

---

# THE RISKS OF INDUSTRY INFLUENCE IN TECH RESEARCH

---

October 27, 2025

Joseph Bak–Coleman<sup>a,b,c,d,e,f,1</sup>, Cailin O’Connor<sup>g,j</sup>, Carl Bergstrom<sup>a,i</sup>, Jevin West<sup>h,i</sup>

<sup>a</sup>Department of Biology, University of Washington, Seattle, WA, USA

<sup>b</sup>Centre for the Advanced Study of Collective Behavior, University of Konstanz, Konstanz, Germany

<sup>c</sup>Department of Collective Behavior, Max Planck Institute of Animal Behavior, Konstanz, Germany

<sup>d</sup>Department of Biology, University of Konstanz, Konstanz, Germany

<sup>e</sup>Santa Fe Institute, Santa Fe, NM, USA

<sup>f</sup>Berkman Klein Center, Harvard University, Cambridge, MA, USA

<sup>g</sup>Department of Logic and Philosophy of Science, University of California Irvine, Irvine, CA, USA

<sup>h</sup>Information School, University of Washington, Seattle, WA, USA

<sup>i</sup>Center for an Informed Public, University of Washington, Seattle, WA, USA

<sup>j</sup>Center for Socially Engaged Philosophy, University of California Irvine, Irvine, CA, USA

<sup>1</sup>Corresponding author: jbakcoleman@gmail.com

## ABSTRACT

Emerging information technologies like social media, search engines, and AI can have a broad impact on public health, political institutions, social dynamics, and the natural world. It is critical to develop a scientific understanding of these impacts to inform evidence-based technology policy that minimizes harm and maximizes benefits. Unlike most other global-scale scientific challenges, however, the data necessary for scientific progress are generated and controlled by the same industry that might be subject to evidence-based regulation. Moreover, technology companies historically have been, and continue to be, a major source of funding for this field. These asymmetries in information and funding raise significant concerns about the potential for undue industry influence on the scientific record. In this Perspective, we explore how technology companies can influence our scientific understanding of their products. We argue that science faces unique challenges in the context of technology research that will require strengthening existing safeguards and constructing wholly new ones.

## Introduction

Digital information technologies are reshaping our world at an unprecedented pace [1]. The swift integration of these technologies into our daily lives has ignited concerns about risks to our social structures, individual well-being, public health, economic stability, democratic function, and ecological sustainability [2–7]. In response, regulators worldwide have started to develop mitigating policy based on relatively scant evidence. For these initiatives to succeed, it is crucial that the scientific community fosters a rich and impartial understanding of technology’s impacts that can serve as a bedrock for evidence-based policy.

Unfortunately, the incentives of technology industries are often misaligned with broader societal well-being. Scientific evidence about an industry’s impacts can motivate regulation that may reduce or eliminate profitability. When such conflicts arise, the affected industries may have incentives to obstruct or distort the gathering of evidence and its translation into policy [8, 9]. The past century is replete with instances where industries under threat disrupted scientific progress, often with considerable success.

To give just a few examples, automotive manufacturers have resisted safety regulations by shifting blame to the behavior of pedestrians and drivers [10]. Tobacco and asbestos companies concealed and obfuscated scientific research on the harms of their products for decades [8, 9, 11, 12]. Fossil fuel companies continue to sow doubt and impede progress on addressing climate change [8]. The soda industry funds and shares distracting narratives about the impacts of sugar on health [13–15]. We should expect much the same regarding the detrimental effects of digital information technologies like social media, as well as search engines and AI.

Studying and regulating technology companies is particularly difficult because these industries restrict access to the data needed to inform policy [16]. Moreover, these companies employ vast teams of internal researchers who operate with limited external scrutiny, with no obligations to share their research, and with research budgets that dwarf those of even the best-funded academic labs. Furthermore, norms regarding the disclosure of industry funding in technology studies are lenient, granting companies significant discretion in shaping research directions within the academy.

These factors raise serious concerns about whether current scientific norms and institutions are sufficiently robust to resist industry influence in this growing area of research. In this perspective, we argue that existing safeguards are inadequate and that industry-driven manipulation of scientific research into information technologies is already well underway. If unchecked, we believe that the accumulation of the necessary evidence to underpin effective policy will be compromised. Indeed, it already has been. We believe these harms are largely unexplored and unacknowledged by those in the relevant research areas, so this perspective is critically important in promoting high quality research in this area. We outline strategies that publishers, funders, scientists, and governments can adopt to protect and advance independent research into emerging information technologies.

## Mechanisms of Industry Influence

Policy that poses a threat to an industry’s profits should be enacted when there is clear evidence of harm and a lack of good policy alternatives [17]. Reaching this point can take time, meaning such policy tends to be reactive rather than proactive. The process typically begins with informal observation or theorizing about potential harms, directing attention toward a potential problem. Researchers then need to establish evidence of clear causal links between an industry’s conduct and undesirable outcomes. Finally, solutions must be proposed, and evidence gathered regarding their practicality and cost-effectiveness. Each link in this chain presents opportunities for industry to exert influence, and they do so in myriad ways. Many of these techniques are subtle and do not involve fraud or other overt violations of scientific norms, allowing industry to avoid risks of public censure and regulation [7, 8].

### Burying Internal Research

Perhaps the simplest, and least detectable, way in which industries impede the accumulation of evidence is by suppressing internal research that proves inconvenient. Companies are often uniquely positioned to detect harms caused by their product. Proximity to the product, involvement in its production process, and engagement with consumers provide them with relevant domain knowledge, privileged access to pertinent data, and strong incentives to address potential public relations or regulatory challenges proactively. Companies engaging in due diligence and product design will routinely uncover issues with their goods and services. For example, leaked internal documents show that Meta staffers grew concerned about children lying about their ages to use VR products designed for older people as early as 2017, and have continued to raise red flags about child safety since [18].

Sometimes, these problems are minimal or are easily remedied. Alternatively, the company may choose to go public with the issue and pay some survivable cost, such as initiating a recall. Yet in other cases, there may be no effective or financially desirable way to mitigate harm—and knowledge of such harm may even pose an existential threat to the industry. The carcinogenic properties of asbestos provide one example, where companies were among the first to notice the impacts of asbestos on lungs as cases emerged among employees [11]. Similarly, in the late 1970s, Exxon-Mobil produced compelling evidence of climate change, alongside remarkably accurate predictions of carbon emissions and global temperature rise [19]. In each of these cases, the discovery had no profit-preserving remedy, and the industries chose not to disclose their findings. The result was decades of delay, profiting the companies at a massive social cost of inaction [8, 11, 19]. Invariably, the severity of harm only become clear over time as independent research accumulated. More generally, if a company has reason even to suspect that such research may reveal existential threats, they have incentive to control or suppress it.

Unlike other industries that must decide to gather and analyze data on the impact of their products, the tech industry is unusual in that many of their products implicitly collect and analyze data—particularly about experiences and impacts on users [20]. Extracting insights from this data is an integral part of daily operations, essential for innovation and maximizing profits. In this regard, they may be more likely than most industries to identify harm in the course of routine

