On the Origins of Polarized Risk Assessment

Zachary Balleisen

Advised by Dan Kahan

Elizabeth K. Dollard Professor of Law

Submitted to the faculty of Cognitive Science in partial fulfillment of the requirements for the degree of Bachelor of Science

Yale University

April 22, 2019

Abstract

This thesis attempts to examine how risk assessments become politically polarized. It addresses four major research questions through two studies, set against the backdrop of artificial intelligence developments. First, it considers whether emerging questions of technological risk have inherent partisan valences. Secondly, it seeks to evaluate the effectiveness of internet comments as a form of social proof. Third, it attempts to determine the extent to which perceptions of an issue's partisan valence influenced object-level beliefs about that issue. And fourth, it hopes to distinguish between having polarized preexisting beliefs and engaging in biased evidence interpretation. It also aims to identify sufficient conditions for biased evidence processing to emerge. I found that internet comments can effectively be used to change the perceived political valence of an issue. However, my other results were inconclusive due to insufficiently strong experimental manipulation. The thesis concludes with proposals for further research, as well as a discussion about how psychological research can contribute to rational democratic discourse.

Table of Contents

Introduction

This thesis is motivated by four questions about emerging risk perception polarization. First, it attempts to answer whether emerging questions of technological risk have inherent partisan orientation, or "valence." Public opinion about many scientific issues, like climate change, exhibits extreme political polarization. Do these issues have inherent ideological valences that determine how existing political coalitions react to them, or are they largely indeterminate, influenced more strongly by the framing of "early belief adopters" and campaigners than by features inherent in the issues themselves?

Secondly, the thesis seeks to assess the effectiveness of internet comments as a form of social proof. To what extent does casually reading comments involve not only absorbing what people say, but also the partisan affiliations of the people saying it? How much exposure is necessary to influence risk perceptions and policy preferences? Are the comments on a single article enough?

Third, the thesis examines the extent to which perceptions of a given issue's partisan valence influenced object-level beliefs about that issue. If people become aware that an issue is polarized, will their views shift toward the perspective of people whom they see as like them? If such influence occurs, does it do so because individuals are directly adopting the views of like-minded people, because they are receptive to the same arguments as co-partisans, or something in between?

And fourth, the thesis hopes determine whether there is a distinction between having polarized preexisting beliefs and engaging in biased evidence interpretation. If partisan internet comments influence the have a polarizing impact on readers views, will that polarization bias their interpretation of any future evidence that they encounter? Does the tone of the original discussion affect their openness to new, countervailing evidence?

This thesis attempts to shed light on these questions with two studies. The introduction section describes prior work in the field, laying out the concepts that underlie its design. The first section discusses how humans form risk perceptions, and why views about risk provide a good backdrop for polarization research. The second section discusses how political polarization affects judgements about risk. The third section discusses public opinion about advances in the field of artificial intelligence, and argues that it is an ideal issue for studying how risk assessment becomes politically polarized, if and when it comes to have that characteristic.

Risk Perception

Risk perceptions play a crucial role in decisions about public policy. Explicitly or implicitly, almost all questions of public policy require balancing the uncertain benefits and costs of possible courses of action. To make good decisions, the public and public officials must not only have a good general knowledge of the risks they face in daily life; they must also be able to update their beliefs and reasoned evaluations in light of new evidence. Even if people initially disagree about the degree of risk, rationale people presumably should respond to new information in a similar way.

This process can be modeled as “Bayesian inference,” a technique for optimally updating uncertain beliefs in light of new information. There is strong evidence that humans use approximate Bayesian processing to learn new concepts, identify probabilities, and infer the beliefs and goals of other actors (Tenenbaum, 1999).

A bevy of studies, however, have shown that many individuals do not perfectly follow Bayesian principles. When evaluating risks, people often employ heuristics, taking shortcuts that make calculations tractable, and are frequently useful. However, these approaches fail under certain conditions. People, for example. tend to exhibit confirmation bias, and often stick with preexisting judgments more than the evidence warrants (Kahneman, 2011).

Many individuals also struggle with assessing the significance of numerical evidence about risk. Thus many people consider only the reliability of a medical test, and not the frequency of the underlying disease. This phenomenon is the source of many false positives and unnecessary procedures in the medical industry (Rottman, Prochaska, & Deaño, 2016), and is known in the psychology literature as “base-rate neglect” (Kutzner, Freytag, Vogel, & Fiedler, 2008).

An even simpler counterintuitive task is “covariance-detection,” or the ability to think proportionally instead of picking the largest (or smallest) value, regardless of any fractional denominator. If presented with the data in Table 1 and asked whether a given health treatment is effective, many individuals use a heuristic that centers on the largest number. First they will check if more people in the “Control” group got better than stayed the same. Next they will check whether more people got better in the “Control” group than in the “Treatment” group. If both of those conditions prove true, they will conclude that the control worked better than the treatment, and if not they will conclude the opposite.

In this case, people following this heuristic would conclude that the control worked better than the treatment, and that the treatment was ineffective. However, the correct answer to this problem is that the treatment appears to have a positive benefit, though one would definitely want a larger sample size to be certain. The covariance detection task is relatively difficult for the general public, and over half of all people employ the heuristic rather than carry out the appropriate calculations (Kahan, Peters, Dawson, & Slovic, 2017; Kahan, Peters, et al., 2012).

	Gets better	Stayed the same
Control	205	73
Treatment	95	21

Table 1. An illustration of a covariance detection task.

One must keep in mind, however, that vulnerability to heuristic reasoning is not uniform, and that some individuals are particularly good at discounting unproductive heuristics and coherently updating their beliefs (Toplak, West, & Stanovich, 2011). Psychologists have proposed a variety of explanations for this variation. Some scientists argue that a high resistance to heuristic errors is associated with a penchant for self-reflection. If individuals are inclined to override their unconscious, implicit mental processing and think deliberatively, it would seem to follow that they will rely less on faulty heuristics. The strength of this inclination can be measured with the cognitive reflection test, which tests how well subjects solve counter-intuitive puzzles (Frederick, 2005).

Another potential explanation lies with open-mindedness. Individuals who readily acknowledge that they may be wrong might be less sure of themselves after an initial evaluation, and therefore more open to considering the significance of new evidence. The Actively Open Minded Thinking scale attempts to measure this willingness by asking a series of self-evaluation questions and producing a composite score (Stanovich & West, 1997).

Finally, another hypothesis is that facility with numbers and numerical reasoning allows people to understand and process risk data more effectively. According to this hypothesis, people with a strong math background are less likely to be confused by the covariance detection task, and they will have less internal leeway to manipulate the evidence. Numeracy can be measured using the Abbreviated Numeracy Scale (Weller et al., 2013).

In a 2013 experiment, researchers asked participants to interpret the results from a simulated study (Kahan, Peters, et al., 2017). The study results were encapsulated in a 2x2 covariance detection table, similar to the one seen in Table 1. Participants were randomly assigned to one of two experimental groups. In the first, they were asked to interpret the simulated study results about the efficacy of a skin cream. (The implications of these "study results" were randomized between participants.) In the second, they were asked to evaluate the results of a gun control measure. The two conditions are shown in Table 2.

To find the best option, participants would have to compare the proportion of situations that improved for both the treatment and the control. However, participants could also take a heuristic shortcut and merely look at the largest number. (The fundamental mistake here is assuming that there are an equal number of situations in each condition.)

	Rash Got Worse	Rash Got Better
Used Skin Cream	223	75
Did Not Use Skin Cream	107	21

	Crime Got Worse	Crime Got Better
Banned Concealed Carry	223	75
Did Not Ban Concealed Carry	107	21

Table 2.The contingency tables from the two experimental groups. Subjects were asked which policy was more successful. Here, the evidence supports using the skin cream and banning concealed carry, but the numbers were also counterbalanced in the other direction (Kahan, Peters, et al., 2017).

On most study topics, such as those investigating whether a skin cream effectively treated a rash, high CRT scores, active open-minded thinking, and numeracy, were all associated with improved covariance detection. This ability to correctly evaluate and interpret numerical evidence did not vary based on the person's preexisting views (Kahan, Peters, et al., 2017).

The three measures are highly correlated with each other. However, the numeracy measurement consistently explained the highest proportion of the variance in covariance detection. Additionally, the numeracy measurement had the best psychometric properties, including a normal distribution and effective discrimination at all capacity levels (Kahan, Peters, et al., 2017).[1]

Political Polarization

The relationship between numeracy and effective evidentiary updating, however, breaks down in contexts where risk assessments are closely linked to a person's self-conception and identification. For example, in the same study, participants were asked to interpret a study on whether a gun control policy increased or decreased crime. As shown in Table 2, the contingency tables in the two situations had the same quantitative data, and they should have been answered in the same way.

However, participants in the gun control condition behaved very differently from participants in the skin cream condition. When the study's conclusions agreed with subjects' political predispositions about gun control, those individuals correctly interpreted the evidence, and higher numeracy correlated with more correct interpretations. When the presented evidence led to conclusions that differed from subjects' preexisting assumptions and political beliefs, highly numerate individuals performed no better than their less numerate peers (Kahan, Peters, et al., 2017).

In the skin cream study, there is a clear, positive relationship between an individual's numeracy score and their performance on the covariance detection task. In the gun control condition, this relationship holds when the evidence supports their identity group's preexisting beliefs. (For Democrats,

[1] In contrast, the Cognitive Reflection Test can only identify the gradations that separate somewhat reflective people from very highly reflective people. A large portion of test takers get no questions correct (Primi, Morsanyi, Donati, & Chiesi, 2014). Alternative measures have been proposed to rectify this issue.

this pattern would mean that the evidence provided showed that gun control decreases crime, while for Republicans, it would mean that the evidence provided showed that gun control increases crime.) However, when the evidence disagreed with their identity-laden views, high numeracy individuals make the covariance detection mistakes at rates similar to low numeracy individuals.

The same trends emerged when one replaces the independent variable with cognitive reflection or active open-minded thinking, rather than numeracy.[2] Crucially, these results have proved to be politically symmetrical. Liberals and conservatives of a given numeracy level displayed equal amounts of motivated, identity-protective thought reasoning. These general results have since been replicated (Kahan & Peters, 2017).

A number of studies have further demonstrated the applicability of this pattern to reasoning about climate change, which researchers have similarly shown to be susceptible to politically polarized risk assessments. This tendency is especially strong when climate change discourse is linked to other cultural values, like belief in free markets. (Cook & Lewandowsky, 2016; Kahan, Jenkins-Smith, Tarantola, Silva, & Braman, 2012; Whitmarsh, 2011).

These findings support the view that when societies become polarized, the resulting divide can extend to areas of life that are far removed from the original central disagreement. Different groups come to inhabit different neighborhoods, consume different products, follow different media, and attend different religious institutions (Green, 2008; Mitchell, Matsa, Gottfried, & Kiley, 2014; Mittal, Malshe, & Sridhar, 2018; Nivola, 2008). Ensconced in different communities, polarized groups develop different understanding of facts, common sense, and risk.

One must be careful, though, to distinguish between the existence of “polarized preexisting beliefs,” on the one hand, and the dynamic of “polarized barriers to belief revisions,” on the other. The former corresponds to a Bayesian prior, and the latter to a Bayesian likelihood ratio. People with different

---

[2] It’s quite ironic that supposedly open-minded thinkers are actually more susceptible motivated reasoning in identity-charged cases. This throws the external validity of the actively open-minded measurement into question.

backgrounds have different initial senses of what is plausible and what is likely, what is safe and what is risky. These differences are a source of healthy debate, and they consistent with sound Bayesian reasoning practices. However, when people fail to update their beliefs when evidence contradicts them, their beliefs about risk levels will no longer converge towards the actual risk levels.

Psychologists lack a comprehensive understanding of how and why some issues become polarized in the first place. Some technological determinists (Postman, 2011) posit that the inherent structure of an emerging issue aligns with fault lines in a society. This correlation makes discussing it inherently “controversial” and indicates what “sides” citizens should take on the issue. In this view, societal conflicts are driven by structural forces, and pragmatic concerns about technological risk are subsumed into the society’s larger structures of ideological framing.

In contrast, other social scientists claim that political controversies are not determined by factors inherent in a particular issue or technology; rather, they are framed by the small groups of elite opinion shapers who happen to stumble upon an opinion first. In this view, the emergence of polarization is a stochastic process. Whether an issue becomes controversial and even the positions of the two sides depend in large part on how the first people considering the new technology or issue frame the risks involved.

Individuals do appear to approach policy issues differently depending on how they first encounter the issues. For example, Americans are more likely to favor allowing the Klu Klux Klan to march in public when the vignette describing the scenario emphasizes free speech instead of public order (Nelson, Clawson, & Oxley, 1997). Framing does not erase citizens’ perspectives, but it does help define what sort of values and questions become most salient to a debate (Brewer & Gross, 2005).

One should keep in mind that weaker versions of these theories are compatible, and that it might be the case that the true cause of risk polarization lies somewhere in-between the most starkly framed hypotheses. Certain issues or technological advancements may be primed to cause political disagreements, and the social implications of a given technology may influence how preexisting political identities align around the question. However, this preexisting lean leaves many undetermined degrees of freedom, and the way the technology is framed by politicians, entrepreneurs, and intellectuals may also have a strong impact on how public opinion around it develops.

At present, we also lack a rigorous way to classify potential types of opinion polarization. Different identity groups may have different preexisting beliefs, or "priors," about a technology. For instance, there is a strong partisan prior about whether the government should require drivers to wear seatbelts. Most conservatives believe taking the risk of not wearing a seatbelt should be a personal choice, and have a strong prior belief against such laws. Some have argued that such a mandate will just cause risker driving behavior (Rock, 1993).[3] Most liberals have a more neutral to positive prior with regard to mandatory seatbelt laws.

