.

*OECD. 2022. "An International Effort Using Behavioural Science to Tackle the Spread of Misinformation." https://doi.org/10.1787/b7709d4f-en.

Oreskes, Naomi. 2019. *Why trust science?* The University Center for Human Values series. Princeton Oxford: Princeton University Press.

*Orosz, Gábor, Benedek Paskuj, Laura Faragó, and Péter Krekó. 2023. "A Prosocial Fake News Intervention with Durable Effects." *Scientific Reports* 13 (1): 3958. https://doi.org/10.1038/s41598-023-30867-7.

Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, et al. 2021. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." *BMJ* 372 (March): n71. https://doi.org/10.1136/bmj.n71.

Pardo, Rafael, and Félix Calvo. 2004. "The Cognitive Dimension of Public Perceptions of Science: Methodological Issues." *Public Understanding of Science* 13 (3): 203–27. https://doi.org/10.1177/0963662504045002.

Paul, Christopher, and Miriam Matthews. 2016. "The Russian "Firehose of Falsehood" Propaganda Model." *Rand Corporation* 2 (7): 1–10.

*Pehlivanoglu, Didem, Tian Lin, Farha Deceus, Amber Heemskerk, Natalie C. Ebner, and Brian S. Cahill. 2021. "The Role of Analytical Reasoning and Source Credibility on the

Evaluation of Real and Fake Full-Length News Articles." *Cognitive Research: Principles and Implications* 6 (1): 24. https://doi.org/10.1186/s41235-021-00292-3.

*Pennycook, Gordon, Tyrone D. Cannon, and David G. Rand. 2018. "Prior Exposure Increases Perceived Accuracy of Fake News." *Journal of Experimental Psychology: General* 147 (12): 1865–80. https://doi.org/10.1037/xge0000465.

*Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. 2020. "Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention." *Psychological Science* 31 (7): 770–80. https://doi.org/10.1177/0956797620939054.

*Pennycook, Gordon, and David G. Rand. 2020. "Who Falls for Fake News? The Roles of Bullshit Receptivity, Overclaiming, Familiarity, and Analytic Thinking." *Journal of Personality* 88 (2): 185–200. https://doi.org/10.1111/jopy.12476.

Pennycook, Gordon, Jabin Binnendyk, Christie Newton, and David G. Rand. 2021. "A Practical Guide to Doing Behavioral Research on Fake News and Misinformation." *Collabra: Psychology* 7 (1): 25293. https://doi.org/10.1525/collabra.25293.

Pennycook, Gordon, Jabin Binnendyk, and David Rand. 2022. "Overconfidently Conspiratorial: Conspiracy Believers Are Dispositionally Overconfident and Massively Overestimate How Much Others Agree with Them," December. https://doi.org/10.31234/osf.io/d5fz2.

Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles, and David G. Rand. 2021. "Shifting Attention to Accuracy Can Reduce Misinformation Online." *Nature* 592 (7855): 590–95. https://doi.org/10.1038/s41586-021-03344-2.

Pennycook, Gordon, and David G. Rand. 2019. "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning Than by Motivated Reasoning." *Cognition* 188 (July): 39–50. https://doi.org/10.1016/j.cognition.2018.06.011.

*Pereira, Frederico Batista, Natália S. Bueno, Felipe Nunes, and Nara Pavão. 2023. "Inoculation Reduces Misinformation: Experimental Evidence from Multidimensional Interventions in Brazil." *Journal of Experimental Political Science*, July, 1–12. https://doi.org/10.1017/XPS.2023.11.

Persson, Emil, David Andersson, Lina Koppel, Daniel Västfjäll, and Gustav Tinghög. 2021. "A Preregistered Replication of Motivated Numeracy." *Cognition* 214 (September): 104768. https://doi.org/10.1016/j.cognition.2021.104768.

Petersen, Michael Bang, Alexander Bor, Frederik Jørgensen, and Marie Fly Lindholt. 2021. "Transparent Communication about Negative Features of COVID-19 Vaccines Decreases Acceptance but Increases Trust." *Proceedings of the National Academy of Sciences* 118 (29): e2024597118. https://doi.org/10.1073/pnas.2024597118.

Pfänder, Jan, Benoît De Courson, and Hugo Mercier. 2025. "How Wise Is the Crowd: Can We Infer People Are Accurate and Competent Merely Because They Agree with Each Other?" *Cognition* 255 (February): 106005. https://doi.org/10.1016/j.cognition.2024.106005.