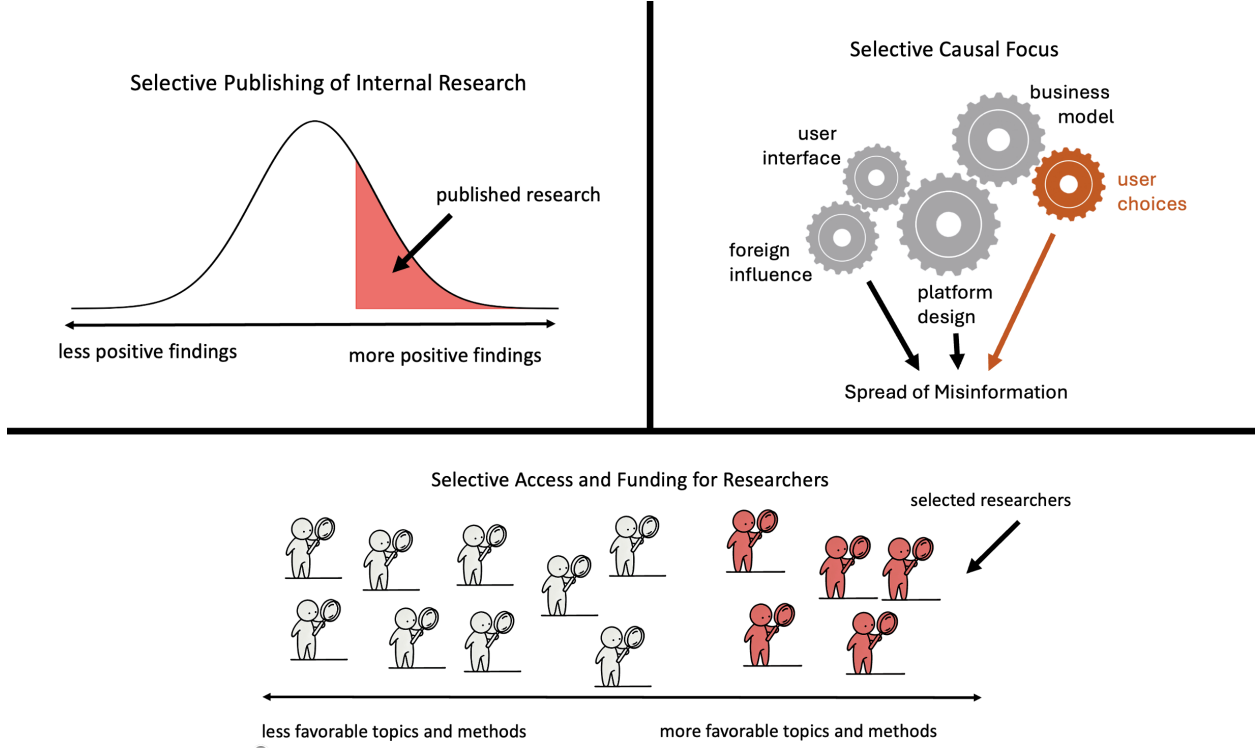


Figure 1: Mechanisms of industry influence. Selective publishing, selective causal focus, and both funding of and promotion of selected researchers are all potential mechanisms for influencing research programs, questions, results, interpretations, and eventually policy. Importantly, none of these mechanisms typically involve fraud or other legal violations.

operations. For example, content moderation undoubtedly provides direct insight into the types and frequencies of problematic posts that appear on a social media platform.

Such harms can easily remain concealed, as tech industry scientists are generally bound by extensive Non-Disclosure Agreements (NDAs) and require approval for their work to be published. It is as if, in the case of climate change, oil and gas giants were the only ones with access to data on how fossil fuels impact climate, and had convincing legal reasons to keep it private. As one might expect, these firewalls have proven effective at delaying disclosure.

There have already been a number of leaks revealing suppressed internal research into harms caused by the tech industry. These are likely indicative of a much larger volume of suppressed research. As noted above, four Meta staffers submitted a cache of internal documents to the US congress supporting their allegations that the company hid research about safety risks to children on their apps and VR devices [18]. They report that Meta deleted and edited evidence of such harms gathered in the research process, and also that Meta lawyers advised researchers not to gather sensitive data, including about children using VR devices, to avoid regulatory concerns. In response, Meta has claimed that any such deletions would have been to protect the privacy of minors.

Earlier documents leaked by Frances Haugen likewise revealed extensive internal research into Meta’s platform impacts, including harmful effects such as Instagram’s impact on teenage mental health [21, 22]. Similarly, former Google employee Timnit Gebru was fired after refusing to retract published work on issues with large language models, now widely deployed [23]. Together, these examples serve as evidence that tech companies routinely detect harm and choose not to disclose their findings, sanction employees who raise concerns, and even willfully avoid producing internal research on suspected harms.

### Selective Publishing

The flip side of undisclosed research is that companies can selectively publish research suggesting benefits or minimizing harms from their product. Companies can engage in internal review before research is made public, meaning work

demonstrating benefits will tend to have an easier time finding its way into the pages of a scientific journal (Fig. 1). This is a time-tested approach. Industry-published tobacco research, for example, painted a much rosier picture of its impacts compared to independent research [8, 24]

With this technique, companies need not rely on fabricated results; instead, they can emphasize the real benefits of their products while neglecting to publish on its harms [25, 26]. This is typically possible since most products have both positive and negative social impacts. They can also focus on research that by spurious chance shows no harm, even when such a harm exists [27]. Relatedly, industry has a history of identifying and promoting favorable independent research that happens to be produced by academic scientists [8, 14, 27, 28]. Together, these strategies create a skewed perception of the net benefits of some product while simultaneously avoiding fraud.

Meta, for example, has published work arguing that features of Facebook aid in grief following loss, improve longevity, benefit well-being, encourage reciprocity, reduce inequality, and mobilize voters in democracies [29–34]. Each of these findings may well be true, but they represent an incomplete set of potential impacts and create a biased perception of net benefits and costs. Companies producing AI products likewise have promoted their potential and prospective benefits from improving crop yields and medical diagnoses to automating scientific discovery [35–37].

Similarly, companies can selectively publish about interventions that worked to reduce their products’ harms, creating a sense that they are actively progressing towards fixing issues whether or not this is true on balance. Social media companies have embraced this strategy, publishing promising results of platform interventions to combat hate speech and misinformation, even though these interventions do little to address the issues at scale [38–40]. Historically, industry has over-emphasized the impacts of similar partially effective interventions like cigarette filters and plastic recycling [26].

Without access to internal records, it is very difficult to know when companies are selectively publishing favorable findings. Recently, a paper published by Facebook employees indicated a substantial reduction in hate speech following a platform intervention that simultaneously banned multiple key accounts [39]. This intervention was enacted in response to an uptick in hate speech during protests in 2020 [41]. As hate speech across all platforms peaked and declined during this period, though, any interventions applied in response to the uptick could spuriously appear to indicate a long-term decline. Notably, hate speech has been a problem for Facebook since its inception. One would certainly expect, and hope, that they have tried multiple strategies in response and evaluated their effectiveness. To date, however, we are not aware of any published research by Facebook on ineffective or actively harmful interventions they have tried. It is impossible for independent researchers to know whether this was the platform’s first attempt at curbing hate speech; or simply the first one that yielded positive results. In our appraisal, the latter seems more likely.

## Design Bias

Another worry about internal research—also demonstrated by the study just mentioned—relates to design bias. In the process of designing a study, researchers must make myriad choices about exactly how to carry it out—what is the exact hypothesis, what exactly will be measured, who will the subjects be, how much data will be gathered, what will the time-frame be, for example. Even if a study is carried out transparently, these choices can have significant impacts on findings, and researchers aiming at a particular finding can make the choices most likely to yield that finding (see figure 2).

Consider the ubiquitous choice of determining criteria for statistical significance and directionality of a test. Stricter criteria and adjustments for multiple comparisons will make null results more likely (and vice versa) [42]. Meta published internal research that used the lenient significance criteria of  $p < .1$  when measuring benefits in terms of well-being [31]. In the context of democratic outcomes, Meta researchers relied on conventional criteria (i.e., 95% C.I.,  $p < .05$ ) to establish benefits to turnout [29, 33]. However, collaborations examining potentially deleterious platform effects on the 2020 election relied on more conservative corrections for multiple comparisons [43–46]. For example, a 2.6% reduction in votes for Trump following withdrawal from the platform was ultimately deemed not significant following correction (Raw:  $p = 0.015$ , Adjusted:  $q = 0.076$ ) [46]. By the criteria used to gauge benefits, this identical result would have led to the conclusion that Meta’s platform did indeed alter voting patterns.

