Eduardo D. Villacis-Calderon, Tabitha L. James, and Paul Benjamin Lowry (2023). "How Facebook's newsfeed algorithm shapes childhood vaccine hesitancy: An algorithmic fairness, accountability, and transparency (FAT) perspective" ***Data and Information Management*** (D&IM) (accepted 25-May-2023) (doi: https://doi.org/10.1016/j.dim.2023.100042).

If you have any questions, would like a copy of the final version of the article, or would like copies of other articles we've published, please contact any of us directly, as follows:

- **Dr. Eduardo D. Villacis-Calderon**
  - Accounting and Information Systems
  - University of Texas, El Paso
  - Email: edvillacisc@utep.edu
  - Website: https://hb2504.utep.edu/Home/Profile?username=edvillacisc
- **Prof. Tabitha L. James**
  - Business Information Technology, Pamplin College of Business
  - Virginia Tech
  - Email: tajames@vt.edu
  - Website: https://sites.google.com/view/tabitha-l-james
- **Prof. Paul Benjamin Lowry,** Eminent Scholar and Suzanne Parker Thornhill Chair Professor
  - Business Information Technology, Pamplin College of Business
  - Virginia Tech
  - Email: Paul.Lowry.PhD@gmail.com
  - Website: https://sites.google.com/site/professorlowrypaulbenjamin/home
  - System to request Paul's articles: https://seanacademic.qualtrics.com/SE/?SID=SV_7WCaP0V7FA0GWWx

How Facebook's Newsfeed Algorithm Shapes Childhood Vaccine Hesitancy: An Algorithmic Fairness, Accountability, and Transparency (FAT) Perspective

ABSTRACT

Vaccine hesitancy is the delay or refusal of vaccination when vaccines are available. Over the last decade, many reports have suggested that the proliferation of vaccine disinformation and misinformation on social media has aggravated the vaccine-hesitancy problem. Access to vaccine dis(mis)information on social media is deemed partly responsible for the resurfacing of vaccine-preventable diseases (e.g., measles). Although studies have examined social media dis(mis)information, including that related to vaccines, the newsfeed algorithm, which determines the content social media users see, has received scant attention in the literature. We examine how people's perceptions of the fairness, accountability, and transparency (FAT) of the Facebook newsfeed algorithm influence their intention to vaccinate their children. We find that people's perceptions of the Facebook newsfeed algorithm's FAT increase their negative attitudes toward vaccination (fairness and transparency). However, they decrease users' perceptions of antivaccination norms on Facebook (fairness, accountability, and transparency). Negative attitudes toward vaccination decrease the intention to vaccinate, as do perceptions of Facebook antivaccination norms. Our findings demonstrate that to decrease the effectiveness of vaccine dis(mis)information, it is critical to educate the public about how social media newsfeed algorithms make content-display decisions.

Keywords: Vaccine hesitancy, vaccine disinformation and misinformation, Facebook newsfeed algorithm, algorithmic fairness, algorithmic accountability, algorithmic transparency, social media, theory of planned behavior (TPB)

1. Introduction

In 2015, the Centers for Disease Control and Prevention (CDC) reported a measles outbreak linked to Disney-theme-park visits, transforming "the Happiest Place on Earth" to "measles ground zero." This outbreak was linked to the presence of unvaccinated individuals visiting the theme park. There were 125 measles cases linked to the Disney outbreak, of which an alarming number of cases (28 of the 125) corresponded to individuals who intentionally avoided vaccinations due to personal beliefs [137]. Unfortunately, the Disney outbreak was followed by other measles outbreaks that experts linked to the increase of *vaccine hesitancy* [19]. In response, the World Health Organization (WHO) [130] classified the delay or refusal of vaccination (i.e., vaccine hesitancy) as one of the top ten threats to global health in 2019. Vaccine hesitancy poses a worldwide threat that may potentially affect *herd immunity*, which is the protection provided to populations by high rates of immunized individuals [43].

The proliferation of vaccine content on social media has resulted in a troublingly high percentage of people in the United States (US) being misinformed about vaccines [117]. A recent study found that of the almost 2,500 people surveyed, 18% thought it was accurate to say that vaccines caused autism, 15% agreed that vaccines are full of toxins, 20% agreed that it was acceptable to delay vaccinating their children (i.e., not follow the CDC's vaccination schedule), and 19% thought that it was better to let herd immunity develop through exposure than through widespread vaccination [117]. Distressingly, disinformation about the death of a participant in a COVID-19 vaccine trial and the use of COVID-19 vaccines to microchip people circulated online before a viable vaccine was even approved [7]. Faced with a resurgence of vaccine-preventable diseases (e.g., measles) and the introduction of newly developed vaccines whose acceptance is critical to global health (e.g., COVID-19), understanding how online dis(mis)information results in vaccine hesitancy is paramount [107]. Our study contributes to the understanding of this problem by examining the role of *newsfeed algorithms*, the technology that feeds users dis(mis)information, in inducing vaccine hesitancy.

Researchers differentiate between two types of incorrect online content (1) misinformation, which results from drawing conclusions from incorrect or incomplete content, and (2) disinformation, which is

purposefully duplicitous and intended to promote an agenda [64]. In the absence of effective content moderators, any content can be posted online regardless of its truthfulness, which leads to the proliferation on the internet of both reliable information and dis(mis)information about topics such as vaccines. In fact, one strategy of Russian internet trolls is to spread both truths and untruths about vaccination to make it appear that the safety and efficacy of vaccines are debatable [11, 16]. Information systems (IS) researchers have only recently begun to investigate online dis(mis)information. Most of the early IS research on disinformation (i.e., *fake news*) emphasized the creation of mechanisms that help users identify fake news [e.g., 71, 72, 96, 135]. Less IS research has investigated the behaviors online dis(mis)information may prompt [e.g., 24], which is the area to which our study contributes. Specifically, we examine how users' perceptions of the fairness, accountability, and transparency of the Facebook newsfeed algorithm influence their intention to vaccinate.

Researchers have suggested that disinformation about vaccines is effective for two primary reasons. First, counteracting vaccine disinformation with accurate information is difficult, because users' social-media-consumption habits result in well-segregated communities (i.e., echo chambers) of people who are either pro- or antivaccination [107, 112]. Second, vaccine disinformation lingers in memory, and researchers have found that it is difficult to update this memory with correct information [103]. These findings suggest that exposure to vaccine dis(mis)information can be harmful and that interacting with such content can lead users into echo chambers, in which they are overwhelmed with information from vaccine opponents. These phenomena are enhanced by the algorithms that run social media *newsfeeds*, which determine the content to which users are exposed [94]. Facebook's newsfeed prioritizes "engaging" content, which can be loosely defined as content people will read, "like," comment on, or share with others [27, 41].

Although some general information about how newsfeed algorithms work is available [27, 41], there is remarkably little public knowledge about how these algorithms determine the content people see [57]. This lack of understanding has recently become a topic of public conversation because, facing pressure about security concerns from the US, TikTok released some details about how its social media feed works [57]. Like other social media companies, TikTok prioritizes engaging content, and it described how engagement-

related factors help determine what is displayed in users' newsfeeds. However, the details TikTok released were familiarly vague; as one reporter noted, some of the "factors matter more than others; and they're all in service of a secret technical sauce that the platform uses to make the best guess as to what it predicts users will want to see" [57, p. 1]. The same can be said for most newsfeed algorithms.

The details of the logic newsfeed algorithms use to determine the content users see are largely unknown and proprietary. Therefore, it is up to users to evaluate whether (1) the logic the newsfeed algorithm employs to display the content they see is fair, (2) Facebook can be held to account for how the newsfeed operates, and (3) they understand how the algorithm determines the content they see. Because the details of the algorithm's logic are unknown, it is important to understand what users think about the FAT of newsfeed algorithms and how those perceptions influence the effectiveness of the dis(mis)information the algorithms convey. Although the FAT framework has been used in the broader artificial intelligence (AI) literature [e.g., 1], our study is unique because it examines the concepts from the perspective of the interaction between general users and a social media newsfeed algorithm. Understanding users' perceptions of algorithmic FAT and the influence those perceptions have on the attitudes and beliefs that may arise from exposure to dis(mis)information on social media is a crucial step of learning how to combat social media dis(mis)information. In this study, we thus examine the following research question.

> **RQ1:** How do users' perceptions of the fairness, accountability, and transparency (FAT) of Facebook's newsfeed algorithm influence the effectiveness of the dis(mis)information about childhood vaccination conveyed on Facebook?

We use the *theory of planned behavior* (TPB) [3] to examine the effectiveness of vaccine dis(mis)information on Facebook, because it is a leading model of behavioral intention and has been used in many contexts, including vaccination [e.g., 2, 18, 129]. Briefly, the TPB posits that people's attitudes toward a behavior, the perceived social norms regarding the behavior, and the perceived control people have to execute the behavior influence their intentions to perform the behavior [3]. We propose that Facebook users' perceptions of the newsfeed algorithm's FAT will shape the TPB factors and thus influence behavioral intent. Consider users whose see more antivaccination than pro-vaccination content in their Facebook newsfeed. If those users think the Facebook newsfeed is fair regarding the logic it uses to

determine the content displayed, they may be more likely to believe that their social connections on Facebook hold antivaccination norms.

