JOURNAL OF SOCIAL COMPUTING
ISSN 2688-5255 04/06 pp221−231
Volume 4, Number 3, September 2023
DOI: 10.23919/JSC.2023.0022

Effect of Artificial Intelligence on Social Trust in American Institutions

Andrew Collins* and Jason Jeffrey Jones

Abstract: In recent decades, social scientists have debated declining levels of trust in American institutions. At the same time, many American institutions are coming under scrutiny for their use of artificial intelligence (AI) systems. This paper analyzes the results of a survey experiment over a nationally representative sample to gauge the effect that the use of AI has on the American public's trust in their social institutions, including government, private corporations, police precincts, and hospitals. We find that artificial intelligence systems were associated with significant trust penalties when used by American police precincts, companies, and hospitals. These penalties were especially strong for American police precincts and, in most cases, were notably stronger than the trust penalties associated with the use of smartphone apps, implicit bias training, machine learning, and mindfulness training. Americans' trust in institutions tends to be negatively impacted by the use of new tools. While there are significant variations in trust between different pairings of institutions and tools, generally speaking, institutions which use AI suffer the most significant loss of trust. American government agencies are a notable exception here, receiving a small but puzzling boost in trust when associated with the use of AI systems.

Key words: social trust; artificial intelligence (AI); algorithm aversion

1 Introduction

The adoption of new technologies by institutions can lead to a rise or fall in the trust individuals place in these institutions. Normalini[1], for example, found that the use of biometric verification technology increased trust in online banking in Malaysia. Conversely, introduction of opioid medications in the United States and the subsequent opioid epidemic significantly decreased Americans' trust in the pharmaceutical industry[2].

Artificial intelligence (AI)—after being moribund for a long winter—is advancing at a rapid pace. As AI-powered technologies become more common in everyday life, institutions and their leaders will need to decide whether to adopt the new capabilities present-day AI affords. However, individuals may push back against adopting AI. The use of AI-powered facial recognition technologies by police departments has already been banned in several states. This is unsurprising, as it is common for new technologies to cause panic, such as the decades-long controversy surrounding violence and video games[3] and widespread paranoia regarding 5G and COVID-19 vaccinations[4].

We hypothesized that the public would become more distrustful of institutions which used AI systems. In the present work, we adapt "trends in public attitudes about confidence in institutions" items from the General Social Survey (GSS) to implement a survey experiment. In this survey experiment, we slightly alter the wording of items to include institutions' use of a technology. Participants were randomly assigned to treatment groups to ensure that observed effects were due to the treatment rather than pre-existing differences among participants. This allows us to make strong inferences about the difference (if any) in trust expressed by respondents when leadership of an institution adopts a technology. Similar to the GSS, we poll sentiment

• Andrew Collins and Jason Jeffrey Jones are with the Department of Sociology, Stony Brook University, Stony Brook, NY 11790, USA.. E-mail: andrew.collins@stonybrook.edu.
∗ To whom correspondence should be addressed.
Manuscript received: 2023-07-29; revised: 2023-11-16; accepted: 2023-11-19

about institutional leaders to reflect trust in institutions.

1.1 Declining social trust

America's declining trust in its social institutions has been a popular topic of debate for social scientists since at least the 1990s, and yet there is still little consensus regarding the causal mechanisms underpinning this alarming trend or the extent to which something significant is even happening. Some particularly influential perspectives have emphasized the role of generational value shifts or long-term socioeconomic changes. Generational arguments emphasize the turnover of values and perspectives between older and younger generations[5–7]. These perspectives argue that trust in institutions responds to highly visible crises or scandals, such as financial crises or the 11th September terrorist attacks. It is thought that such events leave permanent impressions on younger, more impressionable cohorts. More broadly, Inglehart[5] suggested that declines in institutional trust may be attributed to a general cultural shift in the West towards postmaterialist values which place greater primacy on individualism and self-expression.

Other studies emphasize more immediate material conditions. In the US, declines in social trust have been linked to income inequality, with higher levels of income inequality connoting lower levels of social trust[8, 9]. Comparative work in political economy has also noted that countries with lower levels of income inequality tend to be characterized by higher levels of social trust: "a short answer to the question of decreased trust in the US and UK based on these studies could be that economic inequality has increased in these countries" (see Refs. [10, 11]). In fact, a substantial amount of the literature on declining social trust is comparative in nature, with many scholars taking the position that this phenomenon is a common feature of Western democracies[5, 12, 13].

Notably, there is little conclusive evidence suggesting that public esteem in government rests on the actual performance of this institution; there is often a stark contrast between public image and fact[13]. It is also possible that Americans' declining levels of trust in social institutions are an extension of Americans' declining levels of trust in each other. Paxton[6] observed that Americans' reported levels of trust in each other have been declining in tandem with declines in institutional trust. Yet, the decline in interpersonal trust has been more linear and sustained, whereas trust in institutions exhibits a "shock-and-rebound" response to popular perceptions of scandal and systemic failure. Trends in institutional trust vary significantly by institution, but the general trend across institutions appears to be negative. Some scholars have noted "moderate to strong" correlations in institutional trust between institutions which are not directly linked, such as major companies and civil services[13]. This supports the idea that Americans' trust in institutions may be responsive to broader socioeconomic or cultural trends which cannot be definitively reduced to any particular institutional failure.