Beyond statistical significance, researchers can choose study designs biased towards finding no evidence of harm. Most notably, while randomized controlled trials are a common causal workhorse in science, they can underestimate the potential for harm in the context of social media for several key reasons [47]. First, many experiments rely on interventions where subjects take short breaks from a platform, or from some features on the platform. This is because participants actually unfamiliar with popular platforms are difficult to come by. Yet, we expect harms such as polarization, depression, and extremism to accrue over years of use. It is unlikely they should fully revert in a withdrawal experiment lasting a few weeks [47]. In addition, it is difficult to practically isolate someone from social media given its widespread use—akin to being a non-smoker in an office full of smokers. Moreover, such studies can

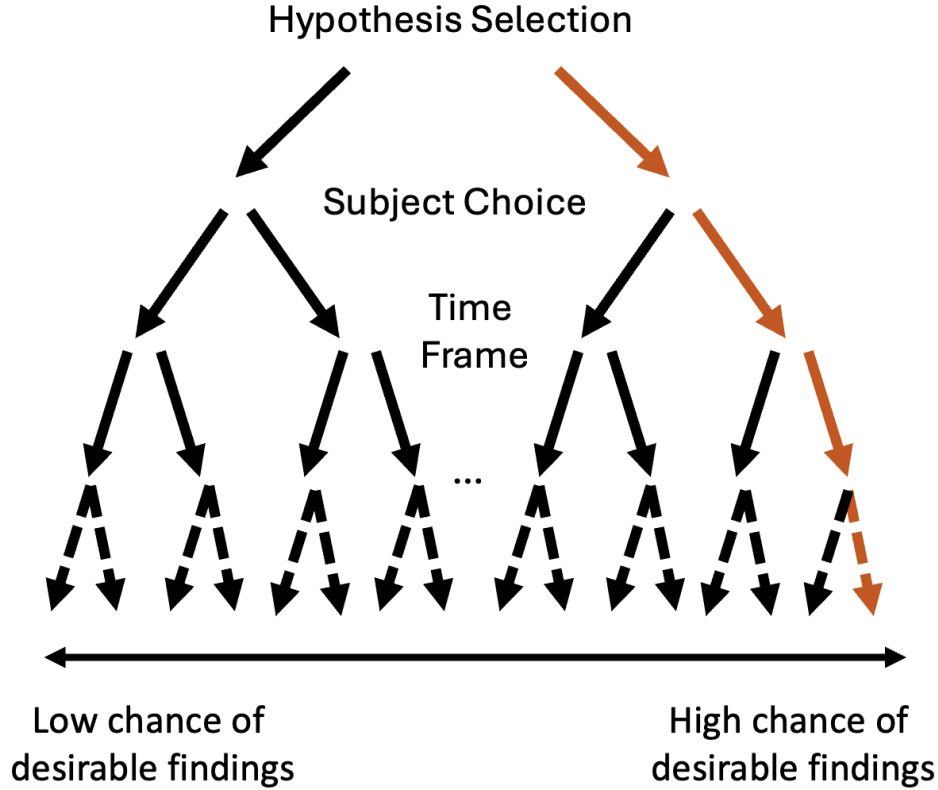


Figure 2: Design bias. Throughout the process of study researchers can make choices more likely to yield desirable findings.

only provide insight into effects on individuals which may be small or undetectable even if societal-scale effects are pronounced [47].

Likewise, the choice of study endpoints can allow for additional bias. For example X (then Twitter) examined whether crowd-sourced fact-checking could combat the spread of misinformation [38]. The authors focused on reduction in user engagement with labeled content, but measured such reductions only after a note had already been applied. The results showed meaningful effects on behavior. However, had the authors opted to measure the total effect on sharing—before and after labeling—they would have come to very different conclusion. This is because over time content shows a natural decrease in engagement, and the measured reductions occurred after the bulk of sharing had already occurred [48]. Indeed, subsequent research has shown that delays in applying notes substantially reduce their effectiveness [40].

These are just a few of the many ways design can shape study findings on tech products. Social media companies possess the capacity to iterate on study design in private. As such, even if published studies have designs that are publicly declared ahead of data collection and analysis, companies may have already selected for design features that increase chances of a desired outcome. Selectively reporting on research with biased design may also lead to proliferation of those designs in the literature since new studies tend to draw on older studies when making design choices—potentially biasing even independent future research.

### Selective Funding

Although industry has significant control over internal research, a major drawback from a company’s perspective is that this research may be viewed as less legitimate than independent research. Supporting academic researchers can increase the appearance of legitimacy while retaining some control over what research is conducted. Across domains, researchers that receive industry support have consistently produced work that better aligns with industry incentives. In medicine, for example, research funded by pharmaceutical or medical device manufacturers routinely reports higher

effectiveness [49]. Tobacco-funded scientists were less likely to find links between second-hand smoke and lung cancer [24]. And systematic reviews by researchers with funding from the soda industry are much more likely to indicate no clear link between sugar and obesity [50].

This does not imply that industry-funded researchers are engaging in ethically dubious, or industry designed, research. Science can be undermined when industry artificially props up otherwise earnest researchers whose work aligns with their incentives. This phenomenon of “industrial selection” is made possible when, from a diverse pool of scientists, industry funds just those with favorable perspectives and practices, and otherwise leaves them to their own devices [51, 52]. (See figure 1.) Given the feedback loop between funding, publishing, and career advancement, this strategy may also have long-term effects as funded scientists are promoted to positions of prestige, elevating their perspectives and practices [51, 53]. Notably, this strategy can be effective even if the funded scientists merely convince regulators that there is ongoing controversy in the relevant domain [8, 54].

A related strategy funds and shares research specifically intended to distract from industrial products as a central cause of public harm [7, 12, 55]. (See figure 1.) Coca Cola, for example, funds extensive research into the link between sedentary lifestyle and obesity [13, 14]. Tobacco similarly promoted research into the link between asbestos and lung cancer [8]. This type of “industrial distraction” or “distraction science” makes use of real, even important, research that can nonetheless mislead about industry harms by suggesting other causes of those harms [12, 26, 55].

Industrial distraction is often used to shift the blame for a product’s impacts from its design to its users. For most industrial products, public harms result from feedback between product design and user behavior. Thus, by emphasizing the causal role of the users in these harms, industry can de-emphasize the role of the product itself. For example, fossil fuel companies sponsored work on carbon footprints, and firearm lobbies have emphasized mental health rather than gun access as the central causal factor in gun violence [56–58].

Tech companies have embraced this strategy by shifting research away from the design of their technology and towards the actions and behaviors of users. Misinformation and polarization research over the past decade provides a compelling example. Meta (then Facebook) created a study which argued that users were more responsible than the algorithm for causing polarization [59]. More recent, industry produced work has argued that detrimental impacts of social media platforms are in part due to the negative mindsets individuals have about the impacts of social media, rather than the media itself [60].

Studies of helpful interventions have likewise focused on users. Google, Facebook, and Twitter all funded or facilitated research on positive “nudges” which tend to involve various forms of labeling designed to deter users from sharing undesirable content [38, 61–63]. There has been particular interest in relying on users to generate labels, increasing the perception that users can and should be responsible for preventing the sharing of misinformation.

### **Inhibiting Independent Research**

Neither selective funding and publication nor burying internal research will reliably prevent independent researchers from gathering data regarding an industry’s impacts. If compelling enough, such findings may lead to financially damaging policy. Companies need some other way to prevent this, and one powerful mechanism involves limiting the ability of independent researchers to gather such evidence.

Most social media companies provided researchers with extensive access to data via application programming interfaces (APIs) throughout the mid 2010s, but this practice was curtailed in the wake of the Cambridge Analytica scandal [16]. The restriction of such data has intensified in the intervening years, with most platforms currently offering little or no API access for research. Meta has eliminated its widely used CrowdTangle platform [64], and X has eliminated affordable academic options for data access [65]. Researchers denied data access must rely on “scraping”—collecting public-facing data through automated tools. Although there are no laws in the US against scraping, technological solutions to reduce scraping are commonplace and can make data gathering difficult [66].

For researchers that rely on API access as well as scraping, fear of losing API access may also reduce willingness to publish research using scraped data. Moreover, researchers who have gathered data directly from users have sometimes faced legal threats. Facebook sued the NYU Ad Observatory for collecting data without their authorization [67]. Elon Musk sued the Center for Countering Digital Hate after they published a paper which used scraped data to show an uptick in hate speech following his acquisition of Twitter [21].

Companies typically require researchers to adhere to a data use agreement if they wish to use data collected via APIs. While fairly standard, these instruments provide ample leverage for industry influence. For example, APIs typically require that researchers delete anything the company has itself removed [68]. Although such restrictions are justified under the auspices of preserving privacy, academic research that systematically lacks access to policy-violating and extreme content will necessarily underestimate the abundance of harmful content on platforms. Moreover, nothing

prevents a company from wielding these terms of service to prevent publication of research. In practice, such terms place sufficiently stringent demands on academic use as to place researchers in persistent legal jeopardy. For example, X sent some researchers notices requiring that they delete all existing data or pay fees of \$42,000 a month [69].

Even when companies do provide access to data, extensive delays or poor data quality can consume resources and frustrate projects. Researchers that joined a data-sharing program with Meta called Social Science One wound up waiting almost two years for the data to become available [70, 71]. Worse yet, it was only discovered after several papers were produced that the dataset was missing nearly half of the users with identified political preferences. Although there is no indication these issues were intentional, lack of incentives to share data correctly and quickly still undermines independent research. As data requests become more prominent under the European Digital Services Act, this may be a recurring issue.