Our study makes important contributions to both theory and practice. Researchers have only begun to study the implications of FAT perceptions for algorithms [e.g., 113], and to the best of our knowledge, they have not examined their influence on intention to perform an offline behavior like vaccination. We contribute to the IS literature by extending the conversation on dis(mis)information to include how algorithmic FAT, or lack thereof, is influencing society. Although we draw from a theoretical base familiar to IS researchers, we extend the TPB in a novel way to examine how social media, and the algorithms that power it, are influencing the intention to perform important offline behaviors like vaccination. We also contribute to the IS research on the persuasiveness of social media, which has investigated how social media content persuades people to make purchases and take political action [e.g., 49, 73]. Comparing advice from algorithms (e.g., recommender systems) and advice from peers (e.g., reviews), IS research has found that both affect users' decision-making [51]. Our study's contribution to this conversation results from its unique combination of the two—the content is posted by users, but the determination of what is shown is left to the algorithms. By studying FAT perceptions in the vaccine dis(mis)information context, we begin the conversation about how people's understanding of the merging of these two elements can influence decision-making.

The practical implications of our findings are notable because we make a case for the importance of societal algorithmic awareness. If people do not understand how the algorithms manipulate the content they see to maximize their engagement, they are more likely to be captured and influenced by social media echo chambers. Our study offers insight into how social media shapes users' vaccination intentions; we offer the conclusion that rather than trying to correct attitudes derived from vaccine dis(mis)information, a more fruitful avenue may be to educate users on how newsfeeds work. This approach may decrease users' confidence in social media content without directly confronting them on an issue they feel strongly about (e.g., vaccination). We also provide justification for removing vaccine dis(mis)information, because as users' FAT perceptions increase, their negative attitudes toward vaccination increase, which lowers

intention to vaccinate. Therefore, users who express confidence in the newsfeed's FAT and who are exposed to vaccine dis(mis)information may be more prone to developing a negative attitude toward vaccination. Although algorithmic fairness and transparency increase negative attitudes toward vaccination, which in turn decreases intention to vaccinate, FAT perceptions decrease normative perceptions that Facebook social connections are antivaccination, which in turn decreases intention to vaccinate. These results may indicate that although the dis(mis)information users see on Facebook can be persuasive and it is thus best to limit exposure, it may be equally important to ensure that users avoid echo chambers and that newsfeeds work in a transparent way.

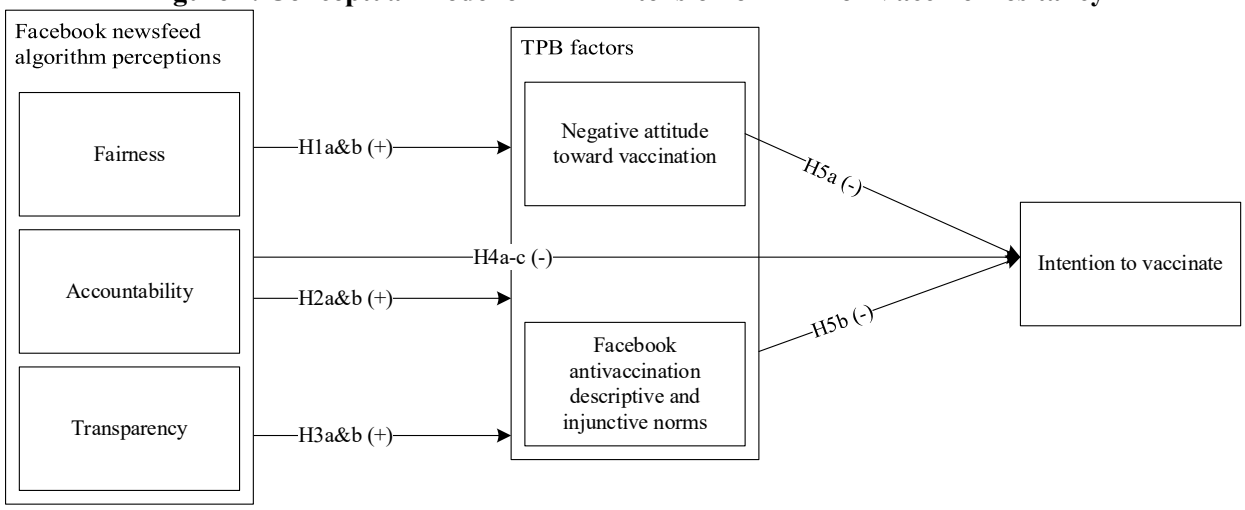
2. Theoretical Development

Our model conceptualizes users' FAT perceptions of Facebook's newsfeed algorithm as antecedents to a traditional TPB model, as shown in Figure 1. We test our model on Facebook users who are connected to more vaccine opponents than proponents. We survey people who have (1) at least one Facebook social connection who has posted antivaccination content and (2) who have more Facebook social connections who post antivaccination content than pro-vaccination content. Vaccine dis(mis)information has been found to be sticky [103], which means that once such dis(mis)information is seen and believed, it is difficult to correct [81]. If vaccine dis(mis)information (i.e., antivaccine content) is effective, people who are connected to more vaccine opponents than proponents are the most likely to be persuaded to not vaccinate their children. We propose that the more fair, accountable, and transparent users perceive the Facebook newsfeed algorithm to be, the more likely users who have been exposed mainly to antivaccination content will be to (1) have a negative attitude toward vaccines and (2) believe their Facebook social connections hold antivaccination descriptive and injunctive social norms. We expect that as users' negative attitudes toward vaccination increase and perceptions of Facebook antivaccination descriptive and injunctive norms increase, users' intention to vaccinate their children will decrease.

2.1 The Theory of Planned Behavior (TPB)

We leverage the TPB in our model of vaccine hesitancy because it is "one of the most influential theories in explaining and predicting behavior" [101, p. 117]. We suggest that the TPB's focus on attitude, beliefs,

Figure 1. Conceptual Model of FAT Extension of TPB for Vaccine Hesitancy



intentions, and social norms, which are natural factors of healthcare decisions, make it particularly useful in explaining vaccine hesitancy. The TPB has been demonstrated to be a valuable theoretical framework for studying a variety of health-related behaviors (e.g., exercise, drug consumption); for reviews, see Downs and Hausenblas [34] and McEachan et al. [90]. Researchers have used the TPB to study intention to vaccinate for human papilloma virus (HPV) [e.g., 10, 18, 78] and influenza [e.g., 2, 28, 48], and intention to vaccinate children [e.g., 35, 62, 129].

The TPB proposes that behavior can be predicted by understanding the different beliefs and attitudes that shape *behavioral intention*: the individual's readiness to perform a determined behavior [45]. The TPB builds on the premise that "the stronger the intention to engage in a behavior, the more likely should be its performance" [3, p. 181]. It proposes that *attitudes*, *norms*, and *perceived behavioral control* shape behavioral intention. Attitude is a personal "latent disposition or tendency to respond with some degree of favorableness or unfavorableness" [45, p. 76] toward an attitudinal object (e.g., object or behavior). A norm is a "specific behavioral prescription or proscription attributed to a generalized social agent" [45, p. 131]. The TPB originally proposed the concept of *subjective norm*, which indicates whether a person believes important others think they should or should not perform a behavior. However, Fishbein and Ajzen [45, p. 131] maintained "that normative prescriptions represent only one source of perceived normative pressure" and that "in addition to believing that particular individuals do or do not want us to perform a given

behavior, we may also experience normative pressure because we believe that important others are themselves performing or not performing the behavior in question." Cialdini et al. [23] termed the former "injunctive norms" and the latter "descriptive norms." Injunctive norms are what a person believes the referent group (e.g., Facebook social connections) thinks they should do with respect to a particular behavior (e.g., not vaccinating their children). By contrast, descriptive norms are what a person believes the referent group is actually doing with respect to the behavior (e.g., their Facebook social connections are actually declining to vaccinate their children). We examine both injunctive and descriptive norms because they have been shown to be distinct in other health-related activities [100]. Perceived behavioral control captures "the degree to which an individual actually has control over performing the behavior" [45, p. 64].

The TPB proposes that the more negative an individual's attitude toward a behavior, the less likely the individual will be to perform it. Therefore, people who have a negative attitude toward vaccination will be less likely to vaccinate their children. Perceptions that a behavior is normative will increase the likelihood that people will perform it. Specifically, if vaccine dis(mis)information is effective, people who believe their Facebook social connections think they should not vaccinate their children (an antivaccination injunctive norm) and who believe their Facebook social connections are not vaccinating their own children (an antivaccination descriptive norm) will be less likely to vaccinate their own children. Finally, the TPB proposes that the more control individuals feel they have over the performance of the behavior, the more likely they will be to perform it. It is unlikely that disinformation will have a significant influence on parents' perceived control over their children's vaccination, thus we retain this TPB factor as a control variable in our model. We do not hypothesize any associations between it and the FAT perceptions.. Our study extends the TPB to examine the influence of FAT perceptions of the Facebook newsfeed algorithm on behavioral intention to vaccinate one's children, and we propose that these relationships are mediated by the individual's negative attitude toward vaccination and perceived Facebook antivaccination descriptive and injunctive norms.