However, this variation does not mean that in all cases a person's identity will reinforce maintenance of prior beliefs regardless of new information, nor that different identity groups cannot converge on a risk assessment as more evidence becomes available.[4] Despite having different prior beliefs, individuals across the political spectrum are responsive to scientific evidence about seatbelt risk, and converge to similar beliefs as the obtain more information (Schenck, Runyan, & Earp, 1985).

With certain controversial topics, however, the interpretation of evidence itself becomes polarized. This dynamic can poison public deliberations and slowly divorce risk assessments from reality. Social scientists have not been able to determine how and why ordinary disputes move into this category, or how anyone can free them from polarization once a negative cycle of evidentiary polarization has started.

Under the stochastic conception, ideas are originally framed by the choices of elite members of an identity group before radiating out through the community via social proof. This way of thinking

[3] This presumed effect is known as risk compensation. It is premised on the idea that people have a risk preference, and reducing the risk in an environment will induce more risky behavior. A risk compensation effect has been observed empirically in some contexts, but there is disagreement about whether seatbelt laws induced such an effect (Schenck, Runyan, & Earp, 1985). There is strong evidence that mandatory seatbelt laws have made road transportation dramatically safer overall (National Highway Traffic Safety Administration (NHTSA), 2009).

[4] Confirmation bias may slow this convergence. However, this bias is not linked to identity, and can be overcome with a sufficiently strong body of evidence.

presumes that all humans are extremely sensitive to the perspectives of the people around them, and that people also track whether others identify with them or not. Thus on this view, minds do not simple track ideas; they also track who expresses them, why, and whom might disagree.

There are many ways social proof can spread. The oldest form comes through direct verbal conversation; points made during arguments often have less impact on one's debate partner than on the observing audience. People excel at tracking viewpoints and disagreements on television (Gamson & Modigliani, 1989; Nelson et al., 1997), and they quickly file away reporting about adversarial viewpoint clashes that they read in news articles.[5] Presumably, people can also automatically identify fault lines and differing viewpoints when reading other forms of text-based communication, like online article comment sections.

However, text-based communication also strips out a lot of useful information that people would glean from in-person interactions. Intonation helps listeners discern the intention, sincerity, and even meaning[6] of a statement, and body language further assists by indicating group affiliations and individual attitudes. In text-based disagreements, it can be hard to determine whether an interlocutor is earnest or jocular, authentic or contrived. This can lead to cycles of escalating uncivil behavior.

The relative paucity of information provided by a text conversation makes it much easier to manipulate opinion. Millions of Americans were unwittingly exposed to Russian sock puppet accounts from the Internet Agency during the 2016 election (Office of the Director of National Intelligence, 2017), most of whom successfully posed as angry American partisans. Pranksters on the websites 4chan and 8chan frequently plan "raids" to manipulate opinion on other websites, and online marketplaces sell astroturfing services with similar aims (Marwick & Lewis, 2017).

[5] Readers tend to find adversarial stories with a clear clash in viewpoints far more engaging, and journalists are encouraged to fit their stories to this mold (Huang, 2005).
[6] For example, a rising pitch at the end of a sentence can be used to express uncertainty. Conveying this meaning in text requires a more explicit treatment (Ching, 1982).

Political Perspectives on Artificial Intelligence

One emerging technological question that seems particularly apt for this type of social analysis is artificial intelligence (AI). Most Americans have heard of artificial intelligence; however, a majority do not understand how modern AI systems work or what they are capable of achieving.

Moderate public awareness of the AI field has existed for decades, but it has never been a top tier political issue, and has received attention more in the realm of science fiction than political debate. In the 1980s, public commentary associated artificial intelligence with so-called "expert systems," where domain-specific knowledge and hand-crafted rules were used to improve business processes. In the 1990s, AI became strongly associated with chess, as the contest between Garry Kasparov and IBM's Deep Blue computer captivated the world. In the 2000s, public focus shifted to semantic search and automated translation, and in the 2010s public views of "AI" centered on personal assistants like Amazon Alexa and Google Home (Fast & Horvitz, 2016).

Despite the periodic shifts in the way the public conceptualizes artificial intelligence, the public's general impression of AI stayed relatively constant for most of this period. According to a linguistic corpus analysis of popular news media, references to AI stayed relatively constant for decades, with positive, optimistic articles slightly outpacing more skeptical, negative ones (Fast & Horvitz, 2016).

However, an inflection point occurred in 2009, when media references to artificial intelligence and related concepts began to grow (Fast & Horvitz, 2016). This heightened coverage coincided with several developments in AI research, including the reemergence of deep neural network architectures powered by GPU acceleration, which has in turn sparked an explosion in automation across the economy (ITU, 2018). In fact, AI systems are playing an increasingly large role in the media industry itself, writing articles (Peiser, 2019) and monitoring comment sections for abuse (Mullin, 2016).

The increasing prominence of AI in the economy has led to a few public opinion polls, though most of the published research on the public's general attitudes have methodological drawbacks. The terms pollsters have used in questions are often vague, and different phrasings produce have large impacts on the results. There is, moreover, little to no public polling on AI topics that has political affiliation crosstabs (2016 CBS Interactive Inc, 2016; West, 2018a, 2018b).

Despite the large and growing impact of AI on Americans' way of life, artificial intelligence has not (yet) become a partisan political issue. AI proponents and skeptics[7] can be found across the political spectrum. In 2017, Senators Maria Cantwell, a Democratic from Washington State, Todd Young, a Republican from Indiana, and Ed Markey, a Massachusetts Democrat, proposed the FUTURE of AI Act, a bill to accelerate the development of AI technology. Cantwell struck a positive tone announcing the bill, saying ""'We expect that artificial intelligence will be an incredibly transformative force for growth and productivity. We need to be ready for it." Young was just as positive, saying "Artificial Intelligence has the ability to drastically boost our economy" (Cantwell, 2017).

Other senators have struck a more skeptical stance about at least some aspects of the technology. Roger Wicker, a Republican from Mississippi, has discussed the need to ensure that "decisions made by AI systems are based on representative data that does not unintentionally harm vulnerable populations or act in an unsafe anticompetitive or biased way." Brian Schatz, a Hawaii Democrat, agreed with this sentiment, and suggested that he was uncomfortable that neural networks make decisions in a "black box" (Vogel-Fox, 2017).

This collection of attributes makes artificial intelligence a good subject for research into polarization formation. It is a well-known technology, but not so well known that research subjects would have a strong foundation of information independent of the research itself. It is easy to construct plausible arguments that favor or disapprove of artificial intelligence advancements from a variety of political perspectives. This malleability facilitates the construction of study elements that incorporate social proof manipulations.

---

[7] Many people believe that AI developments have the capacity to benefit or harm society. In this thesis, I classify a position as optimistic if a person believes that the expected result of our current trajectory will produce a net benefit for society overall. I regard a position as pessimistic or skeptical if they believe that future AI developments will harm society overall. People who wish to stringently regulate AI instead of banning it outright may still be skeptics.

An alternative approach would be to craft a narrative about a completely fake emerging technology. This tack would ensure that all participants would start out without background opinions on the subject, potentially easing the task of creating manipulations. However, describing a credible new technology could be difficult, and the resulting experiment would be less immediately transferable to the real world. These limitations could reduce the experiment's external validity.

For this thesis, I have analyzed relevant results from an existing pilot study, as well as conducted an entirely new study myself that builds on the findings from the pilot study.

Methods and Results

Study 1

Introduction

The first survey was a multipurpose pilot study conducted by the Cultural Cognition Lab with support from the Annenberg Public Policy Center. The study instrument included a CRT, an Actively Open Minded Thinking measure, and several risk assessment questions. While doing exploratory analysis on this dataset I realized that it contained data on the public's perception about several emerging technological risks, including artificial intelligence, driverless cars, and human genomic editing. I used this data to help develop a second study on the emergence of political polarization.

Participants

The initial study was conducted on YouGov, a public opinion and data firm. Participants were drawn from a representative public panel; YouGov oversampled and the selected a subset of data to ensure a reasonable demographic balance. The research team then received data on 2500 subjects.

Due do the oversampling procedure, the sample's demographic makeup roughly mirrored the US's voting age population. The mean age was 51.4, with a standard deviation of 17.2. The sample was 56.2% female, and had representative education levels. Some participants did not complete all the questions necessary for certain statistical tests, so they were excluded from those analyses.

Designs

Participants were first asked for their opinions on four policies: gun control, AI regulation, affirmative action, and bans on human genetic editing. The questions were randomized, and responses from “strongly oppose” to “strongly support” were indicated on a six point Likert scale. Subjects were then asked to rate the severity of a series of risks, including the risks entailed by self-improving artificial intelligence, the risks caused by allowing driverless cars on public roads, and the risks involved in human genomic editing. The risk questions were asked in randomized order, and responses were given on an eight point Likert scale from “no risk at all” to “very high risk.”

Participants then completed a science curiosity evaluation (Kahan, Landrum, Carpenter, Helft, & Jamieson, 2017). The completed metric included questions on news interests, leisure actives, book genres, conversation topics, and social media posts. It also asked participants to self-report their interests, and to select an article to read during the study. Each question asked about the person’s interest in science-related materials and conversations, as well a series of distractor questions on other topics like religion and sports.

At the end of the survey, participants were asked to complete a series of psychological measures presented in a randomized order. First, participants were given the Cognitive Reflection Test, a three question evaluation which tests one’s ability to make counterintuitive judgments. This test attempts to assess a person’s relative tendency to engage in slow deliberative reasoning or fast intuitive reasoning (Frederick, 2005).

Participants also completed the Need for Cognition scale, which tries to measure one’s willingness to engage in effortful thinking or information processing (Cacioppo & Petty, 1982). Participants also took an Actively Open-Minded Thinking questionnaire, a self-report survey which attempts to quantify how open one is to new ideas and evidence (Stanovich & West, 1997).

In order to evaluate scientific knowledge and reasoning ability, subjects took an Ordinary Science Intelligence assessment (version 2.0), which measures one’s ability to recognize scientific evidence and apply it to everyday decisions (Kahan, 2017). They were also asked a question from the National Science Foundation's Science and Engineering Indicator about their views on science.

Participants also reported their partisan affiliation, as well as their personal ideology from very liberal to very conservative. These two variable were normalized, added, and normalized again, creating a composite measure of each subject's political views.

Results

Participants expressed significant concerns about AI risks. Their mean risk evaluation was 5.45 out of 8, signifying that the average subject believed that the "development of 'artificially intelligent' computers capable of reprogramming themselves based on automated information search and experience" presents a "moderate to high" risk.

In contrast, participants were slightly more worried about driverless cars or human genome modification. "Allowing driverless cars to operate on public roads" had a mean risk evaluation of 5.73, and "human genome modification" had a mean risk evaluation of 5.88. These results indicate that the average participant believes that driverless cars and human genome modification present a "high" risk. However, participants' risk perception judgements had a lot of variability, as showed in Figure 1. The effective difference between the distributions is relatively small. Additionally, they are all right skewed, which is probably an artifact of the measurement scale.

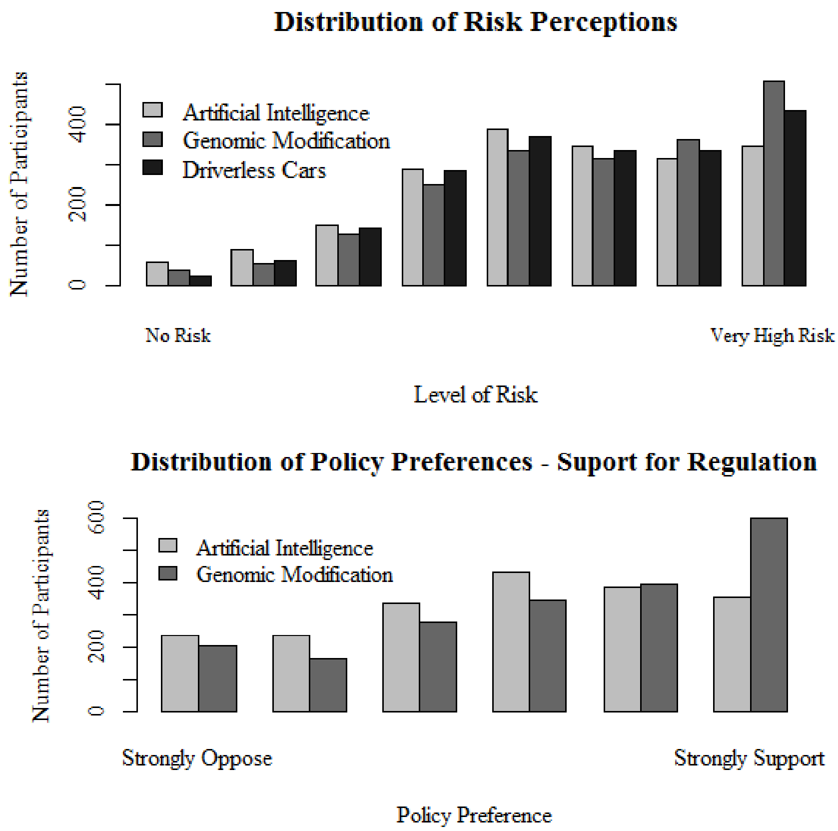


Figure 1. Distribution of risk perceptions and policy preferences for the assessed topics.

The study also indicated a heterogeneity of opinions about regulatory policy. Figure 1 shows that levels of support for stronger AI regulation, in particular, is fairly evenly distributed. In contrast, support for regulations on human genomic modification is very right-skewed.

People who believe AI is risky are more inclined to regulate it, though a linear model still only explains 2.9% of the data's variance. People who have extremely low or high evaluations of AI risk have notably heterogeneous views on AI regulation, which reduces the explanatory power of a linear model. The correlation between AI risk and AI policy preferences is weaker than the correlation between genomic modification risk and genomic modification policy preferences, suggesting that people have not given much thought to AI risks.

To generate a composite measure of each participant's political perspective, their ideological self-evaluation (from very liberal to very conservative) was added to their political affiliation (from strong

Democrat to strong Republican.) The result was then normalized. The relationship between political perspective and technological risk evaluations are shown in

Table 3. The correlation (R value) from a linear model is shown, as well as the P-value from a two way Student's T-Test.