Equally concerning is that researchers have no way of independently cross-checking data that social media companies provide. For example, the URL data provided by Meta to Social Science One could not be cross-checked or verified with any other data [72]. In addition, citing privacy protections, Facebook restricted researchers' data access but provided less restricted data for free to private companies [73].

Note that all this data is further limited in that it only provides insight into the behavior of users. What is missing is information about the platform, from its technological features (e.g., algorithms, spam detection) to its sociological ones (content moderation, monetization). For much of the past decade, independent researchers have been given no industry-provided tools for directly probing the function of algorithmic systems, despite the fact that such tools exist for in-house use [74]. Likewise, data gathering APIs historically have provided some access to the types of data users share or engage with, but very little insight into what users see [16]. Similarly, researchers have not been provided tools for probing the functioning of recommendation systems to evaluate their impacts. As a consequence, research has been forced to focus disproportionately on user behavior rather than platform design, or interaction between these two elements, which, as noted, may shift perceptions of responsibility away from platforms.

### **Selective Access and Performative Collaboration**

When tech companies do share data with independent researchers, they may do so selectively in order to promote industry messaging. We call this "performative collaboration"—collaborations implemented with pragmatic industry goals in mind, rather than with a primary purpose of yielding high quality research.

For example, the long-standing lack of academic access to social media data was recently relaxed through a collaboration between Meta and academic researchers. A pair of academic leads, with existing ties to Meta, were selected to work alongside Meta researchers [75]. They were free to select their academic team, provided members had previously engaged with Meta in their Social Science One project. All experimental designs had to be approved by Meta, raising the possibility of selection for design-biased research. The independent rapporteur charged with overseeing this referred to the process as "independence by permission" [75].

This kind of selective access, besides leading to design bias, may have follow-on impacts on larger research communities. Researchers selected to participate will tend to have privileged opportunities for career growth [53]. Top journals prefer to publish studies with large data sets, and these researchers have access to more data and more funding than most of their peers. It is similarly not uncommon for industry researchers to return to the academy, having published for years with unique access to data and resources.

As these industry associated researchers rise to prominence, they often become gatekeepers—such as editors and peer reviewers—in top journals. We are not suggesting they are engaging in any sort of misconduct, but rather that, in line with our discussion of industrial selection, researchers selected by industry because of their methods, topic choices, or approaches may end up disproportionately impacting the direction of technology research. Something similar happened in cardiovascular research when pharmaceuticals funded and promoted only scientists using the methods that tended to find anti-arrhythmic drugs were safe and effective. These scientists often rose to disciplinary prominence on the basis of that funding [51].

These performative collaborations can also have multiple, positive impacts for companies beyond their impacts on research. Conveying cooperativity and a sense of public concern, they can generate positive associations for the company. In this way, these collaborations can function similarly to green-washing, where companies publicize minimal environmental actions to create a positive public image [76]. As researchers in the field have commented, the Social Science One 'collaboration' following the Cambridge Analytica scandal was an attempt by Facebook to "gain positive news coverage and to reduce political pressure on the company" [73]. Throughout the history of industry influence on science, companies have likewise laundered the reputations of independent scientists in various ways, including by the simple association of academic science with their brand [7, 8]. Recently, for example, Coca Cola has been criticized for

performative academic collaborations in exercise science that both allow them to control which topics get researched and also create the impression that Coke cares about public health [13, 14].

An ongoing collaboration between the Center for Open Science and Meta reveals the degree to which Industry-Academic collaborations have shifted towards performative, data-contained, and design biased.<sup>1</sup> In this collaboration, four selected labs will be provided data about teen users’ of the platform, but are prohibited from requesting the specific content users saw, engaged with, or shared. By contrast, Meta’s internal research seems to have identified content-specific harms arising from exposure to content like that containing body standards, bullying, and unwanted sexual advances [77, 78]. For example, Meta found that teens with body image issues were exposed to feeds containing 10.5% ““eating disorder adjacent content,” compared to 3.3% for those that did not [79]. Yet this data, known to Meta, is not being provided to independent researchers. As a result, any impacts of exposure to harmful content cannot be revealed by the collaboration, resulting in design bias away from plausible sources of harm.

## Insufficiency of existing norms and institutions

### Publishing requirements

Past conflicts between science and industry have led to the development of systems intended to identify and reduce industry influence. Clinical trial registries were developed and strengthened in medicine in response to concerns over misleading presentations of biomedical research [80]. By contrast, journals publishing technological research do not typically require trial registries, even when experimentally evaluating their products’ impacts on health and well-being.

When registries are not deemed necessary, academic journals and institutions nonetheless require scientists to have their work approved by Institutional Review Boards (IRBs) which function to weigh potential risks to study subjects against benefits to science and society. This process provides important safeguards, constraining what science can be conducted. However, for reasons that are unclear to us, journals routinely waive IRB approval as a requirement for industry-produced research. For example, this has been the case for experimental research aiming to alter emotions, voting patterns, and content exposure [33, 39, 74, 81]. As a result, industry is able to produce and publish research that independent scientists ethically cannot. Notice this asymmetry also prevents independent replication of industry findings.

Competing interests declarations are another important mechanism that enables readers to appropriately reduce trust in industry influenced work [82]. However, there is evidence that tech-funded researchers are regularly failing to disclose industry funding or affiliation. Meta’s research blog identifies more than a thousand individuals to whom it has funded or awarded fellowships. (This includes only recipients of open awards—not of informal direct grants.) Yet, a search of OpenAlex indicates only 472 publications produced by independent academics acknowledging Facebook/Meta as a funding source. In writing this piece, we identified numerous instances where academics collaborating with or funded by industry failed to note these competing interests. In addition, in several cases, academics with substantial funding from a given company nonetheless served as editors and peer-reviews of research relevant to their funders.

To be clear, this lack of transparency should not be understood as a deliberate attempt to obscure questionable financial relationships. Rather, it reflects existing norms in the field coupled with a general lack of awareness around the subtle ways industry can impact the progress of research. Indeed we found several instances in which the authors disclosed industry funding yet affirmatively declared no competing interests—clearly in conflict with many journal’s policies on the matter. Our guess is that researchers often perceive their ties to industry as unimportant or irrelevant to their research choices, despite that in other areas of science comparable ties have historically had substantial impacts on research outcomes.

In sum, current publishing requirements aimed at protecting the integrity of science are not always functioning to do so when it comes to tech research.

### Open Science and related reforms

Industry-academic collaborations have leaned on “Open Science” in recent years to address concerns over transparency [75, 83, 84]. A stated goal of open science is to improve transparency through sharing of data and materials as well as *a priori* commitments to design and analysis through preregistration [83]. Yet there are reasons to think these policies cannot fully protect tech research from industry influence. On the contrary, tech companies may sometimes leverage open science to their advantage while independent researchers remain constrained by its requirements [85, 86].

---

<sup>1</sup>Details of the collaboration can be found here: <https://www.cos.io/about/news/meta-partners-with-cos-to-share-data-to-study-well-being-topics>



Preregistration, for example, involves creation of a publicly accessible, time-stamped document outlining study hypotheses and planned analyses. In 2020, a collaboration between Meta and academic scientists evaluated impacts of misleading and inflammatory content on democratic processes. They pointed to preregistrations performed as an assurance that they “didn’t hide any of the results”.<sup>2</sup> Yet it is well established that authors routinely deviate from their plans in ways that are not disclosed or difficult to identify [87]. Although deviations can be justifiable, each provides an opportunity to hide undesirable results.

The preregistered collaboration just described involved many such deviations, with one study disclosing three dozen [44]. For example, it was later revealed Meta deviated by initiating 63 secret, temporary measures targeting harmful content on their platforms. These changes were found to reduce the presence of such content during the study period [43]. By analogy, this amounts to installing temporary air filters on factories only while measuring their pollution levels. As this deviation and its impacts were not disclosed at the time of publication, it was not taken into account by reviewers and editors.

Some industry research—including an ongoing collaboration between the Center for Open Science and Meta—relies on registered reports to reduce the potential for bias. These are similar to preregistration, but involve pre-review of the study, and a guarantee of publication regardless of results. But industry can (in principle) conduct an analysis and examine results before deciding whether to collaborate with academics to submit a registered report. Without independent supervision of the full research process it is impossible to know if registered reports are properly constraining industry research in these cases. Given the publication guarantee, any obfuscation that occurs during data collection and delivery will invariably get the stamp of “peer-reviewed” research.