2.2 Fairness, Accountability, and Awareness (FAT)

AI algorithms increasingly make decisions for people online, from determining what products they might

like to purchase (i.e., recommender systems) to curating the content they see in their social media newsfeeds [32, 113]. How these algorithms make such decisions is often poorly understood by the users whose decisions the algorithms influence. As AI and analytics continue to shape society, it is important to evaluate what users know about how these algorithms influence their lives and whether that understanding, or the lack thereof, shapes people's behaviors. Shin and Park [113] argued that it is important to conceptualize algorithms as sociotechnical systems, that is, to view both the social and technical aspects of a technology as essential characteristics of the technology. They adapted FAT to study how users' perceptions of an algorithm influence algorithmic satisfaction. In the face of accusations of algorithmic bias in the media, FAT[i] has become a popular set of principles for assessing AI[ii] [1, 36]. We follow this approach to examine how users' FAT perceptions of Facebook's newsfeed algorithm contribute to vaccine hesitancy.

In 2014, a paper demonstrated how scientists at Facebook had manipulated people's Facebook newsfeeds to expose some people to more emotionally positive and others to more emotionally negative content [77]. The findings of the study demonstrated that people's emotions could be manipulated by tweaking the newsfeed algorithm. The study showed that on online social networks, "emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness" [77, p. 1]. This study received a great deal of media attention [e.g., 92, 93], arguably because it provided evidence that people can be manipulated by the design of an algorithm. As one reporter said of the experiment, "It was probably legal. But was it ethical?" [93, p. 1]. The capacity of Facebook's newsfeed algorithm to be so powerfully manipulative does pose questions about "social and ethical issues arising from the algorithmic society" [113, p. 278].

It is critical to understand how users view the algorithms that make decisions for them, such as newsfeed algorithms or recommender systems. Shin and Park [113] proposed that people's perceptions of algorithmic FAT could be measured and used to predict their satisfaction with the algorithm. *Algorithmic fairness* means that the logic the algorithm uses to make decisions is fair, just, accurate, unbiased, and nondiscriminatory [113, 133]. The issue of fairness in AI has arisen frequently as businesses and governments have increasingly adopted AI tools. For example, there has been controversy over the fairness

of AI used to determine jail sentences [55] and whom to hire [30] and of AI used for identification [85]. Moreover, it was recently demonstrated that Twitter's algorithm expresses a racial bias in the photos it prioritizes [44].

Algorithmic accountability is the ability to hold an entity responsible for the quality of the decisions made by its algorithms [32, 68, 113]. There has been much debate over accountability in social media, specifically over whom should be held accountable for disinformation [e.g., 111]. It is often argued that algorithmic transparency is a precursor to determining accountability—that it is necessary to understand how an algorithm works before blame can be assigned for its failures [5, 32, 68]. Ananny and Crawford [5, p. 974] suggested that "rather than privileging a type of accountability that needs to look inside systems," we should "instead hold systems accountable by looking across them—seeing them as sociotechnical systems that do not contain complexity but enact complexity by connecting to and intertwining with assemblages of humans and non-humans." *Algorithmic transparency* is thus a user's understanding of how an algorithm works or what they think they know about how the algorithm makes decisions.

2.3 Fairness, Accountability, and Transparency (FAT) and the TPB Factors

In the following subsections, we hypothesize the relationships between users' FAT perceptions of the Facebook newsfeed algorithm, their negative attitudes toward vaccination and perceptions of Facebook antivaccination descriptive and injunctive norms, and their intentions to vaccinate their children.

2.3.1 Fairness and Dis(mis)information Exposure

Perceptions of fairness have been found to affect attitudes in a variety of contexts. For example, one study developed a TPB model to examine users' intention to comply with information security policies and found that the perceived fairness of an information security policy increased users' positive attitudes toward compliance with the policy [13]. In another study, consumers' attitudes toward sellers were found to be better when consumers perceived that the sellers adhered to fair rules for pricing [89]. Research has also shown that *procedural justice*, which is "the perceived fairness of the procedures used to determine outcomes [84]" [42, p. 163], played an important role in employee attitudes toward a company that was conducting drug tests [75]. They found that procedural justice predicted attitudes better than did outcome

fairness. Organizational research has found that fair resource allocation and fair procedures are important predictors of attitudes toward management and organizations [76]. Research has thus determined that perceiving the logic of a policy or request to be fair will result in better attitudes toward what the individual is asked to do (e.g., comply with the policy).

The concept of algorithmic fairness is linked to the concept of procedural justice. The logic a computer algorithm uses to determine its behavior is analogous to the procedures companies or people follow to determine their behaviors. Content that is perceived as the result of a biased (i.e., unfair) algorithm is likely to be taken less seriously and will thus be less likely to influence the user's attitude toward the behavior (e.g., not vaccinating children) the content recommends. Focusing on users who have been exposed to vaccine dis(mis)information on Facebook, we propose that individuals' attitudes toward vaccination will be more susceptible to dis(mis)information when they feel the newsfeed algorithm's logic is fair in how it determines what content is shown to users. Therefore, Facebook users who have been exposed to antivaccination content by an algorithm employing logic that has been deemed fair regarding how it determines the content displayed to the user may hold a negative attitude toward vaccination.

H1a. The perception of users exposed to vaccine-dis(mis)information that the logic used by Facebook's newsfeed algorithm to determine the content displayed is *fair* is <u>positively</u> associated with a *negative attitude toward vaccination*.

Researchers have linked fairness (i.e., justice) to norms in a TPB model of digital piracy, finding that the more unfair piracy was perceived to be, the less likely individuals were to believe that important others thought they should engage in piracy [134]. Notably, research has found that social norms can mediate the relationship between perceived fairness and behavior. Specifically, Lin et al. [83] found that perceptions of procedural justice influenced cooperative norms, which in turn influenced individual helping behaviors. The authors reasoned that the climate of fairness in organizations (i.e., the belief that employees are treated justly) nurtures the cooperative norm and stimulates helping behaviors.

It has been argued that "social media platforms can be effective in facilitating new norms that encourage or discourage risk taking" [63, p. e52]. Research has found that social norms on Facebook (e.g., for sharing content) are learned by watching what others on the social media site do [15, 91]. In one study, students

reported smoking more when exposed to pictures of their friends engaging in risky behavior on social media [63]. Users who feel the Facebook newsfeed algorithm is fair in determining what content it displays, and who have been exposed to antivaccination content from their social connections, may believe that those social connections think people should not vaccinate their children (injunctive norms) and to believe that those connections do not vaccinate their own children (descriptive norms).

> **H1b.** The perception of users exposed to vaccine-dis(mis)information users that the logic used by Facebook's newsfeed algorithm to determine the content displayed is *fair* is positively associated with the perception that their Facebook social connections hold *antivaccination descriptive and injunctive norms*.

2.3.2 Accountability and Dis(mis)information Exposure

Algorithmic accountability has become an increasingly important topic because algorithms have become integral to society. Much has been made of Facebook's newsfeed algorithm's potential to influence the democratic process, but everyday phenomena, such as how much one pays for products on Amazon or for a ride on Uber (i.e., dynamic pricing), are also subject to algorithmic decisions [22, 31, 32, 38, 121]. Algorithms influence people's lives, but research has suggested that people are largely unaware of this influence [54, 106, 108]. For example, in a study of college students, Powers [106] found that most of them were unaware that Google and Facebook tracked people's use or personalized content. The opacity of algorithms combined with limited human input into the decisions they make has led to calls for and discussions about *algorithmic governance* to address and protect society from algorithmic decisions that are inaccurate or unfair [29, 32]. Algorithmic accountability is the user's perception that some entity (e.g., the company that created and uses the algorithm) can be held to account if an algorithm makes irresponsible, inaccurate, or inequitable decisions [29, 32, 113].

Some studies have examined the influence of accountability mechanisms on attitudes. For example, awareness of the consequences of prosocial action has been shown to bolster positive attitudes toward environmentally friendly practices [109]. Perceived consequences of using m-commerce applications have also been shown to increase positive attitudes toward their use [69]. Ascription of responsibility has been demonstrated to positively influence attitude toward energy-saving practices [128]. These studies show that

consequences can affect decision-makers' attitudes toward behaviors. Privacy and security researchers have found that institutional factors can influence people's perceptions of privacy and trust. For example, institutional assurances—mechanisms that can be used to hold a business accountable—have been shown to build trust, which can in turn encourage the use of e-commerce websites [46]. Similarly, the perceived effectiveness of privacy policies has been found to reduce consumers' perceptions of privacy risk [132]. Together, these studies suggest users who believe Facebook can be held accountable if their newsfeed algorithm makes inaccurate, irresponsible, or inequitable decisions about the content it shows have less reason to ignore or doubt the antivaccination content to which they have been exposed.