Both the generational value-shift and materialist perspectives tend to take highly macroscopic analytical perspectives. A popular theoretical account of how declines in institutional trust may emerge at the individual level remained elusive until Putnam's civic participation argument[7]. According to Putnam[7], Americans' decline in civic participation denies them opportunities to establish norms of trust, reciprocity, and collaboration, leading to lower levels of social trust. Putnam[7] claimed that the technological development of society had been a major driver of declines in social trust—in particular, that television and the internet have had an "individualizing" effect on people's leisure time. Social trust is a cornerstone element of social capital as conceptualized by Putnam[7]. Putnam's social capital, which is notably distinct from the Bourdiusian concept of social capital, refers to the "connections among individuals—social networks and the norms of reciprocity and trustworthiness that arise from them". Social capital is understood as a resource that can be leveraged by the constituents of a society to coordinate collective action towards—for example—political and economic processes[7, 14, 15]. In societies with high levels of this resource, individuals readily trust those beyond their immediate contacts to behave in appropriate or reliable ways, facilitating complex networks of cooperation and governance. In a society with a severe deficit of this resource, an individual's horizon of trust may not extend very far beyond his network of personal acquaintances. In theory, declining levels of social capital pose an existential threat to the social unit.

The extent to which the decline in social trust is a problem in the United States is unclear. Social trust is a concept that is difficult to operationalize, and a lack of quality time-series data dedicated to this question has led to a heavy reliance on the GSS[15, 16]. Many of the comparative analyses rely on data from a variety of surveys. Moreover, while social trust in institutions does appear to be declining generally over time, this decline is not monotonic. There are periods of years where trust steadily increases by significant margins. For example, America's trust in government fluctuated considerably between 1958 and 2004, and in the 1990s was at its highest point since the late 1960s[13]. Finally, there is very little quality data on America's levels of social trust before the 1950s, an era of anomalous economic growth and prosperity for the US. If the materialist perspectives hold water, levels of social trust may have been unusually high at this time. In this case, it would make little sense to gauge the severity of the trust problem from our earliest data.

Whatever our point of reference, there is substantial support for the claim that levels of social trust in American institutions have been in a period of decline from at least the 1990s and have recently reached historic lows[14−16]. Given the importance of this resource for society, further research is warranted into the factors affecting Americans' trust in their institutions. While previous research on social trust has questioned the consequences of individuals' relationships with technology[7, 17], in this article, we explore the consequences of institutional use of technology. Specifically, we ask whether an institution's use of controversial new tools such as AI systems affect Americans' social trust in this institution. While much scholarship has been dedicated to characterize the decline in social trust and propose causal mechanisms for it, less work has been done to understand how trends in social trust might be compounded or attenuated when placed in the context of other factors which could influence the public perspective.

1.2 Attitude towards AI

Recent polls suggest that Americans are distrustful of emergent technologies such as AI[18]. Americans also express low confidence in the ability of their social institutions to responsibly use and develop AI. Zhang and Dafoe[19] found that only 31% of polled Americans supported the development of "high-level machine intelligence", and there are more Americans who think AI will be bad for humanity than good. Remarkably, a substantial 12% of Americans believe that high-level machine intelligence could lead to human extinction. This raises an important question: how might the use of AI systems impact Americans' trust in social institutions?

Americans' lack of faith in institutions to responsibly use and manage AI has already sparked controversy and legislative action. In 2016, an investigation by the nonprofit ProPublica reported bias against black people in AI used by Florida courts to predict criminal recidivism. In 2020, Clearview AI, a private corporation using AI for facial recognition, was the subject of extensive litigation brought by the American Civil Liberties Union for its facial recognition index which used more than 20 billion faces mined from the internet. Some states are taking preemptive action to restrict the government's use of AI, leading to bans on government use of facial recognition tools in California, Oregon, and Massachusetts.

Clearly, there is mounting suspicion and fear surrounding the use of AI by social institutions, but the nature of these fears and the conditions under which they result in backlash remain poorly understood. The extent to which the public is comfortable with using AI or living in a society in which the institutional use of AI is commonplace is still an open debate. Some have noted a phenomenon known in the literature as algorithm aversion (AV). While there is no current "best practice" for evaluating algorithm aversion, it could be defined as the expression of bias against algorithmic judgment in favor of human judgment, even when evidence clearly indicates superior or beneficial algorithmic performance[20, 21]. For example: an experiment by Dietvorst et al.[20] showed that their subjects preferred human judgment in forecasting tasks after watching the AI perform these tasks, even when their incentives were tied to accurate outcomes and the AI's performance was observably superior. They found that "people more quickly lose confidence in algorithmic than human forecasters after seeing them make the same mistake."

Other studies have observed an effect known as algorithm appreciation (AA), whereby individuals display a preference for algorithmic judgment over human judgment. For example, Logg et al.[22] present evidence suggesting that people trust algorithmic judgment more than human judgment on tasks involving certain numeric estimates or forecasting the popularity of songs and romantic attraction.