As seen in Table 3, participants' views on artificial intelligence had a very small correlation with participants' political perspectives. Conservative ideology correlated with a slightly higher AI risk evaluation and slightly lower support for AI regulation. However, the strength of these correlations was very weak.

Neither AI risk assessments nor AI policy preferences exhibited a noticeable relationship with Active Open Minded Thinking, Need for Cognition, or religiosity. AI risk assessments did notably correlate with age; young people had a significantly lower evaluation of AI risk than their older peers. However, age had no significant relationship with AI policy preferences.

Table 3. Interaction between political perspective and risk/policy judgements

Variable	R Value (Correlation)	T-Test P Value
AI Risk Level	0.1020	2.87e-05
AI Regulation	-0.0995	8.084e-07
Genomic Modification Risk Level	0.1638	9.982e-13
Genomic Modification Regulation	-0.0380	0.04338
Driverless Car Risk	0.1928	2.2e-16

Discussion

This study suggest that artificial intelligence is a good background issue to study the emergence of political polarization. The results indicate that people have some familiarity with AI as a topic, but they have less certainty than in the case of more established technological risks like human genetic modification and driverless cars. This suggests that the public's views on AI are still developing.

Of the three technological risks covered in this pilot study, AI has the lowest degree of political polarization. There are slight correlations between participants AI views and their political perspectives, and a significant relationship between age and views on AI risk. However, these modest correlations appear very weak, and suggesting that the "AI issue" is not already too polarized to serve as the backdrop of a study on the emergence of political polarization.

Study 2

Introduction

The second study was designed to shed light on how political polarization emerges, and how "identity protective cognition" develops, if it does emerge. It uses an intergroup manipulation of social proof in the hopes of giving participants differing understandings of the political valence of an emerging issue – artificial intelligence. It then examines how that social proof influences subjects' views on this issue, if at all. It also examines how any induced political polarization affects subjects' interpretation of statistical evidence.

Participants

Study 2 was conducted on Amazon Mechanical Turk (MTurk), despite concerns about the representativeness of the typical MTurk sample. Many MTurk workers take academic surveys on a regular basis and may be familiar with techniques like the Cognitive Reflection Test and contingency table evaluations (Chandler, Mueller, & Paolacci, 2014). Additionally, some MTurk workers use scripts to speed up tasks by answering some questions automatically (Msori, 2018).

Validation questions can remove some of this uncertainty; we also used CAPTCHAs, survey timers, and IP location searching to eliminate fraudulent responses from the pool. These measures reduced but did not eliminate concerns about MTurk reliability. Ideally, the study would have used a more consistent service like YouGov, which ensures a more representative sample that is less familiar with standard research techniques. This alternative was not possible due to funding limitations.

A total of 1806 people completed the survey. We ran the first 200 subjects in a first batch, and ran a simple manipulation check to ensure that the study was functioning as anticipated. We did not change the number of participants at this time, or examine any other outcome variables. We then ran an additional two additional batches of participants, and used MTurk's qualification feature to keep people from taking the test twice.

A total of 1806 individuals started the survey, and a total of 144 people were removed from the sample for not completing the survey, for lying about their presence in the US,[8] or for refusing consent for their results to be used.

The remaining subject pool had 1662 participants. Surprisingly, 56.6% of the subject pool was male; in most online studies, women have higher response rates. Politically, the sample skewed somewhat to the left. This can be seen in Figure 2; 25.8% of the study sample identify as strong Democrats, while in Pew polls only 20**%** do (NW, Washington, & Inquiries, 2017).

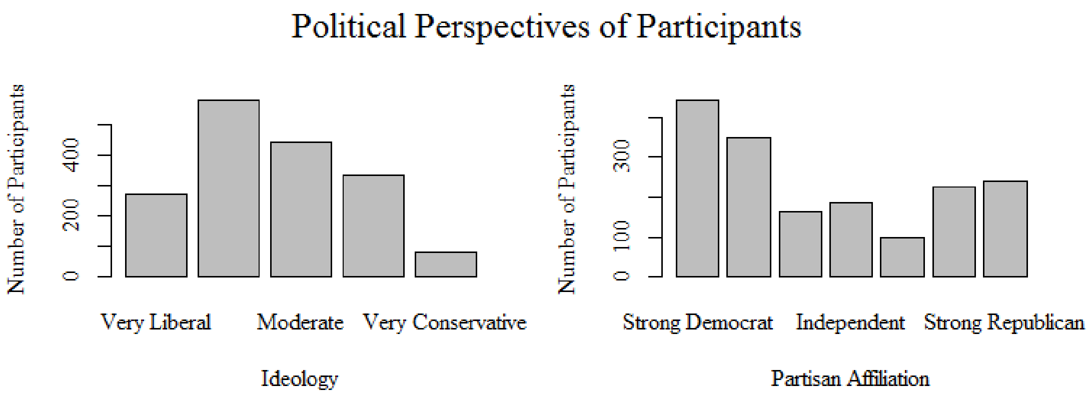


Figure 2. The political perspectives of Study 2's sample.

The study had a 3x2 between subjects design; participants were assigned to a condition randomly. Each of these 6 groups had 275 to 279 participants. For exact numbers and a visual representation of the design, see Table 6.

Design

First, participants were asked a set of basic demographic questions, including their gender, age, and economic well-being. Next, they were questioned about their political identity. In accordance with best surveying practice, we "pushed" avowed independents to express a part preference. Most registered

---

[8] Because the study included references to US politics, it was important to ensure that the sample only included Americans. People with foreign IP addresses or known VPN outlets were blocked before taking the survey.

independents favor one party over the other, and frequently their political perspectives are indistinguishable from those of registered political partisans. Indeed, there are some people who identify with one party for legacy reasons but mostly vote for another party's candidates in elections (Lake, 1984).[9]

Next, participants answered questions about voting behavior. Reliable voters are not necessarily representative of the population at large, and they likely have stronger partisan identities. I then asked participants how they feel about various institutions, both in political and AI spheres. These queries provide a wealth of useful information, including degree of negative partisanship, trust in the tech industry, and background feelings about AI.

Participants were then assigned into one of six groups, in a 3x2 structure. Each group encountered manufactured social proof about public opinions on AI research, through scripted web comments to a neutrally framed mock news article about AI. In the first direction, the political valence of AI risks was varied. For example, in the first two groups, conservatives appeared optimistic about the future of AI, while liberals expressed skepticism and a high risk evaluation. This is denoted as "conservatives for liberals against," or "CFLA." In the second two groups, liberals were presented as optimistic, and conservatives as pessimistic; this is written as "LFCA." In the final two groups, the social proof included a mixture of positive and negative perspectives from across the political spectrum. (This third group acts as a control.)

In the second direction we varied the intensity of the views. In the "implicit" condition, views were expressed with subtle political leanings. For example, a conservative perspective might emphasize the problems of "red tape and government inefficacy," while liberals might discuss "democratization and inclusion." However, the arguments were presented in a respectful non-partisan manner. In contrast, in the "explicit" condition, arguments were presented in an explicitly partisan manner. More invective was

[9] This study is an effort to investigate identity-driven cognition, and it would be a mistake to ignore self-proclaimed identification.

employed in this condition, and AI was presented as a zero-sum matter. Table 6 shows these conditions visually, including the number of people in each condition.

Before being exposed to the manipulated social proof, subjects were asked to read a concocted “news article” about artificial intelligence. The goals here are twofold. First, one potential confound is that people with certain identities might be more familiar with artificial intelligence, and that this background knowledge difference could be the root cause of apparent identity-driven polarization. Making participants read the article gives them a basic artificial intelligence primer, creating a minimum knowledge floor amongst participants and reducing the extent of this confound.

Secondly, the news article allowed us to credibly deliver the desired social proof through manipulated comment sections. Comment sections are text-based, allowing experimenters to assume multiple identities and curate conversations to generate controlled doses of social proof. Humans are adept at understanding the social dynamics of conversations. While participants scroll through a comment section, they will absorb (manipulated) clues about the positions of their co-partisans, as well as a sense of the partisanship of the issue. For example, participants in the explicit LFCA group might be exposed to:

> **UrbanMillennial:** AI is the way of the future, and it will radically improve economic inclusion and reduce the amount of time we spend at work. Services that were once only available to the super-wealthy are being democratized. This is exciting stuff. Let’s not be nostalgic about an imaginary past.
>
> **RedWave77:** In that “nostalgic past” you could get a good job right out of high school if you actually worked hard. They’d train you on the job, and you could stay in that job for your whole career and retire with a nice pension that you earned. Then rich elites got rid of a lot of those jobs and shipped them overseas. Now they want to take the rest and replace human workers with AIs. It’s all part of the same plan.
>
> **AmericanPatriot21:** I hear they still plan to make workers train their own replacements, just like they do with immigrant workers. Instead now it’s called “providing training data” for an algorithm.

Here, both sides are quite strident, and the “commenters” drop clues about their political identities though both their username and the content of their speech. They also exhibit negative partisanship towards the other side. In contrast, in the implicit CFLA group saw much softer comments, with more subtle political affiliation cues. Additionally, the positions of the liberal and conservative "commentators" was reversed:

> **SubwayFan:** Finding more uses of AI will only concentrate more wealth at the top while destroying middle class jobs. This is a genuine crisis in the making.
>
> **AmericanEngineer21:** You don't know that. Every technological revolution has created more jobs than it destroyed, and AI will be no different. Just because we don't know what those jobs are doesn't mean that they won't be created.
>
> **SubwayFan:** Yes, and every technological revolution required heavy regulation to protect society's most vulnerable people from its effects. We're already seeing the beginnings of this with Facebook and the Cambridge Analytica scandal and the data brokers that sell our info on the dark web. We need a General Data Protection Regulation like what they have in Europe.

The mixed condition included both optimistic and pessimistic comments from both sides of the political spectrum. The goal of this condition was to present a blank canvas onto which subjects to project their intuitions. Text-based conversations provide fewer context clues than other forms of conversational manipulation, like a recorded video, reducing the risk that test subjects will see the comments as inauthentic or unrealistic.

After reading the article and its associated "comments," participants examined a "study" about the effects of AI implementation on workers in a fictitious medium-sized logistics and manufacturing firm. They then were presented with a contingency table of "study results." There were two versions of this table, one where the mathematically counterintuitive but correct interpretation was "workers did better in an AI environment," and another where the mathematically counterintuitive but correct interpretation was "workers did better in a traditional environment." This approach will help assess whether both optimistic and pessimistic participants engage in biased evidence interpretation. The contingency tables associated with each version of the covariance detection are shown in Table 4.

	Better, Higher Paying Job	Worse, Lower Paying Job
Ohio (traditional)	190	75
Pennsylvania (artificial intelligence)	78	22

	Better, Higher Paying Job	Worse, Lower Paying Job
Pennsylvania (artificial intelligence)	190	75
Ohio (traditional)	78	22

Table 4. Contingency tables from the study instrument. Each participant was randomly shown one of the two. In the version on the left, workers did better in an AI environment, while in the version on the right, they did better in a traditional environment.

Participants were then asked whether AI policy has a partisan political valence, and if so, which position each side supports. This question functions as a manipulation check, testing whether the comment section actually influenced peoples' views about the political orientation of AI. They were also asked to assess their optimism about AI directly; for example, they were asked "Do you think AI advancements will create more opportunities or problems for society?"

Finally, participants took a standardized numeracy assessment based on Weller et al. These open-ended responses were coded and fed into a two parameter Item Response Analysis, from which I constructed a single, normalized numeracy estimate for each participant. Interested readers can find the completed instrument in the Appendix.

Hypotheses

This study is designed to shed light on several research questions at the same time. The hypothesized causal chain is shown in Figure 3. I predicted that comments would influence participants' understanding of the social context surrounding the AI discussion; specifically, I thought it would modify their beliefs about the relative optimism of liberals and conservatives in the direction of comments they read. I call this variable "perception of partisan issue valence." For example, someone assigned to see optimistic conservative comments and pessimist liberal ones (the CFLA condition) would be more inclined to say conservatives are more optimistic about AI than liberals.

I anticipated that perception of partisan issue valence would interact with participants preexisting political views to influence their economic and personal outlook on AI, a combined measure derived from their responses to the questions “how do you feel personally about new AI technologies?” and “do you think AI developments will impact you economically?” Participants answered these questions on a Likert scale, and their responses were added and normalized.

I further hypothesized that participants whose political views were “congruent” with optimistic AI comments would feel more positively about future AI developments. For example, I expected a typical left-leaning participant in the LFCA condition to be more optimistic about AI developments, while I expected a typical left-leaner to be pessimistic if assigned to the CFLA condition. The expectations for other combinations of political views and comment valences are shown in Table 5. Each combination of preexisting political views and comment valence is hypothesized to make a participant more optimistic about AI developments, more pessimistic about AI development, or have no impact on their views.

Randomly Assigned Comment Valence

		LFCA	CFLA	Mixed
Political Views	Left-wing	Optimistic or Congruent	Pessimistic or Incongruent	Neutral
	Right-wing	Pessimistic or Incongruent	Optimistic or Congruent	Neutral
	Center	Neutral	Neutral	Neutral

Table 5. Coding “congruency with AI support,” the hypnotized manipulation of participants’ optimism about AI developments.

I forecast that the relationship between comment valence and outlook would be mediated by perception of partisan issue valence. I also anticipated that “explicit” comment manipulations would be stronger than “implicit” ones, having a larger impact on participants’ economic and personal outlook.

If participants polarize along the lines of ideological belief, regardless of the comment valence, one could conclude that something about the nature of the artificial intelligence issue is prompting this polarization, and that social proof and framing have little effect. Instead, if social proof (comment valence) strongly impacts their “perception of partisan issue valence,” that would indicate that “artificial intelligence” is a neutral canvas, one that partisans project onto as it rises in popular consciousness.