> **H2a.** The perception of users exposed to vaccine-dis(mis)information that Facebook can be held *accountable* for the content-display decisions made by its newsfeed algorithm is positively associated with a *negative attitude toward vaccination*.

Perceptions of shared responsibility have been linked to individuals perceiving a sharing social norm [82]. In another study, ascription of responsibility for a prosocial behavior (i.e., responsibility for the negative outcomes others experience as a result of not implementing a behavior) has been found to have a positive influence on one's personal norms for environmentally-friendly behavior [128]. Governmental influence has been shown to positively influence subjective norms in an electronic-health-record-adoption context [60]. Organizational formalization, which provides information about what behaviors are valuable to an organization, has been found to positively influence subjective norms [59]. Physicians are accountable for their advice they give, and their recommendations have been demonstrated to positively influence vaccination norms [47]. These findings suggest that accountability mechanisms can be used to encourage the development of norms to ascribe preferred behaviors.

Injunctive social norms represent what is considered the preferred opinion or behavior of a particular group of people [14], and descriptive norms represent what members of the group are actually seen to do [23]. People need to understand what the social norms of a group are, and on Facebook this is accomplished by viewing the sentiments of, or the actions described in the content posted by, the group (e.g., Facebook social connections) regarding a behavior (e.g., positive or negative posts about vaccinations). It makes sense that when Facebook users are exposed to more antivaccination than pro-vaccination content, they will get

the impression that their Facebook social connections have an antivaccination injunctive norm. A user would perceive their Facebook social connections to have an antivaccination descriptive norm if their connections post about their choices not to vaccinate their own children or family members. Research has shown that people will accept the consensus opinion (e.g., product evaluation) when the uniformity of the individual opinions is high [25]. Users who perceive that Facebook can be held accountable for the content-display decisions made by its algorithm have no reason to believe the antivaccination social media content they are exposed to is biased. Thus, users who see vaccine-dis(mis)information on Facebook and believe-that Facebook will be accountable if its algorithm releases inaccurate or irresponsible vaccine information, may believe that their Facebook social connections hold antivaccination descriptive and injunctive social norms.

> **H2b.** The perception of users exposed to vaccine-dis(mis)information that Facebook can be held *accountable* for the content-display decisions made by its newsfeed algorithm is positively associated with the perception that their Facebook social connections hold *antivaccination descriptive and injunctive norms*.

2.3.3 Transparency and Dis(mis)information Exposure

Algorithmic transparency can be defined as the user's understanding of an algorithm's inner workings, its motives, and the characteristics that drive its decisions [33, 127]. Facebook users report that although they know the newsfeed algorithm is attempting to show them relevant content, they are uncertain about how this is accomplished [108], which suggests that Facebook's newsfeed algorithm lacks transparency. However, transparency can have a positive effect on attitudes. For example, awareness of information security policies has been found to positively influence attitudes toward them [13]. Recognizing covert online advertising as such can give consumers a negative attitude toward the ad and brand, but being transparent about the sponsorship of such ads can mitigate this negative effect [39, 40]. Perceived information transparency has also been shown to increase positive attitudes toward information system use [4] and positive attitudes toward socially responsible companies [67]. These results suggest that when people feel they are given enough information to understand the situation, their attitude toward the desired behavior improves.

Perceived transparency strengthens trust in algorithms [98, 127]. Users think they understand how transparent algorithms arrive at decisions (e.g., what content they are shown) and thus might be less likely to question the credibility of dis(mis)information. An algorithm bestowed with professional status is perceived as credible, and users are apt to follow its advice [51]. Research on recommendation systems has found that users feel more confident about recommendations provided by a transparent algorithm [115]. In our context, the newsfeed algorithm is not creating the advice—it is curating advice from others. Users who have some understanding of how the Facebook newsfeed algorithm works can make choices (e.g., engage with content) that affect what content it shows them [6], which means that informed users can make choices that may reduce the antivaccine content they see. However, high algorithmic transparency means users will think they fully understand why they see the content they see, and that may create the illusion that they are in complete control over it. Thus, users who receive vaccine dis(mis)information users who perceive Facebook's newsfeed algorithm as transparent may manifest a negative attitude toward vaccination.

H3a. The perception of users exposed to vaccine-dis(mis)information that Facebook's newsfeed algorithm makes content-display decisions in a *transparent* manner is positively associated with a *negative attitude toward vaccination*.

Learning the social norms in a virtual community requires observation of other people's actions and the consequences of those actions [70]. Gunawan and Huarng [52] examined the influence of social media interaction on subjective norms in the context of purchasing. In their study, social interaction is composed of two components: *social integration*, which is the social network connections' transparency about themselves, and *social influence*, which is the social pressure exerted; the authors suggested that both influence subjective norms. Social influencers are thought to create social norms that drive purchasing on social media [97]. Similarly, sources of antivaccination information can serve as social media influencers and drive the perception that a person's Facebook social connections hold antivaccination descriptive and injunctive norms.

A descriptive social norm describes what others actually do (e.g., cheat), whereas an injunctive social norm describes the perception of what others think a person should do (e.g., not cheat) [80]. Research has found that when people know others are cheating (i.e., the descriptive norm is to cheat), major rule

violations increase [86]. This indicates that seeing whether others are vaccinating their children may be critical to increasing antivaccination behavior. Facebook's newsfeed algorithm determines the content its users see, and if users believe the algorithm is transparent, then they will think they understand how it works and can take steps to adapt the content they see (e.g., engage with content they agree with). Moreover, research has suggested that social media can be used to facilitate the development of social norms [63]. Thus, users who receive vaccine dis(mis)information who believe Facebook's newsfeed algorithm is transparent may think their Facebook social connections hold antivaccination descriptive and injunctive norms.

> **H3b.** The perception of users exposed to vaccine-dis(mis)information that Facebook's newsfeed algorithm makes content-display decisions in a *transparent* manner is positively associated with the perception that their Facebook social connections hold *antivaccination descriptive and injunctive norms*.

2.4 Users' Intentions to Vaccinate Who are Exposed to FAT and Dis(mis)information

Researchers have found that perceived fairness can positively affect intentions to comply with tax regulations [65]. People who feel the tax system is unfair are more likely to rationalize not paying their taxes. It has been noted that when people feel disadvantaged, they may be more likely to intend to participate in activities that promote their values [125, 138]. Perceived unfairness has been found to decrease the intention of employees to participate in professional development activities, with the reasoning being that unfair procedures signal disrespect which motivates uncooperativeness [138]. Similarly, we posit that users who believe the newsfeed algorithm's logic is unfair in its content-display decisions have a reason to ignore or discount the vaccine dis(mis)information they see on Facebook. Dis(mis)information propagated by an algorithm perceived as fair about its content-display decisions is thus likely to be followed.

> **H4a.** The perception of users exposed to vaccine-dis(mis)information that the logic used by Facebook's newsfeed algorithm to determine the content displayed is *fair* is negatively associated with their *intentions to vaccinate* their children.

A study of software agents (i.e., algorithms that support decision-making) found, contrary to expectations, that as perceived accountability increased, individuals were more likely to delegate decisions to the software agent [119]. The authors argued that deferring to the software agent's decision enables

people to shift blame and responsibility to the technology. Accountability has been associated with intention to comply with recommended behaviors [122, 123]. In such cases, mechanisms to hold an individual accountable can reduce unethical behaviors [9]. Similar logic has been applied to behavioral adherence: accountability can encourage people to continue performing healthy behaviors [95]. Organizational commitment has also been demonstrated to increase compliance intentions [58, 105], which suggests that a behavior can be induced by demonstrating that it is valued. These results suggest that people are more comfortable performing recommended behaviors when they feel those behaviors are valued and that some entity will be held accountable for bad advice. Hence, users who receive vaccine dis(mis)information who perceive that Facebook can be held accountable for its newsfeed algorithm making accurate, responsible, and equitable decisions about the content shown may be less likely to vaccinate their children.

> **H4b.** The perception of users exposed to vaccine-dis(mis)information that Facebook can be held *accountable* for the content-display decisions made by its newsfeed algorithm is negatively associated with their *intentions to vaccinate* their children.

Employees who are provided with sufficient information have been found to report stronger intentions to perform change-supportive behaviors [66], which indicates that feeling more informed can increase behavioral intention. Sponsorship transparency can mitigate the negative effects of advertising and can even result in positive consumer intentions and responses [17]. Information transparency has been found to positively affect consumers' purchasing intention [136]. These results suggest that when users feel informed about a behavior, they are more likely to perform that behavior. Feeling confident that they understand how Facebook's newsfeed algorithm works gives users the illusion of choice over the content they see, which gives them less reason to question that content and more reason to follow the advice not to vaccinate their children. Therefore, users who receive vaccine dis(mis)information who think they understand how the Facebook newsfeed algorithm makes content-display decisions may be less to vaccinate their children.