Professional forecasters, however, were the least likely to adhere to algorithmic advice, and AA waned in general when subjects had to choose between their own judgment and the algorithm's recommendations. The tension between the findings on AV and AA suggests that the public's feelings towards AI are largely contextually or socially mediated.

In a review of the literature, Jussupow et al.[21] documented the key empirical findings on AV. This review further highlights the lack of consensus on the topic of algorithmic aversion and conceptualizes AV as a phenomenon stemming from the characteristics of the human agents under study and the performance and features of the algorithms in question. To date, notable human characteristics include expertise and social distance, with perceived expertise and smaller social distance translating to more faith in the judgment provided. Algorithmic characteristics include the algorithm's agency, task performance, perceived capabilities, and the extent to which humans are known to be involved in its development and usage.

1.3 Current study

This article serves the ongoing debates regarding algorithmic aversion and social trust by being the first to study whether the American public's trust in institutions is affected by institutions' use of AI. We compare to other novel tools to check that effects are not driven simply by status-quo bias or complexity of the survey item. We hypothesized that Americans' suspicions about AI would translate to lower levels of trust in institutions when they are said to use AI. Using a survey experiment on nationally-representative samples of American adults, we compared subjects' self-reported levels of trust in American hospitals, corporations, police precincts, and government. We found evidence that Americans trust hospitals, corporations, and police precincts significantly less when they use AI tools.

2 Material and Method

2.1 Survey experiment

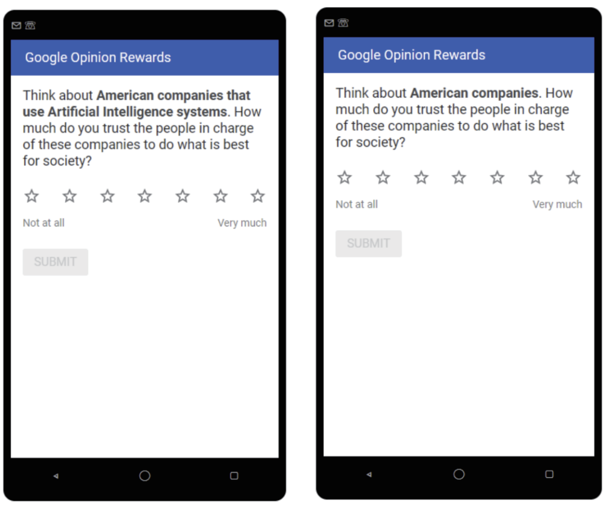
We deployed an original survey experiment using Google Surveys[23]. Respondents were asked to report trust on a 7-point Likert scale. Responses with values above 4 we consider to be representative of trust, whereas values below 4 are taken as indicative of distrust.

Similar to previous works on social trust[15, 16], the flagship question operationalizes social trust based on the General Social Survey's "confidence in institutions" items that have run for many decades. For example, the following item has been run in the GSS since the 1970s: "I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?" We hope that the familiarity of this metric in the literature on social trust will help future scholars make easy use of the results. Figure 1 shows the format of the prompts that the respondents saw.

The survey was administered in four waves approximately every three months from 2nd July 2020 to 14th April 2021. Every item was fielded as a one-item survey to a target sample size of 200 respondents. In total, we collected 18 758 responses across the range of survey items. Demographic information was collected from respondents including gender, age, and state of residence. The "survey" package in *R* was used to apply post-stratification weights to calculate population mean estimates and conduct tests for statistical significance.

2.2 Dimension of the data

In addition to measuring the public's attitude towards AI, we measure attitudes towards machine learning algorithms, smartphone apps, implicit bias training, and mindfulness training. These serve primarily for comparative purposes to validate the significance of the



(a) (b)

Fig. 1 Format of the survey questionnaire. (a) Corresponding to the tool under survey. (b) Establishing baseline values for the institution under survey.

observed effect of AI use on the American public's trust in each institution. We gather data for the following institutions: American companies, American police precincts, American government agencies, and American hospitals (Table 1).

Respondents were asked to rate their level of trust in these institutions in the context of their use of each tool (Table 2). The structure of the prompt is exemplified in the following: "Think about American companies that use Artificial Intelligence systems. How much do you trust the people in charge of these companies to do what is best for society?" In this case, the institution is "American companies," and the tool is "Artificial Intelligence systems". The survey also establishes benchmark levels of trust in each institution to serve as a control from which to estimate the effect of the tool. For example, the item "Think about American companies. How much do you trust the people in charge of these companies to do what is best for society?" was used to measure trust in the American companies institution without any tool modifier. Separate, random samples of respondents received each combination of institution and tool. All analyses estimate between-subject effects. The effect of each tool on trust in the institution was defined as the difference between the institution benchmark mean (baseline) and the mean associated with the institution plus the tool (treatment). As an example result, in the case of American police precincts, we find a 0.8-point difference between the benchmark value of trust (4.53) and the value when using AI (3.73).

Table S1 in the Appendix contains the full text and frequency of every item deployed in the surveys. All of the data and analysis code used in this article are publicly available at https://osf.io/nf7pa/.

Table 1 Counts of unique responses per institution.

Institution	Response count
American company	4337
American police precinct	4125
American government agency	1704
American hospital	1641

Table 2 Counts of unique responses per tool.