An additional worry is that independent researchers are restricted in their ability to preregister novel studies about online platforms. Some proponents of preregistration have argued that it is ineffective or inappropriate for existing, observed data [84, 88]. However, given limited access to data, social media researchers routinely re-analyze existing datasets. Moreover, open science advocates describe work lacking preregistration as exploratory, less credible, and not useful for making inferences to any wider population<sup>3</sup>. Taken seriously, these standards could provide grounds for dismissing the vast majority of independent research as merely exploratory intrinsically as less rigorous than industry sponsored research.

A key tenet of Open Science involves sharing of data and code [83]. Industry affiliated researchers have sometimes publicly embrace this standard. Yet companies are free to quietly publish as much or as little of their data, code and materials as they want. Selective sharing is a concern. In any case, if there is no independent appraisal of this entire pipeline from raw data gathering to publication, there is no reason to trust industry adherence to open data. And, once again, these requirements may, paradoxically, disadvantage independent researchers who are often prohibited from publicly sharing API-collected data [68].

Open science norms surrounding large sample sizes and statistical power may, likewise, advantage industry [89]. Academic sample sizes are constrained by the costs of paying participants and by ethical considerations enforced by IRBs. By contrast, researchers in the tech industry have access to millions—sometimes even billions—of users who can be studied without their knowledge, much less consent or payment. For example, research on X’s algorithmic impacts involved experiments conducted on millions of users without requiring IRB approval, participant consent, or payment [74].

Other standards for rigorous science can also disadvantage independent tech research. For example, it is a common misunderstanding that to establish causal claims in science requires manipulative studies, which are generally only possible for platforms and platform collaborators, rather than independent researchers [47, 90]. In reality, there are no shortage of techniques for establishing causality using observational data, and such approaches routinely and consistently uncover evidence of harm caused by social media [91–95]. Nevertheless, confusion over this point leads journals to prioritize on-platform experimentation. As a result, high-profile reviews and experiments have dismissed non-experimental evidence suggesting harms [43, 44, 96, 97].

Across these examples 1) industry can find ways around the standards of open science, and 2) well-intentioned reforms can become epistemic filters that select for industry-produced science. For industry, selective application of Open Science provides a means of creating a perception of unbiased, transparent work while retaining considerable potential to influence findings. For independent researchers, constraints posed by open science may be perceived as lack of transparency, rigor, or otherwise. For this reason, the otherwise laudable standards of open science may fail to preserve the integrity of tech research.

<sup>2</sup>See <https://research.facebook.com/2020-election-research/>

<sup>3</sup><https://www.cos.io/initiatives/prereg>

## Proposed Solutions

Despite the academy’s best efforts, it is clear that most of the research findings about harmful impacts from tech remain hidden as trade secrets. To protect the public, policies that bring this information to the light of day are needed. In addition, we need policies to manage the tech industry’s influence on academic research and the literature more broadly.

To be clear we are not advocating an absolute firewall between industry and academia. Collaborations with industry can provide academic researchers with access to data that would otherwise be unavailable, and are also a powerful vehicle for transferring academic insights into real-world products and policies. Rather than throwing out the wheat with the chaff, we encourage more careful attention to risks and trade-offs, alongside bulwarking norms and institutions against harmful influence in the field.

A first step is to clarify and enforce existing norms within scientific institutions. One aspect of this would require ethical approval by an independent Institutional Review Board to publish any human subjects research in the technological sphere. Currently, industry-led experiments are publishable without such approval [39]. There is little reason to exempt industry from standard journal policies simply because their data arises from experiments conducted as a part of everyday business. Beyond the obvious ethical benefits, requiring IRB approval, perhaps alongside study registration, would provide insight into the proportion of attempted research that goes unpublished by industry.

In addition, although most scientists are generally aware of the need to report financial conflicts of interest (e.g., direct payments and stocks), there appears to be variation across authors in declaring past funding, unique access to data, and collaboration as competing interests. Given apparent variation in disclosure norms, journals which have well-established policies may need to audit past research to identify undisclosed competing interests. Verification may be difficult at times, as there are few formal mechanisms in place to ensure that disclosures are accurate, or even to check if they are *post hoc*.

To avoid ambiguity going forward, journal policies should be adopted to ensure appropriate disclosures are clearly made by authors, reviewers, and editors, and to treat failures of disclosure as one might any other ethical issue in publishing. Such policies could be as simple as multiple prompts to researchers and reviewers for disclosure during the publication process and clearly communicating when such disclosures are required. As with other publication ethics lapses, journals can enforce consequences for authors (mandatory retraction), reviewers (removal from reviewer pool, or ethical investigation) and academic editors who fail to disclose conflicts of interests. To avoid risks of earnest mistakes, it may be prudent to create cross-journal databases that track COIs.

More generally, it is important that reviewers, editors, and other stakeholders are, at the very least, aware of the biasing risks of industry-affiliated research and weigh those against more obvious strengths. In our experience, we find that most are simply not aware and industry ties are rarely considered when evaluating evidence. This is particularly important when the questions posed by a study could conceivably have outcomes whose discovery is misaligned with industry incentives. Reviewers and editors should explicitly consider whether the nature of industry collaborations could be biased by internal data collection practices or design. Reviewers and editors should pay particular attention when industry scientists could have had a “first look” at the results, enabling perhaps silent deviation, deletion or abandonment of the project conditional on the findings. Absent assurances in these domains, the risk of publishing a misleading study should be weighed against any insight it may provide. In less severe cases, editors and reviewers should consider whether claims surrounding the independence and verifiability provided by Open Science mechanisms are warranted.

Of course, we should not rely on industry to be the sole producers of research on their impacts. For independent research, the constraints it operates under should be taken into account when assessing its contributions. There is no reason to prioritize potentially biased experiments over independent observational approaches to causality, nor large samples with high risk of bias over smaller independent samples [47]. Industry-imposed constraints on capacity for adhering to open science should not be penalized or viewed with suspicion. Likewise, brokers of scientific knowledge—those writing summary reports of literature on technology, or communicating key findings to policy makers—should take care to emphasize and value fully independent research when reporting their findings. For example, there is little reason for it not to be standard practice that tech journalists ask scientists if they have received industry funding or resources when working on a story related to the tech industry. Improving disclosure policy, as noted above, would be an invaluable tool for weighing independence as a strength of research in this domain—as we might consider any other methodological feature believed to substantially reduce potential bias.

Beyond leveraging existing institutions, leveling the playing field may require regulatory action. As noted, the last few years have seen an erosion of researcher access to tech data, jeopardizing the potential for even rudimentary independent research [16, 68, 75]. This trajectory needs to reverse, and the European Union is leading the way with the Digital Services Act that sets the groundwork for mandatory researcher access to data, although its effectiveness remains to be

seen. Nevertheless, this is an essential first step, and additional infrastructure will be needed to ensure these efforts are successful, and that data is accessible to researchers globally.

Data access naturally comes alongside concerns about user privacy, and infrastructure will need to be built to balance trade-offs between privacy and access. Rather than relegating these choices to companies, institutional Review Boards can play an important role, ensuring that risks are justified by benefits when conducting research on digital data. At present, much of the research conducted in this space is considered exempt based on outdated policies of a pre-internet age. An update to the Belmont report, which outlines human subjects ethics, should explicitly consider risks and opportunities of research on digital data [98].

In some cases, additional privacy-minded practices may be required to reduce risks. These could include privacy-preserving datasets, or analysis in physical or virtual “clean-rooms”—as used for U.S. census data—that ensure sensitive data is centrally maintained and not copied off premises [99]. Access to such clean rooms should be determined by independent approval boards free from competing interests, and not limited to those with whom industry wishes to associate. Although digital data will bring unique challenges, existing models of accessing sensitive biomedical data can provide a starting framework.

In cases where there is evidence that internal research has been suppressed, government requirements for industry sharing of internal research may help [100]. Such requirements may state that after some time period, internal industry tech research must be publicly released. A worry is that such requirements might disincentivize industry research in the first place. But even legal requirements to simply disclose the nature of the experiments—if not their results—will help make clear what investigations are being performed, and what is not being publicly reported. There is clear precedent for such disclosures, and their benefits, in biomedicine and little reason it cannot be extended to social media companies conducting human subjects research. If industry is only willing to produce research on the condition that it be kept private, this should be a red flag.