I also hypothesized that participants “congruency” with AI support would influence the way they processed new evidence about AI, using the covariance detection task as a proxy for subjects’ ability to correctly interpret numerical evidence. I expected “congruent” or “optimistic” participants to perform relatively better on the covariance detection task when the correct answer to the task was “AI made workers better off;” I denoted these people as having “agreement” between their optimism and the pro-AI implications of their version of the task. I anticipated optimistic participants to do worse when the correct answer to the task was “a traditional environment made workers better off;” I denoted these subjects as having “disagreement” between their optimism and the anti- implications of their version of the task.

Conversely, I expected “incongruent” participants to perform relatively better on the task when the correct answer to the task was “a traditional environment makes workers better off” (agreement), and worse when the correct answer was “AI makes workers better off” (disagreement). A visual representation of this coding schema can be found in Table 5**.**

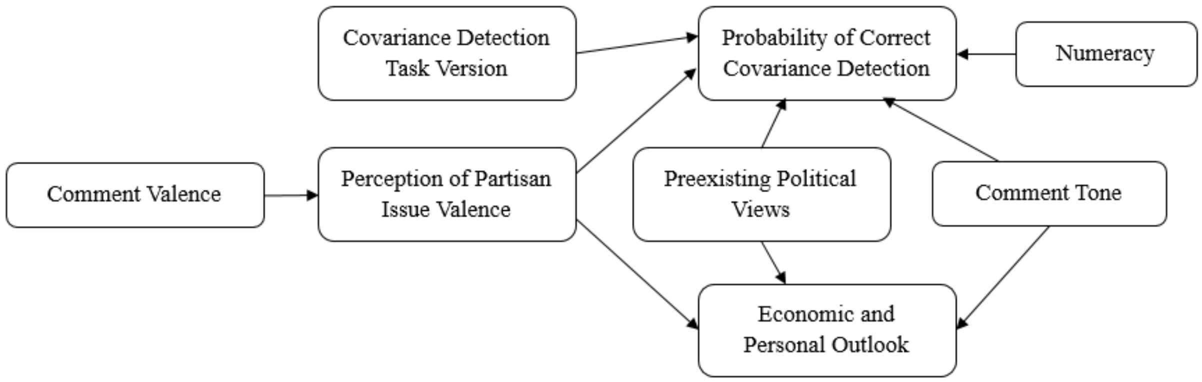
Congruency with AI Support

		Congruent (Optimistic)	Incongruent (Pessimistic)
Task Version	AI Environment Correct	Agreement	Disagreement
	Traditional Environment Correct	Disagreement	Agreement

Table 5. Coding alignment between the correct solutions of the covariance detection task and (hypnotized) manipulation of optimism about AI developments.

Additionally, this study aimed to distinguish between different types of polarization. On certain issues, like gun control and climate change, partisans have employed motivated reasoning to discount evidence that contradict their views, making errors that skew in favor of their perspective when interpreting unfavorable studies (Kahan, Peters, et al., 2017). Does this pattern emerge whenever polarized views exist, or only when it is framed as an identity-infused struggle?

I hypothesized that participants in the explicit condition would experience this “evidence processing divergence,” where “agreement” participants would perform much better at the covariance detection task than “disagreement” participants. Additionally, I expected that the relative difference between “agreement” subjects and “disagreement” subjects would be highest among high numeracy individuals. In contrast, I hypothesized that “agreement” and “disagreement” participants in the implicit condition would have little to no difference in their "probability of correct covariance detection," and that high numeracy individuals will process the data in an unbiased way.



Variable	Meaning
Comment Valence	The political perspective expressed by the comments. LFCA is coded as -1, mixed is coded as 0, and CFLA is coded as 1.
Comment Tone	Whether or not comments exhibit "explicit" or "implicit" partisanship.
Preexisting Political Views	A combination of a participant's political affiliation and ideological self-identification, summed and normalized.
Perception of Partisan Issue Valence	This is measured by the question "do you think liberals or conservatives are more optimistic about advances in artificial intelligence?"
Covariance Detection Task Version	This variable indicates the randomized version of the covariance detection task to which a participant was assigned.
Probability of Correct Covariance Detection	The proportion of participants that correctly solve a covariance detection task. This is a proxy for whether participants accurately interpret evidence.
Economic and Personal Outlook	Whether people think AI will have a positive or negative impact on their lives and livelihoods.
Numeracy	A composite score of numerical ability converted into a z-score with a two parameter item response model.

Figure 3. A flowchart of the hypothesized causal relationships between measured variables. The meaning of each variable name is explained in the associated table. These variable names will be used throughout the results discussion.

Results

The MTurk workers answered the survey very quickly. The median survey taker completed the study in 11.0 minutes, which was less than half the time we predicted from pretesting. This outcome indicates that participants were either rushing or have developed skills for fast survey taking. One of the participants contacted the research team with the information that the time requirement was misjudged, saying "most of us have been doing this type of thing for years so we simply work much faster than an average person or student." This comment suggests that there may be a major difference between MTurk workers and the general population.

Overall, participants felt moderately positive about the future of artificial intelligence. Only 20.6% felt that it would harm them economically, while 21.5% felt it would make no difference. In contrast, 57.87 believe AI advances will help them economically.

This study indicates that American MTurk workers have a strong prior belief that liberals as a group are more optimistic about AI developments. In the "mixed" control condition, where participants were exposed to a variety of optimistic and skeptical views from across the political system, only 27.4% of participants believed that conservatives were more optimistic than liberals. The social proof manipulation did not overcome this prior, as participants in the CFLA group still indicated that liberals were far more optimistic about AI.

The manipulation did have a significant impact on participants' understanding of the political orientation of AI as an issue. A higher proportion of participants in the CFLA treatment believed that conservatives were more optimistic about AI than in the mixed treatment condition; in turn, a higher proportion of people in the mixed condition believed conservatives were more optimistic than in the LFCA condition. A Chi squared test showed that the difference the CFLA and LFCA was significant ($p<0.01$). The difference between the CFLA and mixed conditions approached but fell short of significance ($p=0.0781$).

This manipulation check also held true when the data was split into the explicit and implicit tone conditions. The manipulation tended to be stronger in the explicit case, though this pattern did not hold for the LFCA case. This general trend likely occurred because comments in the explicit condition were much sharper, and would have been more obvious to a casual reader. The proportions of participants in each condition who thought that conservatives were more optimistic than liberals are shown in Table 6. I also constructed a series of logistic regression models for PPIV, which are shown in the Appendix in Table 7.

Perception of Partisan Issue Valence

Tone		CFLA	LFCA	Mixed
	Implicit	$\frac{82}{279} = 29.4\%$	$\frac{55}{279} = 19.7\%$	$\frac{73}{275} = 26.5\%$
	Explicit	$\frac{104}{277} = 37.5\%$	$\frac{57}{277} = 20.6\%$	$\frac{84}{275} = 30.5\%$

Table 6. Proportion of people in each group who believed conservatives are more optimistic than liberals about AI.

Since there was a high correlation between political ideology and partisan affiliation, I summed and normalized the two variables into a composite measure of "political view" or "political perspective." The distribution was shifted so that a "nonpartisan moderate" would be centered at zero, with "left of center" participants scored as less than zero, and "right of center" participants scored as greater than zero. I followed a similar normalization procedure to combine participants' "economic outlook" and "personal outlook" into a single measure.

I then compared each individual's political views with the ideological valence of the comments to which they were exposed, labelling them as "congruent with AI support," "incongruent with AI support," or neutral. This process shown in Table 5; individuals with completely neutral political views were excluded from the following analysis. Using a Wilcoxon nonparametric signed-rank test, I found that participants in the "congruent with AI support" condition were more optimistic than participants in the "incongruent with AI support" condition about the impact of AI on their economic and personal prospects. ($p<0.01$).

I then constructed a series of regression models to examine how assigned comment valence, assigned comment tone, numeracy, and preexisting political perspective combine to influence participants' economic and personal outlook on AI developments. (The "economic and personal outlook" variable was assumed to have a normal distribution; exploratory data analysis confirmed that this assumption was reasonable.) I added these variables of interest in stages. Model 1 examines how the assigned comment valence and preexisting political view interact to influence economic and personal outlook. Model 2 adds an indicator variable representing assigned comment tone. Model 3 includes all the variables from Models 1 and 2, but it also models the effects of the participants' numeracy. The logit coefficients from the three models can be found in Table 8 in the Appendix.

Next, I analyzed subjects' performance on the covariance detection task. As expected, the covariance detection task was difficult, and 32.2% of participants were able to solve it correctly. This matches findings from the literature (Kahan & Peters, 2017). Interestingly, when the right answer to the task was "workers did better in the AI environment," 44.6% of participants answered correctly. When the right answer was "workers did better in the traditional environment," only 19.8% of them answered correctly. This difference was statistically significant.

I constructed a similar set of three logistic regression models to investigate how comment valence, assigned comment tone, numeracy, and preexisting political perspective impact performance on the covariance detection task. The outcome variable indicates whether a given participant gave the counterintuitive but mathematically correct answer to the covariance detection question. (A "0" signifies an incorrect answer, while a "1" signifies a correct response.) The inputs to the regression models were identical to the ones above. The results of these analyses are shown in Table 9.

Figure 4 displays the relationship between numeracy and proportion of correct covariance detection answers. Participants with "answer-view agreement" and participants with "answer-view disagreement are plotted separately. I predicted that subjects with "answer-view agreement" would perform better than subjects with "answer-view disagreement" (See the Hypothesis section and Table 5 for an explanation of my reasoning).

The graph depicts a subtle "U" shape. I believe a substantial number of low numeracy individuals guessed randomly, which would have had a paradoxical effect of raising the proportion they got correct. Individuals with middling numeracy overwhelmingly fell for the mistake, while a substantial number of high numeracy individuals interpreted the contingency table correctly.

While people with "answer-view agreement" do appear to perform better than people with "answer-view disagreement," the difference is fairly small. Additionally, there is no indication that higher numeracy individuals are more polarized than their lower numeracy peers. There is a notable difference between participants who were exposed to “explicit” and “implicit” comments. Explicit comments appear to have substantially boosted the distinction between “answer-view agreement” and “disagreement,” but high numeracy individuals do not exhibit a bigger difference in this regard than low numeracy individuals.

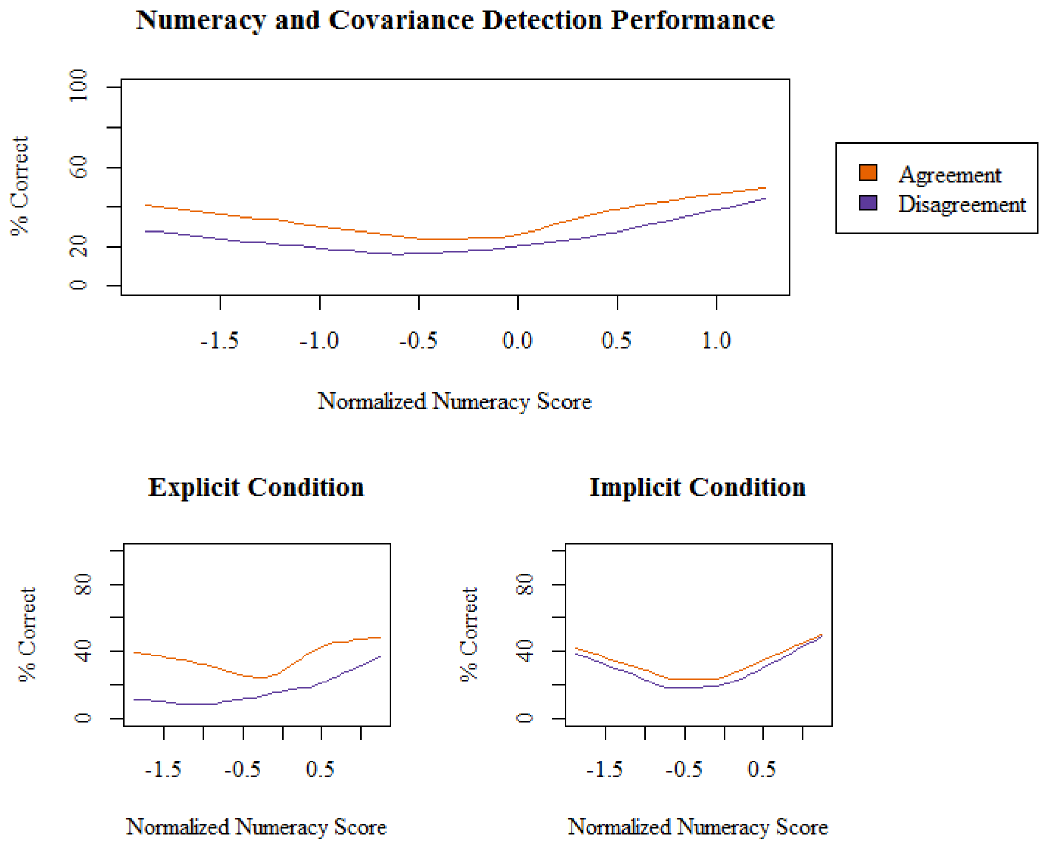


Figure 4. Relationship between answer-view (dis)agreement, numeracy, and evidence interpretative ability on the covariance detection task.

Discussion

Though the manipulation worked, the effect size was small, which hampered the other analyses that I sought to run. The explicit condition had a larger impact on the participants’ perspectives about the benefits and costs of AI; future experiments should probably forgo the implicit condition. In fact, it might make sense to construct even more obvious and vehement comments or to present comments on multiple articles to increase participants’ exposure.

This study also revealed some preexisting beliefs on the part of the survey takers. The social science literature on AI suggests that Americans have a modestly optimistic outlook on AI, and the data from this study is consistent with these findings. However, that body of studies provides no indication that a strong majority of the public believes that liberals are more optimistic about AI developments than conservatives, or that people have a strong preexisting belief that workers will do better when AI is implemented in their workplaces. In fact, the second belief was strong enough to exert a strong confirmation bias. The existence of the unanticipated priors is a reminder that the only way to accurately model public opinion is to measure a representative sample, and to conduct regular surveys so as to pick up shifts in opinion.