Tool	Response count
None (baseline)	5349
Artificial intelligence systems	5573
Implicit bias training	2016
Mindfulness training	1972
Machine learning algorithm	1968

2.3 Analysis and visualization

For each institution (see Table 1), a post-stratified population mean trust estimate at baseline (i.e., no tool treatment) was calculated. Then, means were calculated for each available combination of institution and tool. The "survey" package in *R*[24] was used to apply post-stratification weights. This package was also used to run t-tests to check for statistical significance, with the standard cutoff of $p < 0.05$ being the threshold for significance. Finally, the results were visualized using *R*'s "ggplot2" package[25].

3 Result and Discussion

3.1 Analysis: Mean trust by US institution

We found evidence that Americans' trust in institutions broadly declines when these institutions use artificial intelligence systems and that the size of this effect varies by institution (Fig. 2). Respondents were especially distrustful of the use of AI by American police precincts. The effect of AI use on police precincts corresponds to a 0.8-point decrease in trust, moving this institution from the "trusted" to the "distrusted" region. This was much greater than the effects observed for this institution with respect to other tools. For example: the effect of mindfulness training on trust in police precincts yielded a mere 0.03-point decrease in trust, suggesting general indifference towards this tool. See the results in Table 3.

Respondents were also significantly less trusting of American companies that used AI. The effect of AI on companies corresponded to a 0.42-point decrease in trust, compounding distrust in an institution which is already firmly below the "trusted" cutoff on baseline measures. American hospitals also became less trusted when using AI, but this effect was milder. The most unusual result was the pairing of AI with American government agencies. AI had a slight positive effect on trust in this institution, and this effect was statistically significant ($p < 0.05$). This may be because Americans trust the government even less than they trust AI. Americans' baseline levels of trust in government were the lowest among all institutions, and evidence from the literature on social trust suggests that American government institutions are often among the least trusted, with nearly half of young Americans having hardly any trust

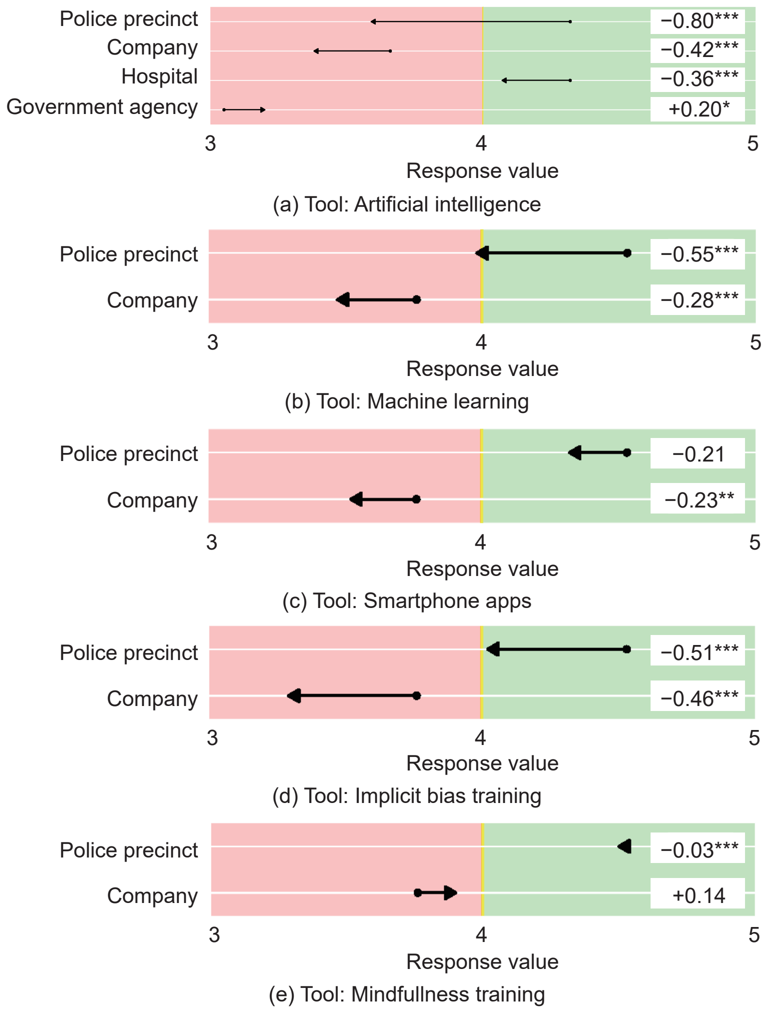


Fig. 2 Effects of tools on trust in US institutions. The response scale ranged from 1 through 7. We interpret the midpoint of 4 as the boundary between mistrust (red area) and trust (green area). Note that we truncate the axes on these plots to highlight the difference in average response between baseline and treatment. * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$.

Table 3 Mean trust values per institution at baseline and per tool treatment.