As noted a key limitation for independent researchers, compared to industry researchers, involves access to funding. Meta reports spending \$43.8 billion on research and development in 2024 alone, with plans to increase that spending going forward. If even a relatively small percentage of this goes to studying the human impacts of their technologies, that is still a sizable research budget for just one company. On the other hand, during the entire four years of Biden administration, only \$267 million federal funding was spent on grants studying misinformation and disinformation, some of these in the private sector. Even if independent researchers can access platform data, the resources they have available to study these platforms will never match those that tech companies can muster.

Given this discrepancy, and the serious risk that emerging technologies pose for public health and democratic function, government bodies should prioritize funding this sort of research where possible. As noted, there is a risk of feedback loops whereby industry-funded researchers are at an advantage when seeking funding from government sources [52]. Here too, disclosure can facilitate directing research funds towards fully independent research. Unfortunately recent trends in US government funding are de-emphasizing or wholly eliminating streams of research on social media and especially misinformation, which may further exacerbate the worries raised throughout this perspective.

In some areas where there are worries about industry influence in science, independent scientists have worked together to produce regular reports about their subject matter. The most notable of these is the UN Intergovernmental Panel on Climate Change (IPCC) which reports on current consensus with respect to climate change. A similar Intergovernmental Panel on Information Technology could protect the integrity of technology research both by producing an independent consensus document assessing new research in tech, and by pressuring tech companies to share data with independent researchers [101]. Reports from the IPCC have been important to policy makers seeking to protect the public from climate change, and a similar report in the tech sphere might serve as an important tool for regulators seeking to reduce the harms of new information technologies. Given existing disclosure norms described above, it would be essential that a panel of this nature surfaces and considers competing interests among surveyed research and scientists nominated to contribute.

Beyond the institutional solutions described above, awareness of the mechanisms of industry influence described above can facilitate informed decision making by academic researchers. As noted throughout, technology companies possess vast access to data, routinely run experiments, and employ large teams of scientists to make sense of the findings. For many research questions, the companies may be more than capable of conducting the research without academic involvement or may already have done so. In such cases, their goals for engaging in collaboration may be misaligned with genuine scientific curiosity. Researchers may instead be sent on performative, design-biased goose-chases or simply laundering the credibility of selected findings that align with industry interests.

For researchers who decide to precede in industry-academic collaboration, it is critical to consider and take responsibility for risks of bias imposed onto ones’ research. For example, do data-use agreements or other restrictions imposed by industry partners introduce design bias? Will the company you’re engaging with have an opportunity to get a “first peek”

at the results and decide whether to continue collaboration based on what they find? Moreover, if you are only being provided with processed data—what potential is there for data integrity issues arising during processing and how might you identify them? What data can you share, and will it allow for earnest independent critique your work? Questions like these are critical for determining whether to continue collaborating and, if so, whether there are any terms of the collaborative agreement need to be modified to ensure the final product is rigorous and unbiased.

## Conclusion

The distortion of scientific progress by industry is a common phenomena that has been documented across a wide range of disciplines and contexts. There is every reason to believe that research on digital technologies has been similarly impacted. The nature of research in these domains advantages industry in numerous ways; they have unique control of data, massive financial and scientific resources, and the capacity to wholly shield much of their product design from scientists' view. To protect inquiry in this area, and public interest more broadly, it is crucial that we find ways to mitigate industry influence in this space.

Unfortunately, few if any safeguards currently exist to protect the legitimacy of technology research. As we outline, there are practical policies that journals can easily implement that may greatly mitigate industry impact on the field. Other reforms, such as requirements for industry sharing of data and internal research, require legislation and are thus less easy to implement, but may be necessary for reliable information environments in an online age.

## Author Affiliations

Joseph Bak-Coleman Department of Biology University of Washington Life Sciences Building 351800 Seattle, WA 98195-1800 USA ORCID 0000-0002-7590-3824

Cailin O'Connor Department of Logic and Philosophy of Science University of California, Irvine 3151 Social Sciences Plaza A Irvine, CA 92697 USA ORCID 0000-0002-8351-2575

Carl T. Bergstrom Department of Biology University of Washington Life Sciences Building 351800 Seattle, WA 98195-1800 USA ORCID 0000-0002-2070-385X

Jevin D. West Information School, Box 352840 University of Washington Seattle, WA, 98195 ORCID 0000-0002-4118-0322

## Acknowledgments

A grant from Templeton World Charities (TWCF-2023-32581), awarded to J.B. and C.B supported this research. The Knight Foundation to the Center for an Informed Public also supported elements of this research.

## Delarations

J.B. has served as a paid consultant to the United Nations on the impact of digital technologies.

## References

- [1] Joseph B. Bak-Coleman, Mark Alfano, Wolfram Barfuss, Carl T. Bergstrom, Miguel A. Centeno, Iain D. Couzin, Jonathan F. Donges, Mirta Galesic, Andrew S. Gersick, Jennifer Jacquet, Albert B. Kao, Rachel E. Moran, Pawel Romanczuk, Daniel I. Rubenstein, Kaia J. Tombak, Jay J. Van Bavel, and Elke U. Weber. Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, 118(27):e2025764118, 7 2021.
- [2] Sachin Banker and Salil Khetani. Algorithm Overdependence: How the Use of Algorithmic Recommendation Systems Can Increase Risks to Consumer Well-Being:. <https://doi.org/10.1177/0743915619858057>, 38(4):500–515, 7 2019.
- [3] P K Haff. Technology as a geological phenomenon: implications for human well-being. *Geological Society, London, Special Publications*, 395(1):301–309, 2014.
- [4] Ethan Kross, Philippe Verduyn, Gal Sheppes, Cory K. Costello, John Jonides, and Oscar Ybarra. Social Media and Well-Being: Pitfalls, Progress, and Next Steps. *Trends in Cognitive Sciences*, 25(1):55–66, 1 2021.
- [5] John Zarocostas. How to fight an infodemic. *Lancet (London, England)*, 395(10225):676, 2 2020.

- [6] Matteo Cinelli, Walter Quattrociocchi, Alessandro Galeazzi, Carlo Michele Valensise, Emanuele Brugnoli, Ana Lucia Schmidt, Paola Zola, Fabiana Zollo, and Antonio Scala. The COVID-19 social media infodemic. *Scientific Reports* 2020 10:1, 10(1):1–10, 10 2020.
- [7] Cailin O’Connor and James Owen Weatherall. *The misinformation age: How false beliefs spread*. Yale University Press, New Haven, 1 2019.
- [8] Naomi Oreskes and Erik Conway. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury Press, London, 2010.
- [9] Fernández Pinto and Manuela Fernández Pinto. To Know or Better Not to. *Science & Technology Studies*, 30(2):53–72, 5 2017.
- [10] Peter D Norton. Street Rivals: Jaywalking and the Invention of the Motor Age Street. *Technology and Culture*, 48(2):331–359, 2007.
- [11] Christy A. Barlow, Jennifer Sahmel, Dennis J. Paustenbach, and John L. Henshaw. History of knowledge and evolution of occupational health and regulatory aspects of asbestos exposure science: 1900–1975. *Critical Reviews in Toxicology*, 47(4):286–316, 4 2017.
- [12] Robert N Proctor. *Golden holocaust: origins of the cigarette catastrophe and the case for abolition*. Univ of California Press, 2012.
- [13] Murray Carpenter. *Sweet and Deadly: How Coca Cola Spread Disinformation and Makes Us Sick*. MIT Press, 2025.
- [14] Susan Greenhalgh. *Soda Science: Making the World Safe for Coca-Cola*. University of Chicago Press, 2024.
- [15] Cristin E Kearns, Laura A Schmidt, and Stanton A Glantz. Sugar industry and coronary heart disease research: a historical analysis of internal industry documents. *JAMA internal medicine*, 176(11):1680–1685, 2016.
- [16] Axel Bruns. After the ‘APIcalypse’: social media platforms and their fight against critical scholarly research. <https://doi.org/10.1080/1369118X.2019.1637447>, 22(11):1544–1566, 2019.
- [17] Linda Rosenstock and Lore Jackson Lee. Attacks on Science: The Risks to Evidence-Based Policy. <https://doi.org/10.2105/AJPH.92.1.14>, 92(1):14–18, 10 2011.
- [18] Jon Swaine and Naomi Nix. Meta suppresses research on child safety, employees say, 2025.
- [19] G. Supran, S. Rahmstorf, and N. Oreskes. Assessing ExxonMobil’s global warming projections. *Science*, 379(6628), 1 2023.
- [20] Shoshana Zuboff. *The Age of Surveillance Capitalism*. Routledge, 3rd edition, 6 2023.
- [21] Bobby Allyn. 4 takeaways from Facebook whistleblower Frances Haugen’s testimony : NPR, 10 2021.
- [22] Dan Milmo. Facebook revelations: what is in cache of internal documents, 2021.
- [23] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big?, 3 2021.
- [24] Suzaynn Schick and Stanton Glantz. Scientific analysis of second-hand smoke by the tobacco industry, 1929-1972. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*, 7(4):591–612, 8 2005.
- [25] Jennifer Jacquet. *The Playbook: How to Deny Science, Sell Lies, and Make a Killing in the ...* - Jennifer Jacquet - Google Books. Penguin, London, 1 edition, 2023.
- [26] David Freeborn and Cailin O’Connor. Industrial distraction, 2025.
- [27] James Owen Weatherall, Cailin O’Connor, and Justin P Bruner. How to beat science and influence people: Policymakers and propaganda in epistemic networks. *The British Journal for the Philosophy of Science*, 71, 2020.
- [28] Stephan Lewandowsky, Toby D Pilditch, Jens K Madsen, Naomi Oreskes, and James S Risbey. Influence and seepage: An evidence-resistant minority can affect public opinion and scientific belief formation. *Cognition*, 188:124–139, 2019.
- [29] William R. Hobbs, Moira Burke, Nicholas A. Christakis, and James H. Fowler. Online social integration is associated with reduced mortality risk. *Proceedings of the National Academy of Sciences of the United States of America*, 113(46):12980–12984, 11 2016.
- [30] William R. Hobbs and Moira K. Burke. Connective recovery in social networks after the death of a friend. *Nature Human Behaviour* 2017 1:5, 1(5):1–6, 4 2017.