Pfänder, Jan, Lou Kerzreho, and Hugo Mercier. 2024. "Quasi-Universal Acceptance of Basic Science in the US," December. https://doi.org/10.31219/osf.io/qc43v.

Pfänder, Jan, Sophie De Rouilhan, and Hugo Mercier. 2025. "Trusting but Forgetting Impressive Science." https://doi.org/10.31219/osf.io/argq5_v1.

Pillow, Bradford H. 1989. "Early Understanding of Perception as a Source of Knowledge."

Journal of Experimental Child Psychology 47 (1): 116–29. https://doi.org/10.1016/0022-0965(89)90066-0.

Popper, Karl R. 2002. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge Classics. London ; New York: Routledge.

Posit team. 2023. *RStudio: Integrated Development Environment for r*. Boston, MA: Posit Software, PBC. http://www.posit.co/.

Pummerer, Lotte, Lukas Fock, Kevin Winter, and Kai Sassenberg. 2024. "Conspiracy Beliefs and Majority Influence." *The Journal of Social Psychology*, September, 1–16. https://doi.org/10.1080/00224545.2024.2397491.

Pustejovsky, James E. 2019. "Simulating Correlated Standardized Mean Differences for Meta-Analysis." https://www.jepusto.com/simulating-correlated-smds/.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rainie, Lee, and Cary Funk. 2015. "Public and Scientists' Views on Science and Society." https://www.pewresearch.org/science/2015/01/29/public-and-scientists-views-on-science-and-society/.

Ranney, Michael Andrew, and Dav Clark. 2016. "Climate Change Conceptual Change: Scientific Information Can Transform Attitudes." *Topics in Cognitive Science* 8 (1): 49–75. https://doi.org/10.1111/tops.12187.

*Rathje, Steve, Jon Roozenbeek, Jay J. Van Bavel, and Sander Van Der Linden. 2023. "Accuracy and Social Motivations Shape Judgements of (Mis)information." *Nature Human Behaviour*, March. https://doi.org/10.1038/s41562-023-01540-w.

Rathje, Steve, Jon Roozenbeek, Jay J. Van Bavel, and Sander Van Der Linden. 2023. "Accuracy and Social Motivations Shape Judgements of (Mis)information." *Nature Human Behaviour*, March. https://doi.org/10.1038/s41562-023-01540-w.

Read, Stephen J., and Amy Marcus-Newhall. 1993. "Explanatory Coherence in Social Explanations: A Parallel Distributed Processing Account." *Journal of Personality and Social Psychology* 65 (3): 429–47. https://doi.org/10.1037/0022-3514.65.3.429.

Richardson, Emory, and Frank C. Keil. 2022. "Thinking Takes Time: Children Use Agents' Response Times to Infer the Source, Quality, and Complexity of Their Knowledge." *Cognition* 224 (July): 105073. https://doi.org/10.1016/j.cognition.2022.105073.

Romeijn, Jan-Willem, and David Atkinson. 2011. "Learning Juror Competence: A Generalized Condorcet Jury Theorem." *Politics, Philosophy & Economics* 10 (3): 237–62. https://doi.org/10.1177/1470594X10372317.

*Roozenbeek, Jon, Rakoen Maertens, Stefan M Herzog, Michael Geers, Ralf Kurvers, and Mubashir Sultan. 2022. "Susceptibility to Misinformation Is Consistent Across Question Framings and Response Modes and Better Explained by Myside Bias and Partisanship Than Analytical Thinking." *Judgment and Decision Making* 17 (3): 27.

Roozenbeek, Jon, Claudia R. Schneider, Sarah Dryhurst, John Kerr, Alexandra L. J. Freeman, Gabriel Recchia, Anne Marthe van der Bles, and Sander van der Linden. 2020. "Susceptibility to Misinformation about COVID-19 Around the World." *Royal Society Open Science* 7 (10): 201199. https://doi.org/10.1098/rsos.201199.

*Rosenzweig, Leah R., Bence Bago, Adam J. Berinsky, and David G. Rand. 2021. "Happiness and Surprise Are Associated with Worse Truth Discernment of COVID-19 Headlines Among Social Media Users in Nigeria." *Harvard Kennedy School Misinformation Review*, August. https://doi.org/10.37016/mr-2020-75.

*Ross, Björn, Jennifer Heisel, Anna-Katharina Jung, and Stefan Stieglitz. 2018. "Fake News on Social Media: The (In)Effectiveness of Warning Messages."