Institution	Mean trust value										
	Baseline	Treat: AI	Sig	Treat: ML	Sig	Treat: APP	Sig	Treat: BT	Sig	Treat: MT	Sig
Company	3.76	3.34	$p<0.001$	3.47	$p<0.001$	3.53	$p<0.01$	3.29	$p<0.001$	3.90	$p>0.05$
Police	4.53	3.73	$p<0.001$	3.98	$p<0.001$	4.32	$p>0.05$	4.02	$p<0.001$	4.50	$p<0.001$
Government	3.05	3.24	$p<0.05$	–	–	–	–	–	–	–	–
Hospital	4.38	4.02	$p<0.001$	–	–	–	–	–	–	–	–

Note: AI–artificial intelligence systems. APP–smartphone apps. ML–machine learning. MT–mindfulness training. BT–implicit bias training. Sig–statistical significance.

in Congress[8]. Presumably, Americans have a more concrete conceptualization of government than AI, and the ambiguity of the latter may have been leveraged by respondents towards a more favorable representation of this institution’s use of the tool.

3.2 Analysis: Mean trust by tool

In Fig. 3, we aggregated reports of trust in each tool for American police precincts and American companies, the two institutions for which data were available across every tool. With the exception of mindfulness training, respondents were not comfortable with these institutions using technology of any sort. For American police precincts, respondents were the most distrustful of AI, followed by machine learning, implicit bias training, smartphone apps, and mindfulness training. Trust in “police precincts that use smartphone apps”

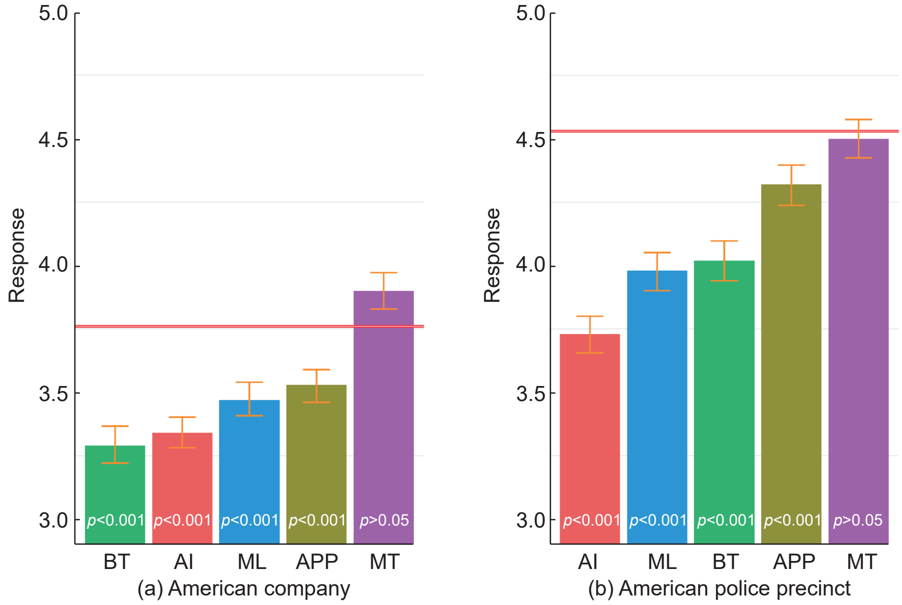


Fig. 3 Mean trust by tool for the two institutions paired with all tools. BT–implicit bias training. AI–artificial intelligence systems. APP–smartphone apps. ML–machine learning. MT–mindfulness training. The heavy red line indicates the baseline level (i.e., no using <tool> phrase) of trust in each institution. Note that the response scale ranged from 1 through 7, and the *y*-axis is truncated to allow easy comparison between mean trust at baseline and each tool treatment.

and "police precincts that use mindfulness training" did not differ reliably from baseline trust in police precincts. For American companies, respondents were least comfortable with the use of implicit bias training, followed by AI, machine learning, smartphone apps, and mindfulness training. However, trust in "companies that use mindfulness training" was not statistically significantly different from baseline levels of trust in companies.

Overall, we observed significant variability between respondents' mean levels of trust in each tool. See the results in Table 3. Americans are markedly concerned about the use of AI but also express distrust towards other new technologies. Most of these effects were statistically significant at the level of $p < 0.001$, suggesting that the sample size provided sufficient power and effects were large enough to be easily detected.

3.3 Limitation and consideration

Our survey does not impose a definition of AI. This is because it is not viable to provide respondents with a definition of AI when there is no stable public or professional consensus on this term. The concept of AI has been prominent in the cultural consciousness for almost a century, and yet the debate over how to define it is ongoing[26, 27]. More practically, the character limit of the survey instrument precluded it. (Each prompt was limited to 175 characters or less). As a result, different respondents might have conceptualized AI in different ways. That might mask variations in public attitudes towards different applications of AI. For example, one might expect differences in public attitudes toward smart weapons systems and self-driving cars. It is still useful to compare attitudes pertaining to institutional use of AI, even if it is impossible to impose a single definition of the term across respondents.

3.4 Discussion

AI has been a key theme of science-fiction since at least the mid-20th century, when super-intelligent machines began to appear in the novels of writers such as Isaac Asimov and Arthur C. Clarke. As such, the popular conception of AI has been mediated by generations of storytellers and futurists, and the impetus to tell a thrilling story has yielded many depictions of AI as a force of violence, superiority, and subjugation. Often, AI is also used in science-fiction to raise uncomfortable existential questions relating to the inevitable obsolescence of human labor or mankind's loss of control over its own destiny[28].