- [31] Moira Burke, Cameron Marlow, and Thomas Lento. Social network activity and social well-being. *Conference on Human Factors in Computing Systems - Proceedings*, 3:1909–1912, 2010.
- [32] René F. Kizilcec, Eytan Bakshy, Dean Eckles, and Moira Burke. Social influence and reciprocity in online gift giving. *Conference on Human Factors in Computing Systems - Proceedings*, 2018–April, 4 2018.
- [33] Robert M. Bond, Christopher J. Fariss, Jason J. Jones, Adam D.I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature* 2012 489:7415, 489(7415):295–298, 9 2012.
- [34] Raj Chetty, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole, and Nils Wernerfelt. Social capital II: determinants of economic connectedness. *Nature* 2022 608:7921, 608(7921):122–134, 8 2022.
- [35] Krishnamurthy (Dj) Dvijotham, Jim Winkens, Melih Barsbey, Sumedh Ghaisas, Robert Stanforth, Nick Pawlowski, Patricia Strachan, Zahra Ahmed, Shekoofeh Azizi, Yoram Bachrach, Laura Culp, Mayank Daswani, Jan Freyberg, Christopher Kelly, Atilla Kiraly, Timo Kohlberger, Scott McKinney, Basil Mustafa, Vivek Nataraajan, Krzysztof Geras, Jan Witowski, Zhi Zhen Qin, Jacob Creswell, Shravya Shetty, Marcin Sieniek, Terry Spitz, Greg Corrado, Pushmeet Kohli, Taylan Cemgil, and Alan Karthikesalingam. Enhancing the reliability and accuracy of AI-enabled diagnosis via complementarity-driven deferral to clinicians. *Nature Medicine* 2023 29:7, 29(7):1814–1820, 7 2023.
- [36] Zerina Kapetanovic, Ranveer Chandra, Tusher Chakraborty, and Andrew Nelson. FarmBeats: Improving Farm Productivity Using Data-Driven Agriculture, 2019.
- [37] Yutaro Yamada, Robert Tjarko Lange, Cong Lu, Shengran Hu, Chris Lu, Jakob Foerster, Jeff Clune, and David Ha. The ai scientist-v2: Workshop-level automated scientific discovery via agentic tree search, 4 2025.
- [38] Stefan Wojcik, Sophie Hilgard, Nick Judd, Delia Mocanu, Stephen Ragain, M. B. Fallin Hunzaker, Keith Coleman, and Jay Baxter. Birdwatch: Crowd Wisdom and Bridging Algorithms can Inform Understanding and Reduce the Spread of Misinformation, 10 2022.
- [39] Daniel Robert Thomas and Laila A. Wahedi. Disrupting hate: The effect of deplatforming hate organizations on their online audience. *Proceedings of the National Academy of Sciences of the United States of America*, 120(24):e2214080120, 6 2023.
- [40] Yuwei Chuai, Haoye Tian, Nicolas Pröllochs, and Gabriele Lenzini. Did the roll-out of community notes reduce engagement with misinformation on x/twitter? *Proceedings of the ACM on Human-Computer Interaction*, 8, 11 2024.
- [41] Cagri Toraman, Furkan Şahinuç, and Eyup Halit Yılmaz. Blacklivesmatter 2020: An analysis of deleted and suspended users in twitter, 6 2022.
- [42] Andrew Gelman and Eric Loken. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*, 348(1-17):3, 2013.
- [43] Andrew M. Guess, Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, Sandra González-Bailón, Edward Kennedy, Young Mie Kim, David Lazer, Devra Moehler, Brendan Nyhan, Carlos Velasco Rivera, Jaime Settle, Daniel Robert Thomas, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Beixian Xiong, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker. How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, 381(6656):398–404, 7 2023.
- [44] Brendan Nyhan, Jaime Settle, Emily Thorson, Magdalena Wojcieszak, Pablo Barberá, Annie Y. Chen, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, Sandra González-Bailón, Andrew M. Guess, Edward Kennedy, Young Mie Kim, David Lazer, Neil Malhotra, Devra Moehler, Jennifer Pan, Daniel Robert Thomas, Rebekah Tromble, Carlos Velasco Rivera, Arjun Wilkins, Beixian Xiong, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker. Like-minded sources on Facebook are prevalent but not polarizing. *Nature* 2023 620:7972, 620(7972):137–144, 7 2023.
- [45] Andrew M Guess, Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, et al. Reshares on social media amplify political news but do not detectably affect beliefs or opinions. *Science*, 381(6656):404–408, 2023.