> **H4c.** The perception of users exposed to vaccine-dis(mis)information that Facebook's newsfeed algorithm makes content-display decisions in a *transparent* manner is negatively associated with their *intentions to vaccinate* their children.

2.5 Perceptions of Facebook Norms and Intentions to Vaccinate

The TPB has been employed to examine a variety of behaviors, including vaccination. Pro-vaccination

social norms and positive attitudes toward vaccination have been found to increase the intention to vaccinate [10, 47]. These findings suggest that users exposed to dis(mis)information with negative attitudes toward vaccination and who believe their Facebook social connections hold antivaccination injunctive and descriptive norms may be unlikely to intend to vaccinate their children. Attitude and norms have also been found to reliably predict a variety of health- and compliance-related behavioral intentions, such as the use of electronic health records or compliance with security policy [e.g., 60, 61, 134]. In this study, we are concerned with people who are exposed to vaccine dis(mis)information on Facebook. If vaccine dis(mis)information is effective, then exposed users will have negative attitudes toward vaccination and perceive that their Facebook connections hold antivaccination descriptive and injunctive norms, which will decrease parents' intentions to vaccinate their children. Research has shown that social media participation can result in offline engagement [26], which suggests that attitudes and beliefs formed from social media content can influence intention to perform offline behaviors.

We follow the TPB literature to propose the following relationships. First, a negative attitude toward vaccination held by parents who receive dis(mis)information is likely to decrease their intention to vaccinate their children. Second, "people tend to conform to what other people do" [14, p. 2], which means that users exposed to dis(mis)information who perceive that their Facebook social connections hold antivaccination descriptive and injunctive norms may be unlikely to intend to vaccinate their children.

H5a. *Negative attitudes toward vaccination* by users exposed to vaccine-dis(mis)information will be <u>negatively</u> associated with their *intentions to vaccinate* their children.

H5b. The *antivaccination descriptive and injunctive norms* of the Facebook social connections of users exposed to vaccine-dis(mis)information will be <u>negatively</u> associated with their *intentions to vaccinate* their children.

3. Methods and Analysis

In this section, we describe the development of our scales and the pre-analysis we performed to assess the scale validity and reliability. We then describe our survey administration and the statistical analysis we performed to test our hypotheses.

3.1 Scale Development and Data Collection

We constructed our survey using preexisting scales that we adapted to our context. The scales and their sources are presented in Table A.2 of Appendix A. After the scales were adapted to our context, we presented the items to two survey experts, who were asked to evaluate them for readability and content validity. The use of an expert panel for survey refinement reduces the likelihood of common method bias [104]. The survey was developed in Qualtrics™ to administer it online, and the items in the online survey were presented to the participants in randomized order, which also decreases the likelihood of common method bias. To collect data, we recruited participants on Amazon Mechanical Turk (MTurk), a crowdsourcing platform used in IS and behavioral research [87, 118]. Crowdsourcing platforms provide access to more diverse demographic samples than other traditional options (e.g., students) and satisfactory data quality as long as researchers follow recommend procedures [12, 87]. Data obtained from US participants on crowdsourcing platforms have properties similar to data obtained from students and consumer panels [118]. To ensure the quality of our responses, we followed the practices for using MTurk recommended in the literature, which include paying a small fee for responses, recruiting only US participants, and providing anonymity to the participants to encourage unbiased and honest answers [87, 118].

The desired sample for this study included parents who had been exposed to vaccine dis(mis)information on Facebook. As recommended, we used MTurk's recruitment filter to restrict participation to the US. We also implemented a set of filters to determine whether participants met our sample criteria. To comply with institutional requirements, we asked participants to confirm that they were at least 18 years old. Participants were then asked to confirm that they had at least one social connection on Facebook who had posted about childhood vaccinations; the survey logic filtered out those who answered in the negative. The next set of filters confirmed that the participants had more Facebook social connections who posted concerns or problems with childhood vaccinations than connections who posted positive information. Finally, participants were asked to confirm whether they were parents or were expecting a child.

We ran a pilot study to test the scales by collecting 150 usable responses from MTurk. The pilot study resulted in an acceptable factor structure, indicating that we could proceed with a full data collection. Of those participants who started the survey, two were filtered out for not being at least 18 years of age, and 257 were filtered out because they reported not being exposed to information or comments about childhood vaccinations on Facebook. We checked that the participants had more social connections who posted negative content about childhood vaccinations than connections who posted positive content, and 262 potential participants were filtered out for not satisfying this criterion. Finally, 250 potential participants were filtered out for failing to confirm they had, or were currently expecting, children. To further ensure the quality of our data, we included attention traps, which encourage participants to be cognitively engaged in the survey [99]. Attention trap items filter out those participants who do not correctly answer them; for example, an attention trap may ask the participant to "select 'disagree' for this question." There were 46 participants who did not pass the first attention trap and 14 who did not pass the second. After incomplete and invalid responses were discarded, the final sample size was n = 505. There were 294 male and 211 female participants in the final sample, the other demographics are presented in Table 1.

Table 1. Sample Demographics (n = 505)

Age distribution		**Education**		**Number of Facebook friends**	
18–20 yrs.	5	High school or less	17	0–50	12
21–25 yrs.	50	Some college	46	51–100	52
26–30 yrs.	121	Four-year college degree	284	101–200	100
31–35 yrs.	109	Master's degree	153	201–300	112
36–40 yrs.	63	Doctorate/law/medical degree	5	301–500	92
41–50 yrs.	81			501–700	66
51+ yrs.	76			701–1000	32
				1001+	39
Number of children		**Employment**		**Number of Facebook antivaccination social connections**	
1	252	Not Employed	9	1–5	97
2	192	Employed part-time	47	6–10	116
3	24	Employed full-time	449	11–20	142
4+	37			21–30	80
				31–40	33
				41+	27

3.2 Model Specification and Testing

We performed partial least squares (PLS) path modeling using SmartPLS 4 to test our structural model [53, 110]. In doing so, we followed key guidelines for employing PLS in behavioral research [53, 88, 102, 120].

First, we conducted extensive pre-analysis and data validation, which are outlined in Appendix B. Specifically, we established the factorial validity and reliability of our scales. We ran two statistical tests to rule out common method bias issues, in addition to the steps we took in the research design to mitigate its likelihood. Finally, we confirmed that multicollinearity was not an issue. Our analysis indicates that the model meets the rigorous validation standards necessary for PLS-based analysis [21, 53, 88, 102].

Our model has three reflective exogenous variables that correspond to users' perceptions of the fairness, accountability, and transparency of the Facebook newsfeed algorithm. The reflective mediating variables are negative attitude toward vaccination and Facebook antivaccination descriptive and injunctive norms. The dependent variable is intention to vaccinate. The structural model results, including the R-squared values, are shown in Figure 2. Table 2 presents the path coefficients and *p*-values for each tested relationship.

Figure 2. Model Results for FAT Extension of TPB for Vaccine Hesitancy

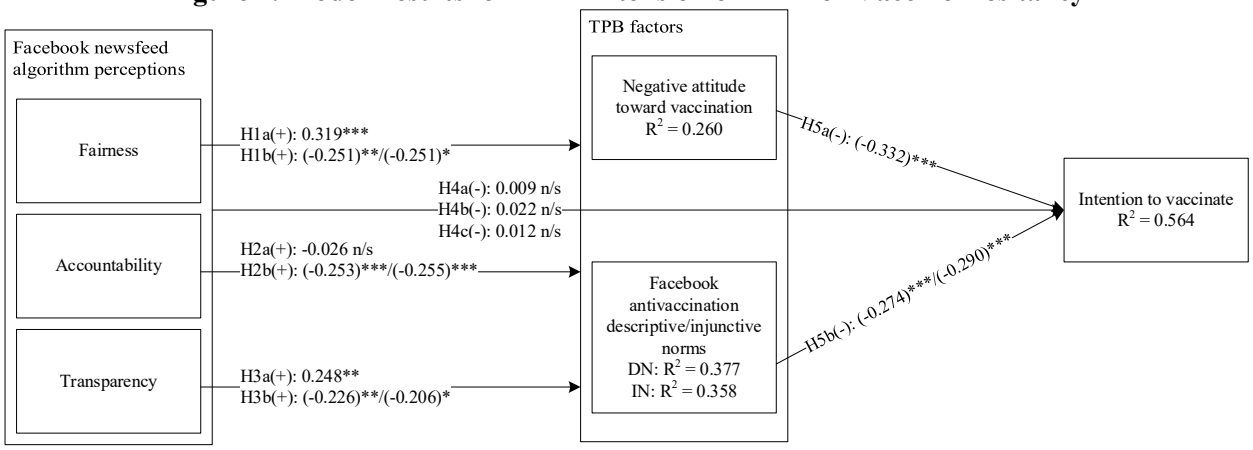


Table 2. Detailed Results of Tested Hypotheses and Control Variables

Tested hypothesis/path	β	*t*-statistic	Support
H1a: Fairness → Negative attitude toward vaccination	0.319	4.257***	Yes
H1b: Fairness → Facebook antivaccination descriptive/injunctive norms	-0.251 -0.251	3.021** 2.351*	No
H2a: Accountability → Negative attitude toward vaccination	(-0.026)	0.442 (n/s)	No
H2b: Accountability → Facebook antivaccination descriptive/injunctive norms	(-0.253) (-0.255)	4.261*** 4.185***	No
H3a: Transparency → Negative attitude toward vaccination	0.248	3.258**	Yes
H3b: Transparency → Facebook antivaccination descriptive/injunctive norms	(-0.226) (-0.206)	2.646** 2.110*	No
H4a: Fairness → Intention to vaccinate	0.009	0.108 (n/s)	No
H4b: Accountability → Intention to vaccinate	0.022	0.401 (n/s)	No
H4c: Transparency → Intention to vaccinate	0.012	0.185 (n/s)	No
H5a: Negative attitude toward vaccination → Intention to vaccinate	(-0.332)	7.073***	Yes