Americans have thus been forming opinions about AI since before the advent of the modern computer, and their opinions about this topic are likely influenced in large part by the many decades of fiction and speculation.

Given this significant cultural baggage, it is perhaps unsurprising that Americans are uneasy with the prospect of the exploitation of this technology by institutions with considerable power over their lives. These concerns are likely redoubled by growing awareness of the risks of algorithmic bias and the potential for the misuse of AI tools in an era where Americans are becoming more distrustful[14−16]. As AI becomes ubiquitous, it is more important than ever to understand Americans' attitudes towards institutions and their use of its new powers.

Our results suggest that Americans' acceptance of the use of AI varies by social institution, and may even be favorable depending on who is using it. There is significant variability in effect sizes between institutions, including differences in the direction of the effect. We should therefore be cautious when asking whether the public is appreciative or averse to algorithms; such generalizations imply a false dichotomy. Rather, we believe it is more fruitful to determine the conditions under which people are particularly appreciative or suspicious of AI. Context matters greatly for questions of trust, and this appears to be true for the range of polled tools and institutions.

Our results agree with previous research noting that Americans are anxious about AI[18, 19]. The current results suggest penalties in trust associated with the use of AI are uniquely severe. Organizational leadership and stakeholders should be considerate of these anxieties when adapting to future advancements in AI. It is possible, however, that these penalties will be attenuated as Americans' engagement with AI becomes more commonplace—as noted in Logg et al.[22], "the extent to which some domains may appear 'algorithmically appropriate' may depend on the historical use of algorithms by large numbers of people." For example, Logg et al.[22] noted that few people take issue with weather forecasting algorithms.

The fact that using AI leads to such large decreases in trust (up to 18%) for social institutions is worrying for two reasons. First, if we are indeed in a social trust crisis, the use of AI systems by social institutions might exacerbate this. The second reason it is worrying is because AI has great potential to support human decision-making. Kleinberg et al.[29] examined bail decisions made by human judges in the state of New York and compared these to decisions based on algorithmic predictions. They evaluated the decisions based on the metric of minimizing the rate of crime committed by released defendants. They found that human judges make bail decisions that are difficult to predict (noisy) and which are poorly correlated with crime risk. A decision rule based on algorithmic prediction of crime risk was demonstrated to produce two beneficial outcomes. First, policies could be chosen to either reduce crime among released defendants or to maintain the current (presumably acceptable) rate of crime while granting release to many more defendants. Second, the decisions made through policies based on algorithmic prediction reduced racial disparities in rates of jailing. It is interesting to note that the current results suggest Americans lose trust in leaders that adopt implicit bias training. Implicit bias training has been deployed widely but not been met with calls for regulation as forcefully as AI. We speculate this is due to the lingering cultural baggage of AI mentioned previously. No one yet has made a summer blockbuster featuring an out-of-control implicit bias trainer as agent of the apocalypse. Machine learning is the cornerstone of many modern AI tools.

Machine learning is a subset of AI that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed. Some popular examples of machine learning include image recognition, natural language processing, and autonomous vehicles. Therefore, the fact that machine learning had a similar effect to AI in every analysis increases our confidence in the validity of our measurements.

4 Conclusion

We find evidence that Americans' trust in institutions is significantly influenced by use of new tools. While there are significant variations in levels of trust between different pairings of institutions and tools, generally speaking, institutions which use AI suffer a significant loss of trust, with the exception of American government agencies. Police precincts and companies suffered particularly large losses of trust. In summary, we find that Americans express significantly lower levels of trust in institutions when they use artificial intelligence systems. When theorizing about social trust, this suggests a need to consider the use of technology not only at the individual level, but also at the institutional level. More work is needed to understand how public anxieties can be attenuated so that society may reap the full benefits of advances in AI.

Appendix

In Table S1, we include the full prompts that were presented to the respondents. The frequency column contains the total number of respondents who received the prompt.

In Fig. S1, we present the overall distribution of

Table S1 Full text of each prompt item.