- [46] Hunt Allcott, Matthew Gentzkow, Winter Mason, Arjun Wilkins, Pablo Barberá, Taylor Brown, Juan Carlos Cisneros, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Sandra González-Bailón, Andrew M. Guess, Young Mie Kim, David Lazer, Neil Malhotra, Devra Moehler, Sameer Nair-Desai, Houda Nait El Barj, Brendan Nyhan, Ana Carolina Paixao de Queiroz, Jennifer Pan, Jaime Settle, Emily Thorson, Rebekah Tromble, Carlos Velasco Rivera, Benjamin Wittenbrink, Magdalena Wojcieszak, Saam Zahedian, Annie Franco, Chad Kiewiet de Jonge, Natalie Jomini Stroud, and Joshua A. Tucker. The effects of facebook and instagram on the 2020 election: A deactivation experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 121:e2321584121, 5 2024.
- [47] Joseph B. Bak-Coleman, Stephan Lewandowsky, Philipp Lorenz-Spreen, Arvind Narayanan, Amy Orben, and Lisa Oswald. Moving towards informative and actionable social media research. *ArXiv*, 5 2025.
- [48] Joseph B. Bak-Coleman, Ian Kennedy, Morgan Wack, Andrew Beers, Joseph S. Schafer, Emma S. Spiro, Kate Starbird, and Jevin D. West. Combining interventions to reduce the spread of viral misinformation. *Nature Human Behaviour* 2022 6:10, 6(10):1372–1380, 6 2022.
- [49] Andreas Lundh, Joel Lexchin, Barbara Mintzes, Jeppe B. Schroll, and Lisa Bero. Industry sponsorship and research outcome. *The Cochrane database of systematic reviews*, 2(2), 2017.
- [50] Maira Bes-Rastrollo, Matthias B. Schulze, Miguel Ruiz-Canela, and Miguel A. Martinez-Gonzalez. Financial Conflicts of Interest and Reporting Bias Regarding the Association between Sugar-Sweetened Beverages and Weight Gain: A Systematic Review of Systematic Reviews. *PLoS Medicine*, 10(12):1–9, 2013.
- [51] Bennett Holman and Justin Bruner. Experimentation by Industrial Selection. *Philosophy of Science*, 84(5):1008–1019, 12 2017.
- [52] Manuela Fernández Pinto and Daniel Fernández Pinto. Epistemic diversity and industrial selection bias. *Synthese*, 201(5):1–18, 5 2023.
- [53] Rachel Heyard and Hanna Hottenrott. The value of research funding for knowledge creation and dissemination: A study of sns research grants. *Humanities and Social Sciences Communications*, 8(1):1–16, 2021.
- [54] Theodore M. Porter. Agnotology: The Making and Unmaking of Ignorance. *Isis*, 100(2):445–446, 6 2009.
- [55] Robert N. Proctor. *Cancer Wars: How Politics Shapes What We Know and Don’t Know About Cancer*. Basic Books, New York, 1995.
- [56] Jayme Walenta. The making of the corporate carbon footprint: the politics behind emission scoping. *Journal of Cultural Economy*, 14(5):533–548, 9 2021.
- [57] T. Higenbottam. The search for safer cigarettes. *BMJ : British Medical Journal*, 299(6706):994, 10 1989.
- [58] Emma E. McGinty, Daniel W. Webster, Marian Jarlenski, and Colleen L. Barry. News Media Framing of Serious Mental Illness and Gun Violence in the United States, 1997-2012. *American Journal of Public Health*, 104(3):406, 3 2014.
- [59] Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239):1130–1132, 2015.
- [60] *Mindsets Matter: How Beliefs About Facebook Moderate the Association Between Time Spent and Well-Being*. Association for Computing Machinery, 4 2022.
- [61] Jennifer Allen, Antonio A. Arechar, Gordon Pennycook, and David G. Rand. Scaling up fact-checking using the wisdom of crowds. *Science Advances*, 7(36), 9 2021.
- [62] Jon Roozenbeek, Alexandra L.J. Freeman, and Sander van der Linden. How Accurate Are Accuracy-Nudge Interventions? A Preregistered Direct Replication of Pennycook et al. (2020). *Psychological Science*, 32(7):1169–1178, 7 2021.
- [63] *Birds of a feather don’t fact-check each other: Partisanship and the evaluation of news in Twitter’s Birdwatch crowdsourced fact-checking program*. Association for Computing Machinery, 4 2022.
- [64] Barbara Ortutay. Meta kills off misinformation tracking tool crowdtangle despite pleas from researchers, journalists. *The Associated Press*. August, 15:2024, 2024.
- [65] Ryan Murtfeldt, Naomi Alterman, Ihsan Kahveci, and Jevin D West. Rip twitter api: A eulogy to its vast research contributions. preprint, 2024.
- [66] Melany Amarikwa. Internet openness at risk: Generative ai’s impact on data scraping. *Rich. JL & Tech.*, 30:533, 2023.
- [67] Andy Sellars. Facebook’s threat to the NYU Ad Observatory is an attack on ethical research | Nieman Journalism Lab, 10 2020.

- [68] Emma Lurie. Comparing Platform Research API Requirements, 3 2023.
- [69] Justine Calma. Twitter just closed the book on academic research , 5 2023.
- [70] Jennifer Allen, Markus Mobius, David M. Rothschild, and Duncan J. Watts. Research note: Examining potential bias in large-scale censored data. *Harvard Kennedy School Misinformation Review*, 2(4), 7 2021.
- [71] Craig Silverman. *Facebook Said It Would Give Detailed Data To Academics. They’re Still Waiting.*, 2019.
- [72] Simon Hegelich. The social science one facebook cooperation: A systemic failure. *Political Data Science Blog*, March 24, 2020.
- [73] Simon Hegelich. Facebook needs to share more with researchers. *Nature*, 579(7800):473–474, 2020.
- [74] Ferenc Huszár, Sofia Ira Ktena, Conor OBrien, Luca Belli, Andrew Schlaikjer, and Moritz Hardt. Algorithmic amplification of politics on Twitter. *Proceedings of the National Academy of Sciences of the United States of America*, 119(1), 1 2022.
- [75] Michael W. Wagner. Independence by permission. *Science*, 381(6656):388–391, 7 2023.
- [76] Sebastião Vieira de Freitas Netto, Marcos Felipe Falcão Sobral, Ana Regina Bezerra Ribeiro, and Gleibson Robert da Luz Soares. Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe*, 32(1):19, 2020.
- [77] U.S. Senate Subcommittee on Privacy Technology and the Law. Social media and the teen mental health crisis, 11 2023.
- [78] Georgia Wells, Jeff Horowitz, and Deepa Seetharaman. Facebook knows instagram is toxic for teen girls, company documents show - wsj, 9 2021.
- [79] Jeff Horowitz. Exclusive: Instagram shows more ‘eating disorder adjacent’ content to vulnerable teens, internal meta research shows | reuters, 2025.
- [80] John P.A. Ioannidis. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*, 294(2):218–228, 7 2005.
- [81] Adam D.I. I Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, 2014.
- [82] Sara Schroter, Richard Smith, Helen Barratt, Julie Morris, and Samena Chaudhry. Does the type of competing interest statement affect readers’ perceptions of the credibility of research? Randomised trial. *BMJ*, 328(7442):742–743, 3 2004.
- [83] B. A. Nosek, G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. A. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. Levy Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni. Promoting an open research culture. *Science*, 348(6242):1422–1425, 6 2015.
- [84] Brian A. Nosek, Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11):2600–2606, 3 2018.
- [85] Berna Devezer, Danielle J. Navarro, Joachim Vandekerckhove, and Erkan Ozge Buzbas. The case for formal methodology in scientific reform. *Royal Society Open Science*, 8(3), 3 2021.
- [86] Manuela Fernández Pinto. Commercial Interests and the Erosion of Trust in Science. *Philosophy of Science*, 87(5):1003–1013, 12 2020.
- [87] Aline Claesen, Sara Gomes, Francis Tuerlinckx, and Wolf Vanpaemel. Comparing dream to reality: an assessment of adherence of the first generation of preregistered studies. *Royal Society Open Science*, 8(10), 10 2021.
- [88] Daniël Lakens, Cristian Mesquida, Sajedeh Rasti, and Massimiliano Ditroilo. The benefits of preregistration and registered reports. *Evidence-Based Toxicology*, 2(1):2376046, 2024.
- [89] Daniël Lakens. Sample Size Justification. *Collabra: Psychology*, 8(1), 3 2022.
- [90] Angus Deaton and Nancy Cartwright. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210:2–21, 8 2018.
- [91] Judea Pearl. Causal inference in statistics: An overview. *Statist. Surv.*, 3(none):96–146, 1 2009.



- [92] Philipp Lorenz-Spreen, Lisa Oswald, Stephan Lewandowsky, and Ralph Hertwig. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour* 2022 7:1, 7(1):74–101, 11 2022.
- [93] Sara Abrahamsson, Aline Bütikofer, Christina Felfe, Catalina Franco Buitrago, Patrick Bennett, Laura Khoury, and Andreas Haller. Smartphone Bans, Student Outcomes and Mental Health Smartphone Bans, Student Outcomes and Mental Health \*. *NHH Dept. of Economics Discussion Paper No. 01*, 1, 2024.
- [94] Luca Braghieri, Ro’ee Levy, and Alexey Makarin. Social media and mental health. *American Economic Review*, 112(11):3660–93, November 2022.
- [95] Joshua D Angrist and Jörn-Steffen Pischke. *Mastering’metrics: The path from cause to effect*. Princeton university press, 2014.
- [96] Ceren Budak, Brendan Nyhan, David M Rothschild, Emily Thorson, and Duncan J Watts. Misunderstanding the harms of online misinformation. *Nature*, 630(8015):45–53, 2024.
- [97] National Academies of Sciences, Engineering, and Medicine and others. Social media and adolescent health. Technical report, 2023.
- [98] Icy Fresno Anabo, Iciar Elexpuru-Albizuri, and Lourdes Villardón-Gallego. Revisiting the belmont report’s ethical principles in internet-mediated research: Perspectives from disciplinary associations in the social sciences. *Ethics and Information Technology*, 21(2):137–149, 2019.
- [99] Piyush Jindal. Privacy-preserving data analysis: Implications of clean rooms. *International Journal of Management, IT & Engineering*, 14(6), 2024.
- [100] Jingyi Wu and Cailin O’Connor. How should we promote transient diversity in science? *Synthese*, 201(2):37, 2023.
- [101] Joseph Bak-Coleman, Carl T Bergstrom, Jennifer Jacquet, James Mickens, Zeynep Tufekci, and Timmons Roberts. Create an ipcc-like body to harness benefits and combat harms of digital tech. *Nature*, 617(7961):462–464, 2023.