General Discussion

Analysis of results

This thesis aimed to address four major research questions. First, it considered whether emerging questions of technological risk have inherent partisan valences. Secondly, it sought to evaluate the effectiveness of internet comments as a form of social proof. Third, it intended to determine the extent to which perceptions of an issue's partisan valence influenced object-level beliefs about that issue. And fourth, it hoped to distinguish between having polarized preexisting beliefs and engaging in biased evidence interpretation. It also aimed to identify sufficient conditions for biased evidence processing to emerge. The empirical portion of this thesis mostly generated inconclusive results. However, it does provide valuable lessons for future research.

Initially, I hypothesized that questions of AI risk would have no inherent "orientation," and that a dose of manipulated social proof could move subjects' views in either direction. There is some evidence to support this view, as valence of the randomly assigned comments did have a significant impact on subjects' perceptions about AI's partisan issue valence (PPIV). However, the comments had a relatively small effect on PPIV; a substantial majority of participants from every group believed that liberals were more optimistic about AI than conservatives.

The existence of this prior does not necessarily mean that AI risk has an inherent partisan lean. Participants may have started out neutral on the question, only to be biased by other pretest materials. For example, AI is strongly associated with California, and some participants have associated AI with liberal Californian policy. However, it does raise significant doubts about AI as a politically neutral issue subject to induced polarization.

Additionally, I am now inclined to argue that evidence suggesting that an issue could be polarized both ways does not imply that the issue lacks an inherent lean. It may be that only elites and experts have the wherewithal to identify how questions of technological risk fit into a larger ideological narrative. Other people could then adapt their views based on the positions taking by the early-adopting elite. In this model, changing social proof about an issue could indeed influence the perspectives of the general public; however, the latent issue orientation could still be very important to the social-proof generating elite. I therefore must conclude that studies of this type cannot answer this research question in principle.

The results from Study 2 indicate that the political perspectives and tone expressed in internet comments on a single article can significantly impact whether participants believe liberals or conservatives are more optimistic about AI. We called this outcome “Perception of Partisan Issue Valence” (PPIV.) Both the Chi-squared tests and the logistic regression analysis in Table 7 confirm that the partisan valence of the randomly assigned comments had a significant impact on participants PPIV evaluations.

The impact of comment tone was slightly less clear. Models 2 and 3 from the regression analysis suggest that the “CFLA explicit” condition influenced a higher proportion of participants than the “CFLA implicit” condition, making them more likely to say that “conservatives are more optimistic about AI.” However this difference was not statistically significant; the “LFCA explicit” and the “LFCA implicit” conditions had no real difference in PPIV responses.

These results largely agree with my hypotheses, though these results do not conclusively show that "explicit" comments have a larger impact than "implicit" ones. However, the manipulation was not strong enough to overcome many participants' prior belief that liberals were more optimistic about AI than conservatives.

Some subjects may have read the article and comments too quickly to absorb the comment valences. Others may have found the comments contrived and discounted them. The political valence in the assigned comments may have been too subtle for many participants recognize it. Finally, the manipulation may have influenced some participants, but not strongly enough to overcome their preexisting belief that liberals are more optimistic about AI developments than conservatives. Whatever its cause, the lack of manipulation strength posed a serious challenge to the remaining research questions.

I had hypothesized that the political valence of the assigned comments would influence participants' economic and personal outlook on artificial intelligence; I believed that participants whose political views aligned with the pro-AI comments in their assigned condition would be more optimistic about AI than participants whose political views aligned with the anti-AI comments. Subjects in the first group were coded as being "congruent with AI support," while those in the second group were coded as being "incongruent with AI support."

The average "congruent" participant had a slightly more positive economic and personal outlook than the average "incongruent" participant, and a Wilcoxon ranked-sign test confirmed that this difference was significant ($p = 0.0095$.) A series of linear regression analyses similarly found that "economic and personal outlook" and "political perspective" (a normalized variable where conservatives were scored as greater than zero) had a positive correlation in the LFCA condition, and a negative correlation in the CFLA condition.

The difference between "CFLA x PP" and the "LFCA x PP" logit coefficients was statistically significant in Model 1, which only considers the assigned comment valence and the subjects political perspective. Models 2 and 3, which consider the comment tone and the participants' numeracy as well, were almost but not quite significant.

Linear Regression Model 3 also suggests that high numeracy individuals are more sensitive to the subtle manipulation of the "implicit" condition. "LFCA x PP x Num x Explicit" has a positive logit coefficient, while "CFLA x PP x Num x Explicit" has a negative one. This suggests that high numeracy individuals are recognizing the ideological valence of comments in both the implicit and explicit conditions, and adjusting their views accordingly. In contrast, low numeracy individuals only seem to recognize the valence of explicitly political comments, and are much less sensitive to implicit ones. It is important to note that this effect is not statistically significant.

I also hypothesized that explicitly partisan comments would produce identity-based political polarization, and that participants in the explicit condition would engage in biased evidence interpretation on the covariance detection task. Participants coded as having "answer-view agreement" were expected to perform better than those with "answer-view disagreement." (See the Hypothesis section and Table 5 for an explanation of my reasoning). I also anticipated that the high numeracy individuals would evince the most bias, consistent with other known cases of identity-protective cognition (Kahan, Peters, et al., 2017) I expected much less motivated reasoning from the implicit group, which I predicted would have less of an identity-based attachment to the issue. Finally, I believed that participants in each of the counterbalanced covariance detection tasks would have comparable error rates.

These predictions proved to be very incorrect. Participants across all experimental conditions answered more covariance detection problems correctly when the counterintuitive but mathematically correct solution was "the AI environment was better for workers" than when the counterintuitive but mathematically correct was "the traditional environment was better for workers."

I'm not sure why this difference emerged. It's possible that participants had a strong preexisting belief that technologically advanced workplaces are better for workers, and that motivated reasoning prompted to give this answer more readily. Many MTurk workers spend a lot of time labeling training data for neural network-based AIs, so this perspective may more popular among MTurk survey takers than the general population. It is possible that some participants were doing the study so quickly that they didn't see that they were answering a question about the well-being of workers, and instead answered based on a general impression that AI does positive things. Finally, there may have been some element of a demand effect, where participants believed that the experimenter wanted them to say “AI is better.”

I conducted a set of logistic regression analyses. Model 1 attempted to predict the probability of correct covariance detection using combinations of comment valence, political perspective, and covariance detection task version. Model 2 added the comment tone, and Model 3 further considered numeracy.

The logit coefficients of the resulting predictor variables mostly have signs in the direction of my hypotheses. However, most of them are not even close to statistically significant, so we cannot safely draw any inferences from them. Nevertheless these extremely provisional results suggest that a stronger manipulation and more power could have satisfied at least as some of my hypotheses.

When the correct answer about the better situation for workers was the “AI environment,” congruent participants like a liberal in the LFCA condition or a conservative in a CFLA condition performed better than their congruent counterparts. In contrast, when the correct answer was the “traditional environment,” incongruent subjects performed a bit better than congruent ones. As expected, these effects were larger in the explicit condition. Numeracy was correlated with increased covariance detection performance, as predicted; however, it was not associated with increased bias in the explicit condition, violating my hypothesis. These relationships can be examined visually in Figure 4.

Ultimately, the weakness of the experimental manipulation led to many inconclusive results. This thesis supports the view the idea that internet comments can be an effective source of social proof, but it generates more questions than answers with respect to the other research queries.

Limitations

Study 1 had very helpful data. However, it only addressed a handful of emerging technological risks, which restricted the potential subjects for Study 2. An ideal pilot study would have cast a larger net, asking about the risks of other emerging technologies like nanotechnology and cyber warfare. Additionally, Study 1 did not provide any guidance on how participants perceived the partisan valence of risks. It is possible that another issue would have had a more neutral partisan valence than AI, which would have helped me address my research questions more clearly.

Study 2 attempted to communicate how people of different political affiliations felt about a particular issue with a single news article and comment section. Unfortunately, the valence of the assigned comments appears to have had a relatively small impact on the variables of interest, like perception of partisan issue valence, economic and personal outlook, and covariance detection performance. This suggests that the comments used in Study 2 may not have sufficiently obvious or extreme enough to allow subjects to pick up on the manipulated political alignment of the comments. A stronger manipulation would have helped across all measures.

Study 2 also had some methodological flaws. I placed questions about participants' political identity at the beginning of the survey, as I was concerned that the questions about the political valence of AI would influence participants' ideological self-assessments. In retrospect, I think it would have made more sense to place questions about demographics and political identity at the end of the study, as the risk of priming people to consider politics may well be greater than the risk of them changing their political perspective based on a single discussion of AI.

I used a binary variable to measure subjects' perception of partisan issue valence. However, a more discriminative Likert scale would have let us distinguish between individuals who did not recognize the comment valence at all and individuals who recognized the comment valence, but did not give it enough evidentiary weight to overcome their prior beliefs.

Humans are generally well attuned to the identity affiliations that emerge in conflicts. Evolutionary psychologists posit that our natural aptitude for locating group fault lines is vitally important for successfully navigating social disagreements, an ability on which all social primates rely (Waal & Waal, 1989)**.** However, the majority of the study took place online, which means that subjects could not assess the body language and verbal cues that help people identify in-group and out-group dynamics. As mentioned earlier, people find it far harder to identify insincerity and lies through text-based communications; it can also be very difficult to identify sarcasm and parody without additional, explicit markers (Aikin, 2013; Hancock, Landrigan, & Silver, 2007).[10] It seems plausible that the removal of these indicators also reduces our ability to identify the political affiliations of various viewpoints, requiring more exposure over a longer period of time to communicate the same social information.

One must further note the obvious limitation that the study here focuses on assessing political polarization around a single issue – artificial intelligence. Artificial intelligence served as a good initial issue to study; it is timely, it has little preexisting polarization, and its meaning is still unclear to the general public (Over half of all participants in Study 2 had heard "nothing," "a little," or "a moderate amount" about AI). Importantly, we believe optimism and skepticism about artificial intelligence to be uncorrelated with political affiliation.[11] However, artificial intelligence is just one topic, and it would be unwise to generalize risk assessment polarization more broadly without conducting similar experiments using other topics, such as nanomaterials, synthetic biology, and genomic manipulation. There is a case for conducting surveys of the sort carried out for this thesis very early in the development of such technologies.

Future directions

Study 2 gave inconclusive answers to several research questions. Nevertheless, I think it would be valuable to run a modified follow-up study with its defects corrected. Hypothetically, this new study would employ a more discriminative metric for perception of political issue valence, and it would put all demographic questions at the end of the survey, or perhaps randomize whether subjects received those questions at the beginning or the end. The mixed condition would be eliminated, as it used up one third of the participant pool without supporting many inferences. Most importantly, I would use more aggressive manipulations. The current "explicit" condition would become the "implicit" condition on the follow-up

[10] There is a well-known internet adage known as Poe's Law which states that without some sort of explicit signifier, it is impossible to create an internet parody so extreme that it is not mistaken for the actual views being parodied (Aikin, 2013).
[11] Surprisingly, while many opinion polls have examined the public's views on artificial intelligence, we could not find any survey or study which had an explicit political affiliation crosstab. We base this assessment on historical attitudes about artificial intelligence and the technology industry expressed by political leaders in Congress. Testing this assumption was one part of Study 2's goals.

study, and a new, even more vitriolic “explicit” condition. Ideally, this survey would be taken by a YouGov panel, reducing some of the concerns about the unrepresentativeness of MTurk workers.

Future studies could also explore other ways to create manipulated social proof. One possibility would be to use news stories from different outlets. Participants could read a “Fox News” story to get a manipulated conservative perspective, and a “MSNBC” story to get a manipulated liberal one. This approach could be combined with comments to further strengthen the manipulation.

While artificial intelligence provides a good initial topic for examining the emergence of polarized probabilistic assessments, other topics should be explored as well. Additionally, further research should explicitly focus on avoiding risk assessment polarization, and on depolarizing conversations where polarization already exists. We know that people with high levels of “science curiosity” exhibit far better capacity to evaluate risks on controversial issues (Kahan, Landrum, et al., 2017). Further research is needed to determine whether this relationship is causal or correlational. If it is the former, we should examine how scientific curiosity can be cultivated in schools, universities, and among the public at large.

Concluding remarks

It is more difficult to communicate science effectively than it is to peddle a conspiracy theory; it is easier to derail a reasoned discussion than it is to facilitate one. Our society faces immensely complex risks, and to make good collective decisions it will not be enough to simply communicate scientific concepts and information to policy makers and voters.

We have an ethical responsibility not just to investigate manipulation techniques, but to examine potential strategies for defeating them. Research ought to be used in service of reason and democracy. However, academic investigations into manipulation techniques must consider the risk that they specify a more detailed roadmap for people who wish to undermine democratic deliberations and polarize an emerging issue along party or ideological lines.

It is obvious why research into reversing the effects of polarization could help facilitate democratic discussion and improve science communication. However, it is important to recognize how research into the initial emergence of polarization might serves these purposes as well. People who run astroturfing campaigns to influence public opinion already have a working understanding of the role that social proof plays in the spreading misinformation and fear, manufactured uncertainty and unjustified doubt. The introduction to this thesis discusses the research into “depolarizing” risk identity-laden risk assessments. However, there has been little experimental evidence to guide the communication strategies for those who seek to keep systematic manipulation efforts from taking hold. The ultimate goal of this research program is to help science communicators prevent issue polarization instead of just trying to cure it after the fact.