Prompt	Frequency
How much do you trust the average American to do what is best for society?	153
Think about **academic research teams that develop new Artificial Intelligence systems**. How much do you trust the people in charge of the teams to do what is best for society?	843
Think about **academic research teams**. How much do you trust the people in charge of the teams to do what is best for society?	814
Think about **American companies that use Artificial Intelligence systems**. How much do you trust the people in charge of these companies to do what is best for society?	847
Think about **American companies that use implicit bias training for workers**. How much do you trust the people in charge of these companies to do what is best for society?	684
Think about **American companies that use machine learning algorithms**. How much do you trust the people in charge of these companies to do what is best for society?	668
Think about **American companies that use mindfulness training for workers**. How much do you trust the people in charge of these companies to do what is best for society?	646
Think about **American companies that use smartphone apps**. How much do you trust the people in charge of these companies to do what is best for society?	619
Think about **American companies**. How much do you trust the people in charge of these companies to do what is best for society?	873
Think about **American government agencies that use Artificial Intelligence systems**. How much do you trust the people in charge of the agencies to do what is best for society?	832
Think about **American government agencies**. How much do you trust the people in charge of the agencies to do what is best for society?	872
Think about **American hospitals that use Artificial Intelligence systems**. How much do you trust the people in charge of these hospitals to do what is best for society?	815
Think about **American hospitals**. How much do you trust the people in charge of these hospitals to do what is best for society?	826
Think about **American police precincts that use Artificial Intelligence systems**. How much do you trust the people in charge of these precincts to do what is best for society?	730
Think about **American police precincts that use implicit bias training**. How much do you trust the people in charge of these precincts to do what is best for society?	672
Think about **American police precincts that use machine learning algorithms**. How much do you trust the people in charge of these precincts to do what is best for society?	645
Think about **American police precincts that use mindfulness training**. How much do you trust the people in charge of these precincts to do what is best for society?	668
Think about **American police precincts that use smartphone apps**. How much do you trust the people in charge of these precincts to do what is best for society?	640
Think about **American police precincts**. How much do you trust the people in charge of these precincts to do what is best for society?	770
Think about **American research labs that create Artificial Intelligence systems**. How much do you trust the people in charge of these labs to do what is best for society?	669
Think about **American research labs that create implicit bias training for workers**. How much do you trust the people in charge of these labs to do what is best for society?	660
Think about **American research labs that create machine learning algorithms**. How much do you trust the people in charge of these labs to do what is best for society?	655
Think about **American research labs that create mindfulness training for workers**. How much do you trust the people in charge of these labs to do what is best for society?	658
Think about **American research labs that create smartphone apps**. How much do you trust the people in charge of these labs to do what is best for society?	621
Think about **American research labs that develop new Artificial Intelligence systems**. How much do you trust the people in charge of these labs to do what is best for society?	837
Think about **American research labs**. How much do you trust the people in charge of these labs to do what is best for society?	816
Think about **the average American**. How much do you trust this person to do what is best for society?	225

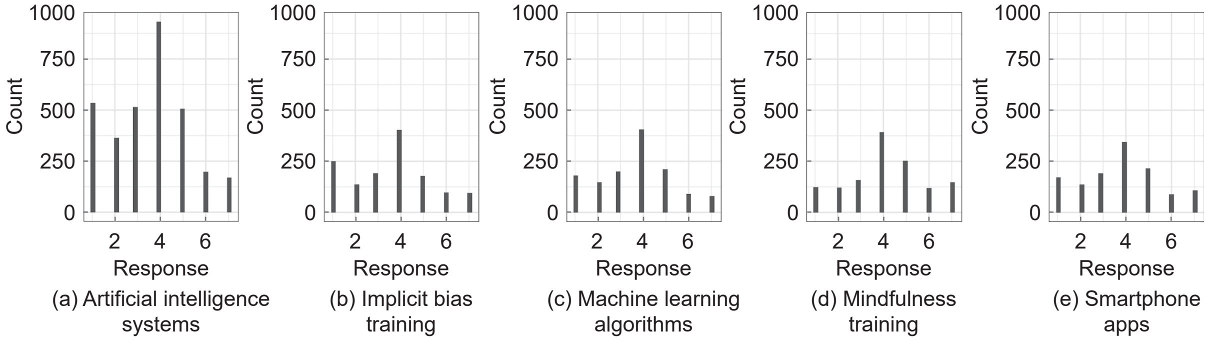


Fig. S1 Response distributions by tool.

responses per tool. Possible responses (levels 1 through 7) are on the horizontal axis, and the raw counts of respondents who chose each response level are on the vertical axis.

Acknowledgment

This work was supported by the National Science Foundation (Nos. IIS-1927227 and CCF-2208664). Jason Jeffrey Jones thanks the Center for Advanced Internet Studies for providing a fellowship in the Fall of 2023. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] M. K. T. R. Normalini, Trust in Internet banking in Malaysia and the moderating influence of perceived effectiveness of biometrics technology on perceived privacy and security, *Journal of Management Sciences*, vol. 4, no. 1, pp. 3–26, 2017.

[2] C. A. Latkin, L. Dayton, G. Yi, A. Konstantopoulos, and B. Boodram, Trust in a COVID-19 vaccine in the U. S.: A social-ecological perspective, *Soc. Sci. Med.*, vol. 270, p. 113684, 2021.

[3] P. Markey and C. Ferguson, Teaching us to fear: The violent video game moral panic and the politics of game research, *Am. J. Play.*, vol. 10, no. 1, pp. 99–115, 2017.

[4] W. Ahmed, J. Vidal-Alaball, J. Downing, and F. L. Seguí, COVID-19 and the 5G conspiracy theory: Social network analysis of twitter data, *J. Med. Internet Res.*, vol. 22, no. 5, p. e19458, 2020.

[5] R. Inglehart, Post-materialism in an environment of insecurity, *Am. Polit. Sci. Rev.*, vol. 75, no. 4, pp. 880–900, 1981.

[6] P. Paxton, Trust in decline, *Contexts*, vol. 4, no. 1, pp. 40–46, 2005.

[7] R. D. Putnam, Bowling alone: America's declining social capital, *J. Democr.*, vol. 6, no. 1, pp. 65–78, 1995.