*Ross, Robert M, David G Rand, and Gordon Pennycook. 2021. "Beyond "Fake News": Analytic Thinking and the Detection of False and Hyperpartisan News Headlines." *Judgment and Decision Making* 16 (2): 22.

Ross Arguedas, A., Sumitra Badrinathan, Camila Mont'Alverne, Benjamin Toff, Richard Fletcher, and R. Nielsen. 2022. "Snap Judgements: How Audiences Who Lack Trust in News Navigate Information on Digital Platforms."

Rousseau, Denise M., Sim B. Sitkin, Ronald S. Burt, and Colin Camerer. 1998. "Introduction to Special Topic Forum: Not so Different After All: A Cross-Discipline View of Trust." *The Academy of Management Review* 23 (3): 393–404. http://www.jstor.org/stable/259285.

Rutjens, Bastiaan T., and Bojana Većkalov. 2022. "Conspiracy Beliefs and Science Rejection." *Current Opinion in Psychology* 46 (August): 101392. https://doi.org/10.1016/j.copsyc.2022.101392.

Salmon, Daniel A., Matthew Z. Dudley, Jason M. Glanz, and Saad B. Omer. 2015. "Vaccine Hesitancy: Causes, Consequences, and a Call to Action." *Vaccine* 33: D66–71.

Scharff, Darcell P., Katherine J. Mathews, Pamela Jackson, Jonathan Hoffsuemmer, Emeobong Martin, and Dorothy Edwards. 2010. "More Than Tuskegee: Understanding Mistrust about Research Participation." *Journal of Health Care for the Poor and Underserved* 21 (3): 879–97. https://doi.org/10.1353/hpu.0.0323.

Scheufele, Dietram A. 2022. "Thirty Years of Science–society Interfaces: What's Next?" *Public Understanding of Science* 31 (3): 297–304. https://doi.org/10.1177/09636625221075947.

Scheufele, Dietram A., and Nicole M. Krause. 2019. "Science Audiences, Misinformation, and Fake News." *Proceedings of the National Academy of Sciences* 116 (16): 7662–69. https://doi.org/10.1073/pnas.1805871115.

Schulz, Anne, Richard Fletcher, and Marina Popescu. 2020. "Are News Outlets Viewed in the Same Way by Experts and the Public? A Comparison Across 23 European Countries." *Reuters Institute for the Study of Journalism*.

Shirikov, Anton. 2024. "Fake News for All: How Citizens Discern Disinformation in Autocracies." *Political Communication* 41 (1): 4565. https://doi.org/10.1080/10584609.2023.2257618.

*Smelter, Thomas J., and Dustin P. Calvillo. 2020. "Pictures and Repeated Exposure Increase Perceived Accuracy of News Headlines." *Applied Cognitive Psychology* 34 (5): 1061–71. https://doi.org/10.1002/acp.3684.

Smith, E. Keith, M. Julia Bognar, and Adam P. Mayer. 2024. "Polarisation of Climate and Environmental Attitudes in the United States, 1973-2022." *Npj Climate Action* 3 (1): 1–14. https://doi.org/10.1038/s44168-023-00074-1.

Smith, John Maynard, and David Harper. 2003. *Animal Signals*. Oxford University Press.

Smith, Tom W., and Jaesok Son. 2013. "Trends in Public Attitudes about Confidence in Institutions." *General Social Survey Final Report*.

Soll, Jack B., and Richard P. Larrick. 2009. "Strategies for Revising Judgment: How (and How Well) People Use Others' Opinions." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35 (3): 780–805. https://doi.org/10.1037/a0015145.

Song, Hyunjin, David M Markowitz, and Samuel Hardman Taylor. 2022. "Trusting on the Shoulders of Open Giants? Open Science Increases Trust in Science for the Public and Academics." *Journal of Communication* 72 (4): 497–510. https://doi.org/10.1093/joc/jqac017.

Sperber, Dan, Fabrice Clément, Christophe Heintz, Olivier Mascaro, Hugo Mercier, Gloria Origgi, and Deirdre Wilson. 2010. "Epistemic Vigilance." *Mind & Language* 25 (4): 359–93.

*Stagnaro, Michael, Sophia Pink, David G. Rand, and Robb Willer. 2023. "Increasing Accuracy Motivations Using Moral Reframing Does Not Reduce Republicans' Belief in False News." *Harvard Kennedy School Misinformation Review*, November. https://doi.org/10.37016/mr-2020-128.