Gaining a better understanding of how politically polarized risk assessments emerge will help science communicators avoid inadvertently tying understandings of risk to identity-laden cultural values. It could guide moderators on web sites, and could inform the structure and tone of arguments used by civil society groups. While this thesis does not contain any actionable insights, it is our hope that its conceptual successors will. There is no easy panacea for polarization in society today, but good science communication can still protect vital political questions from being tied to people's partisan identities.

Author Contributions

Balleisen conducted a literature review with advice and recommendations from Kahan. Kahan and the Cultural Cognition Lab designed, conducted, and coded Study 1. Balleisen analyzed the data. Balleisen conceived of Study 2, and developed the protocol. Extensive feedback from Kahan was incorporated into the design. Balleisen conducted Study 2 on Amazon Mechanical Turk, coded the results, and performed the statistical analysis. Kevin Anderson helped implement the analysis of the numeracy measure. Balleisen wrote the thesis by himself, and incorporated extensive feedback from Kahan.

Acknowledgments

I would like to thank Dan Kahan and the Yale Law School's Cultural Cognition Project for supporting this project. I would also like to thank Meredith Berger for her administrative support. I am grateful to Mark Sheskin for his guidance over the last two years, and to my peer reviewers who (will undoubtedly) help sharpen my ideas and improve my writing. This research would not have been possible without the support of the Yale Cognitive Science program. Additionally, my study was facilitated by generous financial support from the Yale Psychology Department. Finally, I would like to thank my friends and family for their suggestions and encouragement throughout this process.

References

2016 CBS Interactive Inc. (2016, March 18). 60 Minutes/Vanity Fair poll: Artificial Intelligence. Retrieved February 10, 2019, from CBS News website: https://www.cbsnews.com/news/60-minutes-vanity-fair-poll-artificial-intelligence/

Aikin, S. F. (2013). Poe's Law, group polarization, and argumentative failure in religious and political discourse. *Social Semiotics*, *23*(3), 301–317. https://doi.org/10.1080/10350330.2012.719728

Brewer, P. R., & Gross, K. (2005). Values, Framing, and Citizens' Thoughts about Policy Issues: Effects on Content and Quantity. *Political Psychology*, *26*(6), 929–948. https://doi.org/10.1111/j.1467-9221.2005.00451.x

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116.

Cantwell, M. (2017, December 12). Cantwell, Bipartisan Colleagues Introduce Bill to Further Understand and Promote Development of Artificial Intelligence, Drive Economic Opportunity [Official Web Page]. Retrieved April 21, 2019, from U.S. Senator Maria Cantwell's Sentate Press Releases website: https://www.cantwell.senate.gov/news/press-releases/cantwell-bipartisan-colleagues-introduce-bill-to-further-understand-and-promote-development-of-artificial-intelligence-drive-economic-opportunity

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, *46*(1), 112–130. https://doi.org/10.3758/s13428-013-0365-7

Ching, M. K. L. (1982). The Question Intonation in Assertions. *American Speech*, *57*(2), 95–107. https://doi.org/10.2307/454443

Cook, J., & Lewandowsky, S. (2016). Rational Irrationality: Modeling Climate Change Belief Polarization Using Bayesian Networks. *Topics in Cognitive Science*, *8*(1), 160–179. https://doi.org/10.1111/tops.12186

Fast, E., & Horvitz, E. (2016). Long-Term Trends in the Public Perception of Artificial Intelligence. *ArXiv:1609.04904 [Cs]*. Retrieved from http://arxiv.org/abs/1609.04904

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Gamson, W. A., & Modigliani, A. (1989). Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach. *American Journal of Sociology*, *95*(1), 1–37.

Green, E. J. D. and J. C. (2008, February 1). Religion and American Politics: More Secular, More Evangelical…or Both? Retrieved March 25, 2019, from Brookings website: https://www.brookings.edu/research/religion-and-american-politics-more-secular-more-evangelical-or-both/

Hancock, J. T., Landrigan, C., & Silver, C. (2007). Expressing Emotion in Text-based Communication. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 929–932. https://doi.org/10.1145/1240624.1240764

Huang, T. (2005, May 2). Framing Stories. Retrieved March 25, 2019, from Poynter website: https://www.poynter.org/reporting-editing/2005/framing-stories/

ITU. (2018). *Assessing the Economic Impact of Artificial Intelligence* (No. Issue Paper No. 1; p. 80). Retrieved from International Telecommunication Union website: http://handle.itu.int/11.1002/pub/81202956-en

Kahan, D. M. (2017). 'Ordinary science intelligence': a science-comprehension measure for study of risk and science communication, with notes on evolution and climate change. *Journal of Risk Research*, *20*(8), 995–1016. https://doi.org/10.1080/13669877.2016.1148067

Kahan, D. M., Jenkins-Smith, H. C., Tarantola, T., Silva, C. L., & Braman, D. (2012). *Geoengineering and Climate Change Polarization: Testing a Two-channel Model of Science Communication* (SSRN Scholarly Paper No. ID 1981907). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=1981907

Kahan, D. M., Landrum, A., Carpenter, K., Helft, L., & Jamieson, K. H. (2017). Science Curiosity and Political Information Processing. *Political Psychology*, *38*(S1), 179–199. https://doi.org/10.1111/pops.12396

Kahan, D. M., & Peters, E. (2017). Rumors of the “Nonreplication” of the “Motivated Numeracy Effect” are Greatly Exaggerated. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3026941

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*(1), 54–86. https://doi.org/10.1017/bpp.2016.2

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732–735. https://doi.org/10.1038/nclimate1547

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Kutzner, F., Freytag, P., Vogel, T., & Fiedler, K. (2008). Base-Rate Neglect as a Function of Base Rates in Probabilistic Contingency Learning. *Journal of the Experimental Analysis of Behavior*, *90*(1), 23–32. https://doi.org/10.1901/jeab.2008.90-23

Lake, C. (1984). Coding of Independent/Independents and Apoliticals in the Party Identification Summary Code and Apoliticals in the Rolling Cross Section. *ANES Technical Report Series*, (nes010124).

Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online* (p. 106). New York, NY, USA: Data & Society.

Mitchell, A., Matsa, K. E., Gottfried, J., & Kiley, J. (2014, October 21). Political Polarization & Media Habits [Pew Research Center]. Retrieved March 25, 2019, from Trust, Facts, and Democracy website: https://www.journalism.org/2014/10/21/political-polarization-media-habits/

Mittal, V., Malshe, A., & Sridhar, S. (2018). The Unequal Effects of Partisanship on Brands. *Harvard Business Review*. Retrieved from https://hbr.org/2018/03/the-unequal-effects-of-partisanship-on-brands

Msori. (2018). r/mturk - MTurk Scripts [News Aggregation and Discussion]. Retrieved March 30, 2019, from Reddit website: https://www.reddit.com/r/mturk/comments/6gbkyj/mturk_scripts/

Mullin, B. (2016, September 20). NYT to partner with division of Google's parent company to expand comments. Retrieved March 30, 2019, from Poynter website: https://www.poynter.org/tech-tools/2016/nyt-to-partner-with-division-of-googles-parent-company-to-expand-comments/

National Highway Traffic Safety Administration (NHTSA), D. of T. (US). (2009). *Traffic Safety Facts: Occupant Protection* (No. DOT HS 811 160; pp. 3–4). Washington, D.C.: NHTSA.

Nelson, T. E., Clawson, R. A., & Oxley, Z. M. (1997). Media Framing of a Civil Liberties Conflict and Its Effect on Tolerance. *The American Political Science Review*, *91*(3), 567–583. https://doi.org/10.2307/2952075

Nivola, W. A. G. and P. S. (2008, May 11). Vote Like Thy Neighbor: Political Polarization and Sorting. Retrieved March 25, 2019, from Brookings website: https://www.brookings.edu/articles/vote-like-thy-neighbor-political-polarization-and-sorting/

NW, 1615 L. St, Washington, S. 800, & Inquiries, D. 20036 U.-419-4300 | M.-419-4349 | F.-419-4372 | M. (2017, October 24). Political typology: Partisanship and political engagement | Pew Research Center. Retrieved April 14, 2019, from https://www.people-press.org/2017/10/24/1-partisanship-and-political-engagement/

Office of the Director of National Intelligence. (2017). *Assessing Russian Activities and Intentions in Recent US Elections* [Declassified Intelligence Report]. Washington, D.C.

Peiser, J. (2019, February 11). The Rise of the Robot Reporter. *The New York Times*. Retrieved from https://www.nytimes.com/2019/02/05/business/media/artificial-intelligence-journalism-robots.html

Postman, N. (2011). *Technopoly: The Surrender of Culture to Technology*. Retrieved from https://books.google.com/books?id=gYrIVidSiLIC

Primi, C., Morsanyi, K., Donati, M. A., & Chiesi, F. (2014). Item Response Theory analysis of the Cognitive Reflection Test: Testing the psychometric. *Issue: Proceedings of the Annual Meeting of the Cognitive Science Society, 36*.

Rock, S. M. (1993). Risk compensation and the Illinois seat belt use law. *Accident; Analysis and Prevention*, *25*(5), 537–544.

Rottman, B. M., Prochaska, M. T., & Deaño, R. C. (2016). Bayesian reasoning in residents' preliminary diagnoses. *Cognitive Research: Principles and Implications*, *1*(1), 5–5. https://doi.org/10.1186/s41235-016-0005-8

Schenck, A. P., Runyan, C. W., & Earp, J. A. (1985). Seat belt use laws: the influence of data on public opinion. *Health Education Quarterly*, *12*(4), 365–377.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*(2), 342–357. https://doi.org/10.1037/0022-0663.89.2.342

Tenenbaum, J. B. (1999). *Bayesian Modeling of Human Concept Learning*. 59–65. Retrieved from http://dl.acm.org/citation.cfm?id=340534.340562

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7), 1275. https://doi.org/10.3758/s13421-011-0104-1

Vogel-Fox, E. (2017, December 12). Congress debates the role of artificial intelligence in America [News Website]. Retrieved April 21, 2019, from WJLA website: http://wjla.com/news/nation-world/congress-debates-the-role-of-artificial-intelligence-in-america

Waal, F. B. M. de, & Waal, F. de. (1989). *Peacemaking Among Primates*. Harvard University Press.

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and Testing of an Abbreviated Numeracy Scale: A Rasch Analysis Approach. *Journal of Behavioral Decision Making*, *26*(2), 198–212. https://doi.org/10.1002/bdm.1751

West, D. M. (2018a, May 21). Brookings survey finds worries over AI impact on jobs and personal privacy, concern U.S. will fall behind China. Retrieved February 4, 2019, from Brookings website: https://www.brookings.edu/blog/techtank/2018/05/21/brookings-survey-finds-worries-over-ai-impact-on-jobs-and-personal-privacy-concern-u-s-will-fall-behind-china/

West, D. M. (2018b, June 21). Brookings survey finds 52 percent believe robots will perform most human activities in 30 years. Retrieved February 4, 2019, from Brookings website: https://www.brookings.edu/blog/techtank/2018/06/21/brookings-survey-finds-52-percent-believe-robots-will-perform-most-human-activities-in-30-years/

Whitmarsh, L. (2011). Scepticism and uncertainty about climate change: Dimensions, determinants and change over time. *Global Environmental Change*, *21*(2), 690–700. https://doi.org/10.1016/j.gloenvcha.2011.01.016

Appendix

Regression Tables

Outcome Variable: Perception of Partisan Issue Valence

	Model 1		Model 2		Model 3	
Predictors	Coef.	SE	Coef.	SE	Coef.	SE
(Intercept)	0.80	0.10	0.89	0.14	0.93	0.15
LFCA	0.53	0.14	0.50	0.21	0.49	0.21
CFLA	-0.17	0.13	-0.05	0.20	-0.07	0.20
PP	-0.42	0.09	-0.42	0.14	-0.42	0.14
LFCA x PP	0.08	0.14	0.21	0.20	0.18	0.21
CFLA x PP	0.18	0.13	0.27	0.19	0.26	0.19
Explicit			-0.18	0.20	-0.21	0.20
LFCA x Explicit			0.03	0.29	0.07	0.30
CFLA x Explicit			-0.25	0.27	-0.27	0.28
PP x Explicit			0.01	0.19	0.01	0.19
LFCA x PP x Explicit			-0.27	0.28	-0.26	0.29
CFLA x PP x Explicit			-0.19	0.26	-0.27	0.27
Num					0.39	0.18
LFCA x Num					-0.14	0.26
CFLA x Num					-0.44	0.24
PP x Num					-0.10	0.17
Num x Explicit					-0.31	0.24
LFCA x PP x Num					-0.07	0.26
CFLA x PP x Num					0.21	0.24
PP x Num x Explicit					0.10	0.24
LFCA x Num x Explicit					0.28	0.36
CFLA x Num x Explicit					0.49	0.34
LFCA x PP x Num x Explicit					-0.39	0.36
CFLA x PP x Num x Explicit					-0.65	0.33

Table 7. Multivariate logistic regression analysis. N = 1662. Outcome variable is "perception of partisan issue valence." The predictor estimates logit coefficients are labelled "Coef" and the standard error is labeled "SE." The outcome is a binary indicator variable (0 = incorrect answer, 1 = correct answer.) The predictor estimates are logit coefficients. "LFCA," "CFLA," and "Explicit" are indicator variables reflecting randomized condition assignment (0 = unassigned, 1 = assigned). "Mixed" and "Implicit" are reference variables. "PP" and "Num" represent combined "political view" and "Num" measures respectively. "PPIV" stands for "Perception of Partisan Issue Valence," an observational indicator variable.