[8] E. M. Uslaner, The moral foundations of trust, *SSRN Electron. J.*, doi: 10.2139/ssrn.824504

[9] I. Kawachi, B. P. Kennedy, K. Lochner, and D. Prothrow-Stith, Social capital, income inequality, and mortality, *Am. J. Public Health*, vol. 87, no. 9, pp. 1491–1498, 1997.

[10] C. A. Larsen, *The Rise and Fall of Social Cohesion: The Construction and De-construction of Social Trust in the US, UK, Sweden and Denmark*. Oxford, UK: Oxford University Press, 2013.

[11] O. P. Hastings, Less equal, less trusting? Longitudinal and cross-sectional effects of income inequality on trust in U. S. States, 1973–2012, *Soc. Sci. Res.*, vol. 74, pp. 77–95, 2018.

[12] A. J. G. M. Bekke and F. M. V. D. Meer, *Civil Service Systems in Western Europe*. Cheltenham, UK: Edward Elgar Publishing, 2000.

[13] S. Van de Walle, S. Van Roosbroek, and G. Bouckaert, Trust in the public sector: Is there any evidence for a long-term decline, *Int. Rev. Adm. Sci.*, vol. 74, no. 1, pp. 47–64, 2008.

[14] A. K. Clark, Rethinking the decline in social capital, *Am. Polit. Res.*, vol. 43, no. 4, pp. 569–601, 2015.

[15] J. M. Twenge, W. K. Campbell, and N. T. Carter, Declines in trust in others and confidence in institutions among American adults and late adolescents, 1972–2012, *Psychol. Sci.*, vol. 25, no. 10, pp. 1914–1923, 2014.

[16] J. Mewes, M. Fairbrother, G. N. Giordano, C. Wu, and R. Wilkes, Experiences matter: A longitudinal study of individual-level sources of declining social trust in the United States, *Soc. Sci. Res.*, vol. 95, p. 102537, 2021.

[17] B. B. Neves, Social capital and Internet use: The irrelevant, the bad, and the good, *Sociol. Compass*, vol. 7, no. 8, pp. 599–611, 2013.

[18] L. Rainie, C. Funk, M. Anderson, and A. Tyson, AI and human enhancement: Americans' openness is tempered by a range of concerns, https://www.pewresearch.org/internet/2022/03/17/ai-and-human-enhancement-americans-openness-is-tempered-by-a-range-of-concerns/, 2022.

[19] B. Zhang and A. Dafoe, U. S. public opinion on the governance of artificial intelligence, in *Proc. AAAI/ACM Conf. AI, Ethics, and Society*, New York, NY, USA, 2020, pp. 187–193.

[20] B. J. Dietvorst, J. P. Simmons, and C. Massey, Algorithm aversion: People erroneously avoid algorithms after seeing them err, *J. Exp. Psychol. Gen.*, vol. 144, no. 1, pp. 114–126, 2015.

[21] E. Jussupow, I. Benbasat, and A. Heinzl, Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion, https://ideas.repec.org/p/dar/wpaper/138565.html, 2020.

[22] J. M. Logg, J. A. Minson, and D. A. Moore, Algorithm appreciation: People prefer algorithmic to human

judgment, *Organ. Behav. Hum. Decis. Process.*, vol. 151, pp. 90–103, 2019.
[23] K. Sostek and B. Slatkin, How Google surveys works? http://g.co/surveyswhitepaper, 2017.
[24] T. Lumley, Analysis of complex survey samples, *J. Stat. Soft.*, vol. 9, no. 8, pp. 1–19, 2004.
[25] H. Wickham and G. Grolemund, R for data science, https://r4ds.had.co.nz/, 2017.
[26] P. Wang, On defining artificial intelligence, *Journal of Artificial General Intelligence*, vol. 10, pp. 1–37, 2019.
[27] D. Monett, C. W. P. Lewis, K. R. Thórisson, J. Bach, G. Baldassarre, G. Granato, I. S. N. Berkeley, F. Chollet, M. Crosby, H. Shevlin, et al., Special issue "on defining artificial intelligence"—Commentaries and author's response, *J. Artif. Gen. Intell.*, vol. 11, no. 2, pp. 1–100, 2020.
[28] I. Hermann, Artificial intelligence in fiction: Between narratives and metaphors, *AI Soc.*, vol. 38, no. 1, pp. 319–329, 2023.
[29] J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan, Human decisions and machine predictions, *Q. J. Econ.*, vol. 133, no. 1, pp. 237–293, 2018.

Andrew Collins received the MS degree in political sociology from The London School of Economics and Political Science (LSE), UK. He is currently pursuing the PhD degree at the Department of Sociology, Stony Brook University, USA. He is a computational social scientist and political sociologist interested in Science, Technology, and Society (STS), social media, and social network analysis. His current research focuses on how individuals' interactions with digital spaces inform their social identity formation, development, and expression.

Jason Jeffrey Jones received the PhD degree from University of California, San Diego, USA in 2011. He is an associate professor at the Department of Sociology and the Institute for Advanced Computational Science, Stony Brook University, USA. He leads the Computational Social Science of Emerging Realities Group and has authored 25 peer reviewed publications. His research interests include the effects of artificial intelligence on society and trends in individual identity