Stagnaro, Michael N., Ben M. Tappin, and David G. Rand. 2023. "No Association Between Numerical Ability and Politically Motivated Reasoning in a Large US Probability Sample." *Proceedings of the National Academy of Sciences* 120 (32): e2301491120. https://doi.org/10.1073/pnas.2301491120.

Stockemer, Daniel, and Jean-Nicolas Bordeleau. 2024. "Understanding Climate Change Conspiracy Beliefs: A Comparative Outlook." *Harvard Kennedy School Misinformation Review* 5 (6). https://doi.org/10.37016/mr-2020-162.

Street, Chris N. H., and Jaume Masip. 2015. "The Source of the Truth Bias: Heuristic Processing?" *Scandinavian Journal of Psychology* 56 (3): 254–63. https://doi.org/10.1111/sjop.12204.

Strevens, Michael. 2020. *The Knowledge Machine: How Irrationality Created Modern Science*. New York: Liveright Publishing Corporation.

Sturgis, Patrick, and Nick Allum. 2004. "Science in Society: Re-Evaluating the Deficit Model of Public Attitudes." *Public Understanding of Science* 13 (1): 55–74. https://doi.org/10.1177/0963662504042690.

Sturgis, Patrick, Ian Brunton-Smith, and Jonathan Jackson. 2021. "Trust in Science, Social Consensus and Vaccine Confidence." *Nature Human Behaviour* 5 (11): 1528–34. https://doi.org/10.1038/s41562-021-01115-7.

*Sultan, Mubashir, Alan N. Tump, Michael Geers, Philipp Lorenz-Spreen, Stefan M. Herzog, and Ralf H. J. M. Kurvers. 2022. "Time Pressure Reduces Misinformation Discrimination Ability but Does Not Alter Response Bias." *Scientific Reports* 12 (1): 22416. https://doi.org/10.1038/s41598-022-26209-8.

Tappin, Ben M., Gordon Pennycook, and David G. Rand. 2020. "Bayesian or Biased? Analytic Thinking and Political Belief Updating." *Cognition* 204 (November): 104375. https://doi.org/10.1016/j.cognition.2020.104375.

Ten Hove, Debby, Terrence D. Jorgensen, and L. Andries van der Ark. 2024. "Updated guidelines on selecting an intraclass correlation coefficient for interrater reliability, with applications to incomplete observational designs." *Psychological Methods* 29 (5): 967–79. https://doi.org/10.1037/met0000516.

Trouche, Emmanuel, Petter Johansson, Lars Hall, and Hugo Mercier. 2018. "Vigilant Conservatism in Evaluating Communicated Information." Edited by Alexander N. Sokolov. *PLOS ONE* 13 (1): e0188825. https://doi.org/10.1371/journal.pone.0188825.

Ulusoy, Ezgi, Dustin Carnahan, Daniel E. Bergan, Rachel C. Barry, Siyuan Ma, Suhwoo Ahn, and Johnny McGraw. 2021. "Flooding the Zone: How Exposure to Implausible Statements Shapes Subsequent Belief Judgments." *International Journal of Public Opinion Research* 33 (4): 856–72.

Uscinski, Joseph, Adam Enders, Casey Klofstad, Michelle Seelig, Hugo Drochon, Kamal Premaratne, and Manohar Murthi. 2022. "Have Beliefs in Conspiracy Theories Increased over Time?" *PLoS ONE* 17 (7): e0270429. https://doi.org/10.1371/journal.pone.0270429.

Van Duyn, Emily, and Jessica Collier. 2019. "Priming and Fake News: The Effects of Elite Discourse on Evaluations of News Media." *Mass Communication and Society* 22 (1): 29–48. https://doi.org/10.1080/15205436.2018.1511807.

Van Stekelenburg, Aart, Gabi Schaap, Harm Veling, Jonathan Van 'T Riet, and Moniek Buijzen. 2022. "Scientific-Consensus Communication About Contested Science: A Preregistered Meta-Analysis." *Psychological Science* 33 (12): 1989–2008. https://doi.org/10.1177/09567976221083219.

Vaupotič, Nina, Dorothe Kienhues, and Regina Jucks. 2021. "Trust in Science and Scientists: Implications for (Higher) Education." In, edited by Bernd Blöbaum, 207–20. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-72945-5_10.

Većkalov, Bojana, Sandra J. Geiger, František Bartoš, Mathew P. White, Bastiaan T. Rutjens, Frenk van Harreveld, Federica Stablum, et al. 2024. "A 27-Country Test of Communicating the Scientific Consensus on Climate Change." *Nature Human Behaviour*, August, 1–14. https://doi.org/10.1038/s41562-024-01928-2.