Outcome Variable: Economic and Personal Outlook

	Model 1		Model 2		Model 3	
Predictors	Coef.	SE	Coef.	SE	Coef.	SE
(Intercept)	-0.03	0.05	-0.08	0.06	-0.08	0.06
LFCA	0.02	0.06	0.05	0.09	0.05	0.09
CFLA	0.01	0.06	0.02	0.09	0.01	0.09
PP	-0.08	0.04	-0.10	0.06	-0.10	0.06
LFCA x PP	-0.04	0.06	-0.01	0.08	-0.01	0.09
CFLA x PP	0.11	0.06	0.14	0.08	0.14	0.09
Explicit			0.10	0.09	0.10	0.09
LFCA x Explicit			-0.07	0.13	-0.05	0.13
CFLA x Explicit			-0.01	0.13	-0.05	0.13
PP x Explicit			0.04	0.08	0.04	0.09
LFCA x PP x Explicit			-0.05	0.12	-0.04	0.12
CFLA x PP x Explicit			-0.05	0.12	-0.10	0.12
Num					-0.01	0.08
LFCA x Num					-0.01	0.11
CFLA x Num					-0.01	0.11
PP x Num					0.05	0.08
Num x Explicit					-0.10	0.11
LFCA x PP x Num					-0.07	0.11
CFLA x PP x Num					-0.08	0.11
PP x Num x Explicit					-0.08	0.11
LFCA x Num x Explicit					0.19	0.15
CFLA x Num x Explicit					-0.06	0.16
LFCA x PP x Num x Explicit					0.18	0.15
CFLA x PP x Num x Explicit					-0.05	0.15

Table 8. Multivariate linear regression analysis. N = 1662. Outcome variable is "economic and personal outlook," a combined normalized measure. The predictor estimates are labelled "Coef" and the standard error is labeled "SE." "LFCA," and "CFLA," and "Explicit" are indicator variables reflecting randomized condition assignment (0 = unassigned, 1 = assigned). "PP" and "Num" represent combined "political perspective" and "numeracy" measures respectively. "PPIV" stands for "Perception of Partisan Issue Valence," an observational indicator variable. (0 = participant says conservatives are more optimistic about AI, 1 = participant says liberals are more optimistic about AI.)

Outcome Variable: Probability of Correct Covariance Detection

	Model 1		Model 2		Model 3	
Predictors	Coef.	SE	Coef.	SE	Coef.	SE
(Intercept)	-1.62	0.17	-1.53	0.24	-1.57	0.25
LFCA	0.37	0.23	0.05	0.33	0.04	0.36
CFLA	0.13	0.24	0.20	0.32	0.18	0.34
PP	-0.14	0.16	-0.17	0.22	-0.13	0.22
CDTV	1.23	0.22	1.17	0.31	1.22	0.31
LFCA x PP	0.23	0.22	0.13	0.32	0.07	0.35
CFLA x PP	-0.23	0.23	-0.19	0.30	-0.09	0.32
LFCA x CDTV	-0.11	0.29	0.08	0.42	0.07	0.44
CFLA x CDTV	0.15	0.30	0.21	0.42	0.19	0.43
PP x CDTV	0.14	0.20	0.14	0.28	0.11	0.29
LFCA x PP x CDTV	-0.28	0.28	-0.01	0.40	0.03	0.42
CFLA x PP x CDTV	0.43	0.29	0.60	0.39	0.52	0.41
Explicit			-0.20	0.35	-0.23	0.37
LFCA x Explicit			0.61	0.46	0.53	0.51
CFLA x Explicit			-0.17	0.49	-0.27	0.55
PP x Explicit			0.07	0.32	0.01	0.34
CDTV x Explicit			0.13	0.44	0.15	0.46
LFCA x PP x Explicit			0.17	0.44	0.39	0.48
CFLA x PP x Explicit			-0.14	0.46	-0.25	0.52
LFCA x CDTV x Explicit			-0.39	0.59	-0.28	0.63
CFLA x CDTV x Explicit			-0.09	0.61	-0.12	0.67
PP x CDTV x Explicit			-0.03	0.41	0.03	0.43
LFCA x PP x CDTV x Explicit			-0.51	0.56	-0.70	0.60
CFLA x PP x CDTV x Explicit			-0.25	0.58	-0.35	0.64
Num					0.55	0.32
LFCA x Num					0.42	0.46
CFLA x Num					0.14	0.44
PP x Num					-0.09	0.31
CDTV x Num					-0.37	0.42
Explicit x Num					0.29	0.47
LFCA x PP x Num					0.20	0.46
CFLA x PP x Num					0.04	0.41
LFCA x CDTV x Num					-1.05	0.57
CFLA x CDTV x Num					-0.06	0.56
PP x CDTV x Num					0.08	0.40

LFCA x Explicit x Num	-0.36	0.64
CFLA x Explicit x Num	0.23	0.71
PP x Explicit x Num	0.22	0.44
CDTV x Explicit x Num	-0.27	0.59
LFCA x PP x CDTV x Num	-0.15	0.57
CFLA x PP x CDTV x Num	-0.28	0.54
LFCA x PP x Explicit x Num	-0.52	0.62
CFLA x PP x Explicit x Num	0.19	0.66
LFCA x CDTV x Explicit x Num	0.87	0.79
CFLA x CDTV x Explicit x Num	-0.86	0.87
PP x CDTV x Explicit x Num	-0.25	0.56
LFCA x PP x CDTV x Explicit x Num	0.25	0.77
CFLA x PP x CDTV x Explicit x Num	-0.44	0.82

Table 9. Multivariate logistic regression analysis. N = 1662. Outcome variable is "answer to the covariance detection task, given the task version." The predictor estimates logit coefficients are labelled "Coef" and the standard error is labeled "SE." This is a binary indicator variable (0 = incorrect answer, 1 = correct answer.) The predictor estimates are logit coefficients. "LFCA," "CFLA," and "Explicit" are indicator variables reflecting randomized condition assignment (0 = unassigned, 1 = assigned). "Mixed" and "Implicit" are reference variables. "PP" and "Num" represent combined "political view" and "Num" measures respectively. "PPIV" stands for "Perception of Partisan Issue Valence," an observational indicator variable. CDTV is a indicator variable that represents the "covariance detection task version." (0 = traditional, 1 = AI.)

Study 2 Instrument

Initial Questions

What is your age?

What is your gender?

What is your highest level of educational attainment?

Which state do you live in?

In politics today, do you consider yourself a Republican, a Democrat, Independent, or something else?

(If a Democrat) Do you consider yourself a strong Democrat or a weak Democrat?

(If a Republican) Do you consider yourself a strong Republican or a weak Republican?

(If an Independent or Other) Generally speaking, do you lean towards the Democratic or Republican Party?

Did you vote in 2016?

Did you vote in 2018?

How would you describe your political views?

On a scale from cold to warm, how do you feel about ______?

- The federal government
- Silicon Valley
- Google
- Apple
- Facebook
- The Democratic Party
- The Republican Party

How are you doing economically right now?

How much have you heard about artificial intelligence (AI)?

Article

What Is AI?

Joshua Colman
The Associated Press

Over the past decade, artificial intelligence (AI) has exploded in the news. Media organizations have tracked the prospects of exciting new startups and trumpeted new products like Siri and Google Translate. Others have worried about the effects of AI advances on privacy and the economy. At the same time, some supposed "AI innovations" have turned out to be more hype than reality.

What is AI? Simply, an artificial intelligence is a human-made system that can think and take action like a person. Researchers divide AI systems into two categories: "strong AIs" that can emulate humans' reasoning, creativity, planning capacity, and common sense, and “weak AIs” that achieve human-like performance at one specialized, narrowly defined task, like recognizing faces or summarizing a document.

Currently, advancements in the field mostly target the creation of weak AIs. A complete strong AI seems to be a ways off in the future; the question of if and when that date will come is subject to furious debate in Silicon Valley.

Most recent AI developments have occurred through forms of pattern recognition using neural networks. For quite a long time, mathematicians have well understood the basic structure of neural networks, but until recently computers were not powerful enough to do the necessary computations to identify patterns using this method.

In the last decade, engineers realized that they could implement the necessary algorithms on modern graphics processing units (GPUs). GPUs were originally designed to allow the many simultaneous calculations needed to display modern computer games, but their parallel architecture also excels at neural network-based calculations to spot patterns, whether in text, numbers, sound, or images.

This breakthrough has led to rapid advancements in many areas of AI research, including speech recognition, information retrieval, automated labeling, and strategic gameplay. AI’s can write convincing symphonies and recognize human faces. Additionally, computer scientists have made rapid strides in using AI to undertake many other tasks, like preparing legal documents, doing quality control, and even driving a car, since they depend so heavily on pattern recognition.

Leading economists disagree about the likely effects of AI on working and middle-class jobs. Some claim that adoption of AI in the workplace will be uneven, and will marginally improve productivity while having a negligible impact on wages or employment. Others argue that AI technologies will fundamentally shift job descriptions and change the distribution of profits in the economy, even as overall GDP stays relatively stable.

Only time will tell which group is correct. Only one thing is for sure: the future of AI will be unpredictable.

Treatments

There will be six different treatment conditions, in a three by two design. In the first dimension, there will be differences in the way that the issues are presented. On the one dimension, the orientation of the different groups will vary, with either conservatives or liberals being for and against artificial intelligence. There will be an additional group with a more neutral background with more heterogeneous viewpoints, which will act as a control.

The second dimension will vary the way these arguments are presented. In the implicit condition, commenters will make arguments and use vocabularies that have a clear partisan valiance. However, the overall discussion will remain a civil and officially non-partisan exchange of ideas. In the other explicit condition, commentators will adopt far more aggressive language and a more zero-sum perspective that actively blames the other side. The treatment conditions are summarized in the chart below.

	Conservative Pro AI, Liberal Anti	Liberal Pro AI, Conservative Anti	Neutral Heterogeneous
Implicit	Group 1	Group 3	Group 5
Explicit	Group 2	Group 4	Group 6

I decided to use “web comments” to provide social proof instead of Facebook posts, tweets, or other possible delivery mechanisms. I believe that this will create logical continuity with the introductory article, and will aid the believability of the overall study. The different tenors of the discussions will all be plausible; to the study taker, they will reflect on the character and culture of the website which published the article.

Group 1: Conservative pro, liberal anti -- implicit

AmericanEngineer21: Luddites have always stood in the way of progress, and they've always be wrong. The free market is better placed to evaluate artificial intelligence than government bureaucrats.

SubwayFan: Finding more uses of AI will only concentrate more wealth at the top while destroying middle class jobs. This is a genuine crisis in the making.

AmericanEngineer21: You don't know that. Every technological revolution has created more jobs than it destroyed, and AI will be no different. Just because we don't know what those jobs are doesn't mean that they won't be created.

SubwayFan: Yes, and every technological revolution required heavy regulation to protect society's most vulnerable people from its effects. We're already seeing the beginnings of this with Facebook and the Cambridge Analytica scandal and the data brokers that sell our info on the dark web. We need a General Data Protection Regulation like what they have in Europe. Otherwise the problem will only get worse!

AProudWesterner: Subway Fan wants to regulate AI? That will just allow China and Russia to take first place. AI is happening, and we should ride the wave instead of being hit by it. There's a reason why the Europeans aren't leading in this field.

AmericanEngineer21: Yeah, the future of AI is too important. We can't let a Communist country like China take the lead here. We don't want AI to be shaped by an authoritarian regime.

Mom21: You can make $19 dollars an hour working from home like me! Go to jobsfrromhome.com to learn more!

Sara17: But AI and big data will bring social control mechanisms to the US; it's a natural tool of the powerful. For example Amazon uses machine learning AIs set prices as high as possible while giving people a misleading impression that their getting a good deal. Law enforcement agencies have systems to track every car by its license plate as it travels around, and then analyze that data to learn about people's daily lives.

Sara17: Would you want to be judged by AI? It's already happening – courts in Wisconsin are using proprietary AI tools to predict how likely people who are convicted of crimes will reoffend. Their prison sentences are based off this. Of course, studies have shown that this system discriminates against poor people and people of color, even when their rap sheet is the same.

AmericanEngineer21: Why wouldn't you want to use a tool like that, if it works as advertised? If it accurately predicts recidivism, it can reduce the cost of imprisonment on the taxpayer while keeping dangerous felons locked up. It would almost certainly reduce our prison population.

Sara17: How can we know if it works as advertised though? We have no way of knowing why AI's make the decisions they do. Are you comfortable being judged by a black-box machine? I'd take a human with empathy any time.

Group 2: Conservative pro, liberal anti -- explicit

AmericanPatriot21: Only a Luddite or a liberal would oppose AI development. The free market is better placed to evaluate AI than government bureaucrats!

NYCforever: Finding more uses of AI will give billionaires billions more while sticking it to the rest of use. It's trickledown economics with no trickle down. How is undercutting middle class jobs a good idea? Are you rich, or do you just like voting against your interests?

AmericanPatriot21: Every technological revolution has created more jobs than it destroyed, but you think we should trust the Democrat micromanagers who want to pick winners and losers? Just because we don't know what those jobs are doesn't mean job creators won't make them.

NYCforever: Yes, and every technological revolution required heavy regulation to protect society's most vulnerable people from its effects. We're already seeing the beginnings of this with Facebook and the Cambridge Analytica scandal, though conservatives may like to play that one down because it helped your guy. We need a General Data Protection Regulation like what they have in the EU. Otherwise corporations will just get more powerful!

Mom21: You can make $19 dollars an hour working from home like me! Go to jobsfrromhome.com to learn more!

ProudWesterner: NYCforever wants to regulate AI? That will just allow China and Russia to take first place. We need an America First approach! Burying our innovators and our military in a mountain of red tape will only hold them back from beating our adversaries, and then we won't make the rules, China will!

AmericanPatriot21: Yeah, the future of AI is too important. We can't listen to the Libs and let a Communist country like China take the lead here. We don't want AI to be shaped by an authoritarian regime.

BlueWave42: But AI will bring social control mechanisms to the US; it's a natural tool of the powerful, including big corporations. Facebook ran a study showing that they could change people's moods by manipulating their feeds, but I guess conservatives are used to that over at Fox News. Doesn't it bother you that the government has systems to track every car by its license plate as it travels around, and then analyze that data to learn about people's daily lives?

ProudWesterner: If you're so concerned about "social control," why don't you get rid of the discrimination against conservatives on social media?