H5b: Facebook antivaccination descriptive/injunctive norms → Intention to vaccinate†	(-0.274) (-0.290)	3.960*** 4.270***	Yes
Controls			
Perceived behavioral control → Intention to vaccinate	0.260	4.191***	Yes
Gender → Intention to vaccinate	0.024	0.426 (n/s)	No
Age → Intention to vaccinate	0.065	2.414*	Yes
Education → Intention to vaccinate	0.040	1.261 (n/s)	No
Employment → Intention to vaccinate	0.034	0.855 (n/s)	No
Number of Facebook antivaccination social connections → Intention to vaccinate	(-0.017)	0.534 (n/s)	No
Number of Facebook friends → Intention to vaccinate	(-0.023)	0.697 (n/s)	No
Number of children → Intention to vaccinate	0.001	0.009 (n/s)	No

*** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, n/s = not significant.

3.3 Moderation Testing

We proposed that negative attitude toward vaccination and Facebook antivaccination descriptive and injunctive norms would mediate the relationship between users' perceptions of the Facebook newsfeed algorithm and their intention to vaccinate. To test for mediation, we used the bootstrapping method [56, 114, 123]. This method has advantages over the traditional Baron and Kenny [8] and Sobel [116] methods, including greater statistical power, not assuming a normal distribution, and allowing for direct measurement of the mediating effects. The results of the bootstrapping test are presented in Table 3.

Table 3. Bootstrapped Confidence Interval Tests for Mediation Model

Proposed relationship	Proposed mediator	Mediation test (*ab*) (indirect effects)			Full/partial mediation test (*c′*)			Type of mediation relationship
		5% lower bound	95% upper bound	Include zero?	2.5% lower bound	97.5% upper bound	Include zero?	
F-Att-Int	Att	-0.159	-0.053	No	-0.167	0.169	Yes	Full
F-DN-Int	DN	0.021	0.119	No	-0.167	0.169	Yes	Full
F-IN-Int	IN	0.005	0.157	No	-0.167	0.169	Yes	Full
A-Att-Int	Att	-0.030	0.046	Yes	-0.082	0.142	Yes	None
A-DN-Int	DN	0.026	0.121	No	-0.082	0.142	Yes	Full
A-IN-Int	IN	0.031	0.123	No	-0.082	0.142	Yes	Full
T-Att-Int	Att	-0.138	-0.028	No	-0.113	0.144	Yes	Full
T-DN-Int	DN	0.012	0.135	No	-0.113	0.144	Yes	Full
T-IN-Int	IN	0.005	0.124	No	-0.113	0.144	Yes	Full

F = fairness; A = accountability; T = transparency; Att = negative attitude toward vaccination; DN/IN = antivaccination descriptive/injunctive norms; Int = intention to vaccinate

4. Discussion

Our results, summarized in Table 4, show that users' FAT perceptions do influence attitudes toward vaccination and the perceptions of vaccination descriptive and injunctive norms held by users' Facebook social connections. However, not all the associations were significant or in the expected direction, which

indicates that the influence of FAT on the vaccine attitudes and beliefs of users exposed to dis(mis)information is more nuanced than anticipated. Overall, our findings demonstrate the importance of educating the public about newsfeed algorithms and how they make content-display decisions, because such education could support efforts to combat vaccine dis(mis)information on social media.

Table 4. Summary of Results

Hypothesis	**Support?**
H1a. Newsfeed algorithm is fair → negative attitude toward vaccination.	Yes
H1b. Newsfeed algorithm is fair → antivaccination descriptive/injunctive norms.	No, negative and significant association
H2a. Facebook is accountable → negative attitude toward vaccination.	No, not significant
H2b. Facebook is accountable → antivaccination descriptive/injunctive norms.	No, negative and significant association
H3a. Newsfeed algorithm is transparent → negative attitude toward vaccination.	Yes
H3b. Newsfeed algorithm is transparent → antivaccination descriptive and injunctive norms.	No, negative and significant association
H4a. Newsfeed algorithm is fair → negative association with intention to vaccinate.	No direct association, but fully mediated by TPB factors
H4b. Facebook is accountable → negative association with intention to vaccinate.	No direct association, but fully mediated by TPB factors
H4c. Newsfeed algorithm is transparent → negative association with intention to vaccinate.	No direct association, but fully mediated by TPB factors
H5a. Negative attitude toward vaccination → negative association with intention to vaccinate.	Yes
H5b. Antivaccination descriptive/injunctive norms → negative association with intention to vaccinate.	Yes

The perception of users who receive vaccine dis(mis)information that the logic used by Facebook's newsfeed algorithm is fair is positively associated with a negative attitude toward vaccination. We also find that negative attitudes toward vaccination is negatively associated with parents' intentions to vaccinate their children. These findings suggest that vaccine dis(mis)information may be effective in turning people against vaccinating their children when that dis(mis)information is displayed by a newsfeed algorithm that users believe employs fair logic. One implication of this result is that vaccine dis(mis)information can be persuasive and that its proliferation on social media by newsfeed algorithms that prioritize *engaging* content may be exacerbating the problem.

Conversely, we find that the perception of users who receive vaccine dis(mis)information that the logic used by Facebook's newsfeed algorithm is fair is negatively associated with the belief that their Facebook social connections hold antivaccination descriptive and injunctive norms, contrary to our expectations. This

finding suggests that even if the newsfeed algorithm is perceived as fair, users are unwilling to assign antivaccination norms to their social connections on Facebook, even when more of their social connections are vaccine opponents as opposed to vaccine proponents. One explanation may be that people holding antivaccination beliefs are more vocal about those beliefs than those holding pro-vaccination beliefs [79, 131]; thus, Facebook users may believe that they are not seeing pro-vaccination content in their newsfeed because their social connections who are proponents of vaccination are unlikely to post about it, rather than attributing it to an unfair algorithm. Consequently, Facebook users who see more antivaccination content than pro-vaccination content, and perceive the newsfeed algorithm to be fair, may still not believe their social connections have antivaccination social norms on a whole because they interpret the silence of their other social connections to be in support of vaccination.

Moreover, Facebook users typically have several social connections that are known to them offline [37] and thus may have evidence from offline conversations about the pro-vaccination beliefs of some of their online social connections. The number of strong antivaccination proponents is small [50], and thus Facebook users may attribute antivaccination norms only to the vocal minority of their social connections. However, belief that one's social connections hold antivaccination descriptive and injunctive norms is negatively associated with the intention to vaccinate one's children. This suggests that when Facebook users find themselves in an echo chamber where antivaccination beliefs is the norm, this could potentially reduce vaccination intentions. A possible positive interpretation of these results is that vocal antivaccination users may not be influential at persuading others to believe their views are the majority views. However, if such antivaccination users are successful at convincing others that they hold majority views, it could decrease vaccinations.

Contrasting these two findings is significant because together they demonstrate that dis(mis)information displayed by a newsfeed algorithm perceived as fair may be effective in changing attitudes but not in changing normative perceptions, and this has key implications for combating vaccine dis(mis)information. Specifically, we find that users who receive vaccine dis(mis)information and perceive that their Facebook social connections hold antivaccination descriptive and injunctive norms have

decreased intention to vaccinate their children. This highlights the danger of social media users involved in disinformation echo chambers. Our results show that when the newsfeed algorithm is perceived as fair, even having more antivaccine than pro-vaccine Facebook social connections may not be enough to convince users that the norms of the group are antivaccine. This means that techniques for countering echo chambers, such as Google's project that offers alternative views to extremist content [124] or modifying the newsfeed algorithm to prioritize pro-vaccine content, may help prevent the perception of antivaccination norms, which could help encourage vaccination. Our results suggest that in order for such techniques to be maximally effective, it is important that the newsfeed algorithm is perceived as fair in the logic it uses. This opens the possibility that techniques for decreasing the perception of antivaccination norms could be used to counteract the damage of vaccine dis(mis)information, especially if future research determines that normative social influence is stronger than persuasive dis(mis)information.