Viechtbauer, Wolfgang. 2010. "Conducting Meta-Analyses in *r* with the **Metafor** Package." *Journal of Statistical Software* 36 (3). https://doi.org/10.18637/jss.v036.i03.

Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. "The Spread of True and False News Online." *Science* 359 (6380): 1146–51. https://doi.org/10.1126/science.aap9559.

Vranic, Andrea, Ivana Hromatko, and Mirjana Tonković. 2022. ""I Did My Own Research": Overconfidence, (Dis)trust in Science, and Endorsement of Conspiracy Theories." *Frontiers in Psychology* 13 (July). https://doi.org/10.3389/fpsyg.2022.931865.

Weingart, Peter, and Lars Guenther. 2016. "Science Communication and the Issue of Trust." *Journal of Science Communication* 15 (05): C01. https://doi.org/10.22323/2.15050301.

Weisberg, Deena Skolnick, Asheley R. Landrum, Jesse Hamilton, and Michael Weisberg. 2021. "Knowledge about the Nature of Science Increases Public Acceptance of Science Regardless of Identity Factors." *Public Understanding of Science* 30 (2): 120–38. https://doi.org/10.1177/0963662520977700.

Wellcome Global Monitor. 2018. "Wellcome Global Monitor 2018." https://wellcome.org/reports/wellcome-global-monitor/2018.

———. 2020. "Wellcome Global Monitor 2020: Covid-19." https://wellcome.org/reports/wellcome-global-monitor-covid-19/2020.

———. 2021. "Public Trust in Scientists Rose During the Covid-19 Pandemic | News." https://wellcome.org/news/public-trust-scientists-rose-during-covid-19-pandemic.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wilholt, Torsten. 2013. "Epistemic Trust in Science." *The British Journal for the Philosophy of Science* 64 (2): 233–53. https://doi.org/10.1093/bjps/axs007.

*Winter, Stephan, Sebastián Valenzuela, Marcelo Luis Barbosa Santos, Tobias Schreyer, Lena Iwertowski, and Tobias Rothmund. n.d. "(Don't) Stop Believing: A Signal Detection Approach to Risk and Protective Factors for Engagement with Politicized (Mis)information in Social Media." https://doi.org/10.31234/osf.io/84c36.

Wintterlin, Florian, Friederike Hendriks, Niels G. Mede, Rainer Bromme, Julia Metag, and Mike S. Schäfer. 2022. "Predicting Public Trust in Science: The Role of Basic Orientations Toward Science, Perceived Trustworthiness of Scientists, and Experiences with Science." *Frontiers in Communication* 6 (January). https://doi.org/10.3389/fcomm.2021.822757.

Xie, Belinda, and Brett Hayes. 2022. "Sensitivity to Evidential Dependencies in Judgments Under Uncertainty." *Cognitive Science* 46 (5). https://doi.org/10.1111/cogs.13144.

Yaniv, Ilan. 1997. "Weighting and Trimming: Heuristics for Aggregating Judgments Under Uncertainty," 13.

———. 2004. "Receiving Other People's Advice: Influence and Benefit." *Organizational Behavior and Human Decision Processes* 93 (1): 1–13. https://doi.org/10.1016/j.obhdp.2003.08.002.

Yaniv, Ilan, Shoham Choshen-Hillel, and Maxim Milyavsky. 2009. "Spurious Consensus and Opinion Revision: Why Might People Be More Confident in Their Less Accurate Judgments?" *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35 (2): 558–63. https://doi.org/10.1037/a0014589.

Yaniv, Ilan, and Eli Kleinberger. 2000. "Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation." *Organizational Behavior and Human Decision Processes* 83 (2): 260–81. https://doi.org/10.1006/obhd.2000.2909.

Yin, Jun, Zikai Xu, Jing Lin, Wenying Zhou, and Xiuyan Guo. 2024. "Smartly Following Others: Majority Influence Depends on How the Majority Behavior Is Formed." *Journal of Experimental Social Psychology* 115 (November): 104644. https://doi.org/10.1016/j.jesp.2024.104644.

Yousif, Sami R., Rosie Aboody, and Frank C. Keil. 2019. "The Illusion of Consensus: A Failure to Distinguish Between True and False Consensus." *Psychological Science* 30 (8): 1195–1204. https://doi.org/10.1177/0956797619856844.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge university press.