BlueWave42: Would you want to be judged by AI? It's already happening – courts in Wisconsin are using proprietary AI tools to predict how likely people who are convicted of crimes will reoffend. Their prison sentences are based off this. Off course, studies have shown that this system discriminates against poor people and people of color, even when their rap sheet is the same. Just what you would expect from conservatives when they get power.

AmericanPatriot21: Why wouldn't you want to use a tool like that? If it accurately predicts recidivism, it can reduce the cost of imprisonment on the taxpayer while keeping dangerous felons locked up. Why do Libs always have to bring race in when it has nothing to do with the issue at hand?

BlueWave42: How can we know if it works as advertised though? We have no way of knowing why neural network based AI's make the decisions they do. Are you comfortable being judged by a black-box machine? I'd take a human with empathy any time.

Group 3: Liberal pro, conservative anti – implicit

49ers4dawin: AI is the way of the future, and it will radically improve economic inclusion and reduce the amount of time we spend at work. Services that were once only available to the super-wealthy are being democratized. This is exciting stuff. Let's not be nostalgic about the past.

ChestertonsFence11: In that "nostalgic past" you could get a good job right out of high school. They'd train you on the job, and you could stay in that job for your whole career and retire with a nice pension. Then they got rid of a lot of those jobs and shipped them overseas. Now they want to take the rest and replace human workers with AIs. It's the same story in my book.

AllLives43: I hear they still plan to make workers train their own replacements. Instead now it's called "providing training data" for an algorithm.

AllLives43: And you know, we're paying for it. Our tax dollars fund the university research that is driving this "AI transition." What a cruel joke.

49ers4dawin: The past wasn't so good for everyone, you know. Not if you were a woman or a person of color. Jobs didn't come so easily to them.

ProudMillennial: The true promise of AI is that is that it will flatten hierarchies and allow us to actually judge things on their merits. Sure, there's going to be more change, but change can be good. New jobs will replace old ones, and America will stay a world leader. We don't want the Russians to gain more power over us in this domain; if we don't invest, China and Russia will dominate.

Mom21: You can make $19 dollars an hour working from home like me! Go to jobsfrromhome.com to learn more!

ChestertonnsFence11: Maybe some mega cities like New York and LA will do well. Not so much in the rest of the country. I don't get why you make such a big thing about Russia, by the way. US research is driving this train, and if we turn off the spigot, this undesirable social transformation may well burn out. We should also lower taxes on companies that hire humans instead of investing in capital and machines. Put American workers first!

ProudMillennial: So you'd be happy using a foreign search engine, giving up Siri, and paying more for potato chips? (Thy actually use AI for potato sorting now!) We need to encourage, not discourage investment in these fields; the gains vastly outweigh the costs, and we can use public policy to ensure the benefits are distributed fairly.

AllLives43: That won't happen. What we need to do is break up some of these big technology cartels like Facebook. They're trying to force one lifestyle (from California) on the rest of the country. Also, the government shouldn't try to ensure that benefits are "distributed fairly." It should stop putting its thumb on the scale and let chips fall where they land.

ProudMillennial: AI is a transformative technology, and it has the potential to make society more efficient and fairer. We only need support for basic research and some lightweight regulation to make this happen. In fact, cities are already using machine learning to make smarter decisions. Why shouldn't the government take advantage of this possibility?

Group 4: Liberal pro, conservative anti – explicit

UrbanMillennial: AI is the way of the future, and it will radically improve economic inclusion and reduce the amount of time we spend at work. Services that were once only available to the super-wealthy are being democratized. This is exciting stuff. Let's not be nostalgic about an imaginary past.

RedWave77: In that "nostalgic past" you could get a good job right out of high school if you actually worked hard. They'd train you on the job, and you could stay in that job for your whole career and retire with a nice pension that you earned. Then rich elites got rid of a lot of those jobs and shipped them overseas. Now they want to take the rest and replace human workers with AIs. It's all part of the same plan.

AmericanPatriot21: I hear they still plan to make workers train their own replacements, just like they do with immigrant workers. Instead now it's called "providing training data" for an algorithm.

AmericanPatriot21: And you know, we're paying for it. Our tax dollars fund the university research that is driving this "AI transition." Get that? We pay for liberal universities to figure out how to take our jobs.

UrbanMillennial: The past wasn't so good for everyone, you know. Not if you were a woman or a person of color. Jobs didn't come so easily to them.

Mom21: You can make $19 dollars an hour working from home like me! Go to jobsfrromhome.com to learn more!

ProudLibtard51: The true promise of AI is that is that it will flatten hierarchies and allows to actually judge things on their merits. Sure, there's going to be more change, but change can be good. New jobs will replace old ones, and America will be a forward-thinking leader. If we don't invest, China and Russia will dominate, and they will keep interfering in our elections.

RedWave77: Maybe AI will help you ship jobs to elite coastal cities. Not so much for flyover country; but I guess we don't show up in your equation. I don't get why you make such a big thing about Russia, by the way. We should also lower taxes on companies that hire Americans instead of investing in robots or illegals. Trump won by promising to put American workers first, and this is the exact sort of policy elitists don't understand.

UrbanMillennial: So you'd be happy using a foreign search engine, giving up Siri, and paying more for potato chips? (Thy actually use AI for potato sorting now!) You're so busy trying to "own the libs" that you can't even see your own best economic interests. And if you feel that your local economy is being harmed by technological progress, why don't you do some free online courses to learn how to apply machine learning itself?

AmericanPatriot21: What we need to do is break up some of these big technology cartels like Facebook. They're trying to force California-style liberal speech codes on the rest of the country. Also, the government shouldn't try to ensure that benefits are "distributed fairly." That's socialism. It should stop putting its thumb on the scale and let chips fall where they land.

UrbanMillennial: AI is a transformative technology, and it has the potential to make society more efficient and fairer. In fact, cities are already using machine learning to make smarter decisions. Why shouldn't the government take advantage of this possibility? Older generations won't be around to suffer from this lack of investment, and its young people who will pay the price for falling behind on AI.

Group 5: Neutral – simple mixed views

ProudMillennial: AI is the way of the future, and it will improve economic inclusion and reduce the amount of time we spend at work. Services that were once only available to the super-wealthy are being democratized. This is exciting stuff.

YoMama41: AI is will not bring about a utopia. Without careful hands-on management, it will undermine workers bargaining power while making our society more susceptible to manipulation and control by powerful people.

SubwayFan: Yeah, I don't trust people organizations like Facebook or Amazon to behave responsibly without oversight. The slogan "move fast and break things" perfectly captures why I have concerns about them. Society doesn't need to be broken in some experimental technological crusade.

Mom21: You can make $19 dollars an hour working from home like me! Go to jobsfrromhome.com to learn more!

ProudMillennial: But if we start imposing a complex regulatory burden on AI innovators, we'll just create barriers to entry that will raise costs for startups and entrench the big players in power. Also, we'll fall behind Russia and China, who will then make the rules about how machine learning is conducted. That's the worst outcome.

SubwayFan: Well it seems like you'd rather have no rules at all.

DogWisperer12: Look, I don't think that's fair. Developing some basic rules of the road is different than having an expensive mandate like the EU's new General Data Protection Regulation.

ProudMillennial: Yeah, and AI has the potential to create highly-skilled jobs in the United States. That's exactly what we want!

YoMama41: But it will displace millions of jobs, disrupting lives and changing ways of life. And we don't know what the outcome will be down the road. We don't know that it will create jobs. It could just give a few wealthy people at the top of the pyramid more money and power.

DogWisperer12: What if we had followed the advice of skeptics when electricity was invented? Do you think the world would have been better off in that case? Big data/AI is a tool, and it has the prospect to improve all our lives. We shouldn't fear it just because there is a possibility of misuse.

Group 6: neutral –heterogeneous explicit views

49ers4dawin: AI is the way of the future, and it will radically improve economic inclusion and reduce the amount of time we spend at work. Services that were once only available to the super-wealthy are being democratized. This is exciting stuff. Let's not get stuck in traditional thinking or nostalgia about an imaginary past.

OldSchoolLiberal16: AI will not magically bring forth a utopia. Without careful hands-on management, it will undermine workers bargaining power while making our society more susceptible to manipulation and control by powerful people.

DogWisperer12: Technologies have natural tendencies, and AI's tendency is to centralize power in the hands of the already powerful. True progressives know that every technological revolution required heavy regulation to protect society's most vulnerable people from its effects. We need a General Data Protection Regulation like what they have in Europe. Otherwise the problem will only get worse!

AdorableDeplorable43: Look I'm a conservative, but, I don't trust people organizations like Facebook or Amazon to behave responsibly without oversight. Their slogan "move fast and break things" perfectly captures why I have concerns about them. Society doesn't need to be broken in some experimental technological crusade.

Mom21: You can make $19 dollars an hour working from home like me! Go to jobsfrromhome.com to learn more!

InvisibleHand5: The true conservative position is to let the free market do its work. If we start imposing a complex regulatory burden on AI innovators, we'll just raise costs for startups and entrench the big players. Also, we'll fall behind Russia and China, who will then make the rules about how machine learning is conducted. That's the worst outcome.

OldSchoolLiberal16: Well it seems like you'd rather have no rules at all. Also, the explosion of AI services and products in the marketplace isn't just a result of the free market. US government-sponsored research is driving this train. We should also raise taxes on companies that by machines instead of hiring people. Trump won by promising to put American workers first, and this is the exact sort of policy elitists don't understand.

Smallbusinessowner: Look, I don't think that's fair. Developing some basic rules of the road is different than having an expensive mandate like the EU's new General Data Protection Regulation. AI has the potential to create highly-skilled jobs in the United States.

AdorableDeplorable43: But it will displace millions of jobs, disrupting our lives and changing our way of life. And we don't know what the outcome will be down the road. This "AI revolution" will continue to cluster economic success in big coastal cities, while imposing California values on the rest of us.

49ers4dawin: What if we had followed the advice of skeptics when electricity was invented? Do you think the world would have been better off in that case? AI is a tool, and it has the prospect to improve all our lives. We shouldn't cripple it just because there is a possibility of misuse.

InvisibleHand5: Also, where's the proof that jobs will be destroyed? We need hard numbers here. As a conservative, I strongly believe that we shouldn't start regulating an industry unless there's hard data about its harms.

49ers4dawin: As a liberal, I agree. We're just speculating here. Congress needs to investigate this issue and develop policy from the facts, not develop the facts to justify a desired policy.

WashingtonIsBROKEN: Look, I'm pretty moderate, and I've voted both ways in the past. We need to keep in mind that AI "innovators" usually just find ways to capture value that already exists. We need tough rules now or the damage will be permanent.

After Experiencing the treatment, participants will be asked to interpret a study which presents an opportunity to make a the same proportion comparison mistake as in Kahan et al 2017

Economics researchers hoping to understand how artificial intelligence advances effect jobs did a case study of Alowa Dynamics, a medium-sized logistics and manufacturing firm with locations in Pennsylvania and Ohio. The CEO, Thomas Campbell, believes in giving significant autonomy to regional managers, allowing them to make different long-term capital investment decisions. He invited outside researchers to observe the results.

At the Ohio site, management has maintained a traditional workflow, with a relatively static production line that produces batches of goods. Customer service is provided by human specialists, and a human employee wraps, packs, and ships each individual order, as part of a longstanding system of quality control.

In Pennsylvania, management decided to apply the most cutting edge AI tools as soon as they became available. The assembly line is entirely automated, and job ordering is prioritized around efficiency. A new Google-powered phone and chat system handles most customer inquiries. All wrapping and packing is done by an "intelligent packer," and shipping decisions are optimized to maximize profit. Management has also created a retraining program to help displaced workers find new roles in the company.

Economists tracked the company's performance for five years. They also tracked the workers to examine the effects of these policies. While many employees stayed with Alowa, the economists made sure to follow up with any original workers that left the company. The data displayed below shows how individual workers were doing at the end of the five year period.

	Number of Better, Higher Paying Jobs	Number of Worse, Lower Paying Jobs
Ohio (traditional)	190	75
Pennsylvania (artificial intelligence)	78	22

(Note: a counterbalanced version with Ohio and Pennsylvania reversed will also be shown.)

According to this study, who did better, the workers in a traditional setting or the workers who were working in a cutting edge AI environment?

How do you feel personally about new AI technologies?

How do you think AI developments will impact you economically?

Do you think AI advancements will create more opportunities or problems for society?

Do you think AI developers are ethical people?

Do you think that AI policy is a politically polarized issue? Do you think that liberals and conservatives have different perspectives on artificial intelligence advancements?

If you had to say, do you think liberals or conservatives are more optimistic about advances in artificial intelligence?

Do you think artificial intelligence will become a big political issue in the future?

Numeracy

Suppose you have a close friend who has a lump in her breast and must have a mammogram. Of 100 women like her, 10 of them actually have a malignant tumor and 90 of them do not. Of the 10 women who actually have a tumor, the mammogram indicates correctly that 9 of them have a tumor and indicates incorrectly that 1 of them does not. Of the 90 women who do not have a tumor, the mammogram indicates correctly that 81 of them do not have a tumor and indicates incorrectly that 9 of them do have a tumor. The following table summarizes the information provided. Imagine that your friend tests positive (as if she has a tumor). What is the likelihood that she actually has a tumor?

	Has tumor	Does not have tumor
Tests positive	9	9
Tests negative	1	81

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

In the ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1000. What percent of tickets of ACME PUBLISHING SWEEPSTAKES win a car?

In the BIG BUCKS LOTTERY, the chance of winning a $10.00 prize is 1%. What is your best guess about how many people would win a $10.00 prize if 1000 people each buy a single ticket from BIG BUCKS?

Imagine that we roll a fair, six-sided die 1000 times. Out of 1000 rolls, how many times do you think the die would come up as an even number?

If the chance of getting a disease is 20 out of 100, this would be the same as having a _____% chance of getting the disease.

If the chance of getting a disease is 10%, how many people out of 1000 would be expected to get the disease?