The relationship between the perceived accountability of Facebook for the content-display decisions its algorithm makes and negative attitudes toward vaccination is nonsignificant. However, users exposed to vaccine dis(mis)information who perceive that Facebook can be held accountable for the content-display decisions made by its newsfeed algorithm are less likely to believe that their Facebook social connections hold antivaccination descriptive and injunctive norms. Again, this was the opposite of our expectations. This result similarly indicates that perceptions of accountability can help prevent the perception of antivaccination norms and thus help mitigate the damage resulting from dis(mis)information.

Users who receive vaccine dis(mis)information who think they understand how the Facebook newsfeed algorithm makes content-display decisions are more likely to have negative attitudes toward vaccination but contrary to our expectations, are less likely to think their Facebook social connections hold antivaccination descriptive and injunctive norms. These results suggest that the users who think they understand why they are seeing the content they are seeing are more prone to developing negative attitudes toward vaccination and thus less likely to intend to vaccinate their children. This finding is another confirmation that vaccine dis(mis)information is pernicious and often convincing.

However, users are hesitant to assign antivaccination norms to their Facebook social connections even

when they have been exposed to dis(mis)information and think they know how the algorithm decided to show them the content they saw. One explanation is similar to the one provided above for the negative relationship between fairness and antivaccination norms: users who receive vaccine dis(mis)information who believe they understand how the algorithm works may be confident that, despite the antivaccination content they see, the silent majority of their social connections are vaccine proponents. Again, this may be a favorable result because despite exposure to antivaccination content, it may be difficult to convince users who believe in the fairness, accountability, and transparency of the algorithm that antivaccination beliefs of a vocal minority are the norm, which may have positive ramifications for childhood vaccinations, especially if normative influence is powerful. These results reinforce the need to help users avoid echo chambers through the use of an algorithm that users understand and perceive as fair. Our results also provide justification for removing vaccine mis(dis)information, because even when users do not perceive their entire social media community to be anti-vaxxers, exposure to mis(dis)information from an algorithm perceived as fair and transparent can contribute to negative attitudes toward vaccination and thus decreased intention to vaccinate.

4.1 Contributions to Research and Theory

Our study makes three primary contributions to research and theory. First, the newsfeed algorithm is the key mechanism that determines what many social media users see. On Facebook and several other social media platforms, such as TikTok, the algorithm is proprietary and how it works is largely unknown [57]. Because of the critical role social media is playing in shaping the dis(mis)information landscape, it is important to begin a conversation about the way users perceive newsfeed algorithms. We draw on the FAT principles [113] from the AI literature to examine users' perceptions of the Facebook newsfeed algorithm. In doing so, we contribute to the nascent IS literature on disinformation (i.e., "fake news") [e.g., 24, 71, 72, 96, 135] by examining social media users' perceptions of the algorithm that feeds the dis(mis)information to them rather than their perceptions of the dis(mis)information itself. Moreover, we examine how these perceptions can influence users' intentions to perform an offline behavior—vaccinating their children.

Second, our study contributes to the IS literature on the persuasiveness of social media. Studies have

investigated how social media content persuades people to, for example, make purchases and take political action [e.g., 49, 73]. We take on another pernicious problem—vaccine hesitancy. Researchers have investigated both advice generated by algorithms (e.g., recommender systems) and advice generated by peers (e.g., reviews) [51]. Our context provides a unique mix of the two, because the advice (e.g., antivaccination content) may come from peers or algorithms (e.g., bots), but whether users see that advice is determined by an algorithm. This makes it challenging for users to comprehend the biases that may be implicitly a part of the algorithm's logic, because Facebook users likely see engaging content (i.e., content preferred by others in some way) rather than being exposed to the full conversation. Therefore, their perception of the newsfeed algorithms' FAT is critical to their interpretation of that content within the broader social context.

Finally, our results confirm that negative attitudes toward vaccination and perceptions that one's Facebook social connections hold antivaccination norms can negatively influence intention to vaccinate. However, our study extends this research on vaccine hesitancy to clarify the role the newsfeed algorithm plays in this pernicious problem. Therefore, we examine the interaction between the technology and the people who use it, contributing to the sociotechnical systems literature that takes this approach [e.g., 20, 74, 126]. We find that users' FAT perceptions influence attitudes toward the behavior and normative beliefs in different ways. One interpretation of our results is that users who believe the algorithm is fair and transparent can be swayed by even a little exposure to antivaccination content. Perceptions that Facebook social connections hold antivaccination norms is negatively associated with vaccination intentions, but FAT perceptions decrease perceptions of antivaccination norms, which suggests that it is advisable to keep people out of social media echo chambers in which antivaccination norms may be perceived and use a fair, accountable, and transparent algorithm to serve content that reduces the appearance of antivaccination norms. Our results thus show that FAT perceptions can influence attitudes and normative perceptions differently and that dis(mis)information served by a fair, accountable, and transparent algorithm may be effective without its message being perceived as the social norm. Our results indicate that because of the complexities of the social and technical systems operating jointly to cause the problem, there may be

multiple ways to combat disinformation by (1) better designing the newsfeed algorithms around the social realities, (2) educating users about the purpose and functioning of newsfeed algorithms, and (3) giving users more control over their newsfeeds.

4.2 Implications for Society and Practice

Our findings have two important implications for society and practice. First, it is critical for social media companies to be more transparent about the logic of their newsfeed algorithms or, alternatively, to give users more control over adapting the newsfeed algorithm's display logic (e.g., prioritizing, snoozing, blocking content). It is equally important to educate the public about what is and is not known about how the algorithms work. These two steps are critical, because our results suggest that people are making judgments about algorithmic FAT, and it would be preferable for them to make informed judgments. Specifically, we show that FAT perceptions can influence attitudes and beliefs, and through these, intention to vaccinate. Educating the public about how the algorithms make content-display decisions will thus allow users exposed to dis(mis)information to make more informed judgements about the algorithm and thus potentially temper the effectiveness of vaccine dis(mis)information.

Second, our results suggest that exposure to dis(mis)information by an algorithm deemed fair and transparent, and for which it is thought the company can be held accountable, may not be enough for users to assign antivaccination descriptive and injunctive norms to their Facebook social connections at large. Especially if normative influence is powerful, this leaves open the possibility that keeping users out of dis(mis)information echo chambers may be a way to minimize its effectiveness, because perceptions of antivaccination norms may decrease intention to vaccinate and thus it is potentially useful to encourage tactics that may discourage perceptions of antivaccination norms. Social media platforms may want to consider tactics like those tested by Google, which involve adding counterpoints to vaccine dis(mis)information to users' newsfeeds [124] or prioritizing such counterpoint posts made by their friends, so that dis(mis)information and verified truths are balanced. The reason for this may be less to convince the users of truths versus untruths and more to preserve the appearance that an antivaccine norm does not exist—that their social connections are not uniform in their opinions. Such techniques, combined with

education, may provide ways to exert normative pressure or inform users about why they see the content they see, which could be a less contentious approach than refuting actual vaccine dis(mis)information.

4.3 Limitations and Future Research

Our study is not without limitations, and these limitations point to opportunities to expand on our findings. First, we conducted a cross-sectional study and surveyed Facebook users who have received vaccine dis(mis)information. This method provided initial findings that demonstrate the associations between the FAT perceptions, attitudes and beliefs, and intention to vaccinate. However, a follow-up study could employ a qualitative method that could more deeply explore the connections between the technical—the newsfeed algorithm—and the social—interactions in users' social media communities and vaccine behaviors.

Second, we chose to survey Facebook users who were connected to more antivaccine than pro-vaccine content sharers. Our interest was in gaining initial insights into the role of FAT perceptions in the childhood-vaccine-hesitancy problem. This study should be extended to examine particular groups of interest, such as anti-vaxxers (e.g., diehard antivaccine advocates) or those opposed to specific vaccines (e.g., COVID-19). Our findings indicate that the social dynamics at play in the disinformation ecosystem are important and may have different properties than those that make the dis(mis)information itself effective. Future studies could consider social dynamics along with technical mechanisms to examine ecosystems characterized by more robust dis(mis)information.

Finally, we chose Facebook as our social media platform because of its prominence and the characteristics of its newsfeed algorithm. The newsfeed algorithms of other social media platforms work differently (e.g., Instagram), have different technical characteristics (e.g., SnapChat posts disappear), or are tailored to particular segments of society (e.g., Parler). Our results show that to better combat dis(mis)information, it is useful to develop a better understanding of how perceptions of the technical components of social media platforms can contribute to human decision-making processes. Future research should expand our study to explore different technological features and different social media platforms.

5. Conclusion

This study examines the role of users' FAT perceptions of the Facebook algorithm in the childhood-vaccine-hesitancy problem spurred by dis(mis)information propagated on social media. We find that perceptions of algorithmic fairness and transparency increase the likelihood that Facebook users who receive dis(mis)information will have negative attitudes toward vaccination, which decrease intention to vaccinate. Contrary to our expectations, algorithmic FAT perceptions decrease the likelihood that these users will perceive that their Facebook social connections have antivaccination social norms, which decreases intention to vaccinate. Our results thus suggest that methods for combatting vaccine dis(mis)information may include technical remedies that decrease the prominence of dis(mis)information but also that social remedies, such as educating the public about social media newsfeeds, may be advisable.

6. References

1. Addis, C and Kutar, M. AI management an exploratory survey of the influence of GDPR and FAT principles. Presented at *2019 IEEE SmartWorld*, Leicester, United Kingdom, 2019, pp. 342-347.
2. Agarwal, V. A/H1N1 vaccine intentions in college students: An application of the theory of planned behavior. *Journal of American College Health*, 62, 6 (2014), 416-424.
3. Ajzen, I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 2 (1991), 179-211.
4. Al-Jabri, IM and Roztocki, N. Adoption of ERP systems: Does information transparency matter? *Telematics and Informatics*, 32, 2 (2015), 300-310.
5. Ananny, M and Crawford, K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20, 3 (2018), 973-989.
6. Bakshy, E; Messing, S; and Adamic, LA. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348, 6239 (2015), 1130-1132.
7. Ball, P. Anti-vaccine movement could undermine efforts to end coronavirus pandemic, researchers warn. *Nature*, (2020), Date last accessed: August 1, 2020, https://www.nature.com/articles/d41586-020-01423-4
8. Baron, R and Kenny, D. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 6 (1986), 1173-1182.
9. Beu, D and Buckley, MR. The hypothesized relationship between accountability and ethical behavior. *Journal of Business Ethics*, 34, 1 (2001), 57-73.
10. Britt, RK; Collins, WB; Wilson, KM; Linnemeier, G; and Englebert, AM. The role of eHealth literacy and HPV vaccination among young adults: Implications from a planned behavior approach. *Communication Research Reports*, 32, 3 (2015), 208-215.
11. Broniatowski, DA; Jamison, AM; Qi, S; AlKulaib, L; Chen, T; Benton, A; Quinn, SC; and Dredze, M. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *American Journal of Public Health*, 108, 10 (2018), 1378-1384.
12. Buhrmester, MD; Talaifar, S; and Gosling, SD. An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological Science*, 13, 2 (2018), 149-154.
13. Bulgurcu, B; Cavusoglu, H; and Benbasat, I. Roles of information security awareness and perceived fairness in information security policy compliance. Presented at *Proceedings of the*

Fifteenth Americas Conference on Information Systems, San Francisco, CA, 2009, pp. 1-9.
14. Burchell, K; Rettie, R; and Patel, K. Marketing social norms: Social marketing and the 'social norm approach'. *Journal of Consumer Behaviour*, 12, 1 (2013), 1-9.
15. Burke, M; Marlow, C; and Lento, T. Feed me: Motivating newcomer contribution in social network sites. Presented at *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA, 2009, pp. 945-954.
16. Burki, T. Vaccine misinformation and social media. *The Lancet Digital Health*, 1, 6 (2019), e258-e259.
17. Campbell, C and Evans, NJ. The role of a companion banner and sponsorship transparency in recognizing and evaluating article-style native advertising. *Journal of Interactive Marketing*, 43, (2018), 17-32.
18. Catalano, HP; Knowlden, AP; Birch, DA; Leeper, JD; Paschal, AM; and Usdan, SL. Using the theory of planned behavior to predict HPV vaccination intentions of college men. *Journal of American College Health*, 65, 3 (2017), 197-207.
19. CDC. Measles cases in 2019. *Cases and Outbreaks*, (2019), Date last accessed: July 19, 2019, https://www.cdc.gov/measles/cases-outbreaks.html
20. Chai, S and Kim, M. A socio-technical approach to knowledge contribution behavior: An empirical investigation of social networking sites users. *International Journal of Information Management*, 32, 2 (2012), 118-126.
21. Chin, WW. The partial least squares approach for structural equation modeling. *Modern methods for business research.*, *Methodology for business and management.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 1998, pp. 295-336.
22. Cho, J; Ahmed, S; Hilbert, M; Liu, B; and Luu, J. Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization. *Journal of Broadcasting & Electronic Media*, 64, 2 (2020), 150-172.
23. Cialdini, RB; Reno, RR; and Kallgren, CA. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, 6 (1990), 1015-1026.
24. Clarke, J; Chen, H; Du, D; and Hu, YJ. Fake news, investor attention, and market reaction. *Information Systems Research*, Forthcoming, (2020), 1-18.
25. Cohen, JB and Golden, E. Informational social influence and product evaluation. *Journal of Applied Psychology*, 56, 1 (1972), 54-59.
26. Conroy, M; Feezell, JT; and Guerrero, M. Facebook and political engagement: A study of online political group membership and offline political engagement. *Computers in Human Behavior*, 28, 5 (2012), 1535-1546.
27. Constine, J. How Facebook News Feed Works. (2016), Date last accessed: August 1, 2020, https://techcrunch.com/2016/09/06/ultimate-guide-to-the-news-feed/
28. Cornally, N; Deasy, EA; McCarthey, G; McAuley, C; Moran, J; and Weathers, E. Student nurses' intention to get the influenza vaccine. *British Journal of Nursing*, 22, 21 (2013), 1207-1211.
29. Danaher, J; Hogan, MJ; Noone, C; Kennedy, R; Behan, A; De Paor, A; Felzmann, H; Haklay, M; Khoo, S-M; and Morison, J. Algorithmic governance: Developing a research agenda through the power of collective intelligence. *Big Data & Society*, 4, 2 (2017), 1-21.
30. Dastin, J. Amazon scraps secret AI recruiting tool that showed bias against women. (2018), Date last accessed: August 3, 2020, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G
31. Diakopoulos, N. How Uber surge pricing really works. (2015), Date last accessed: September 20, 2020, https://www.washingtonpost.com/news/wonk/wp/2015/04/17/how-uber-surge-pricing-really-works/
32. Diakopoulos, N. Accountability in algorithmic decision making. *Communications of the ACM*, 59, 2 (2016), 56-62.

33. Diakopoulos, N and Koliska, M. Algorithmic transparency in the news media. *Digital Journalism*, 5, 7 (2017), 809-828.
34. Downs, DS and Hausenblas, HA. Elicitation studies and the theory of planned behavior: A systematic review of exercise beliefs. *Psychology of Sport and Exercise*, 6, 1 (2005), 1-31.
35. Dubé, E; Gagnon, D; Ouakki, M; Bettinger, JA; Witteman, HO; MacDonald, S; Fisher, W; Saini, V; and Greyson, D. Measuring vaccine acceptance among Canadian parents: A survey of the Canadian Immunization Research Network. *Vaccine*, 36, 4 (2018), 545-552.
36. Eisenstat, Y. The Real Reason Tech Struggles With Algorithmic Bias. August 2, 2020 (2019), Date last accessed: https://www.wired.com/story/the-real-reason-tech-struggles-with-algorithmic-bias/
37. Ellison, NB; Steinfield, C; and Lampe, C. The benefits of Facebook "friends": Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, 12, 4 (2007), 1143-1168.
38. Epstein, R and Robertson, RE. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences*, 112, 33 (2015), E4512-E4521.
39. Evans, NJ; Phua, J; Lim, J; and Jun, H. Disclosing Instagram influencer advertising: The effects of disclosure language on advertising recognition, attitudes, and behavioral intent. *Journal of Interactive Advertising*, 17, 2 (2017), 138-149.
40. Evans, NJ; Wojdynski, BW; and Grubbs Hoy, M. How sponsorship transparency mitigates negative effects of advertising recognition. *International Journal of Advertising*, 38, 3 (2019), 364-382.
41. Facebook. How News Feed Works. (2020), Date last accessed: August 1, 2020, https://www.facebook.com/help/1155510281178725
42. Fassina, NE; Jones, DA; and Uggerslev, KL. Relationship clean-up time: Using meta-analysis and path analysis to clarify relationships among job satisfaction, perceived fairness, and citizenship behaviors. *Journal of Management*, 34, 2 (2008), 161-188.
43. Fine, PE. Herd immunity: History, theory, practice. *Epidemiologic Reviews*, 15, 2 (1993), 265-302.
44. Fingas, J. Twiitter has 'more analysis to do' after algorithm shows possible racial bias. (2020), Date last accessed: September 23, 2020, https://www.engadget.com/twitter-responds-to-algorithm-racial-bias-claims-203337604.html
45. Fishbein, M and Ajzen, I. *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press, 2010.
46. Gefen, D; Karahanna, E; and Straub, DW. Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, 27, 1 (2003), 51-90.
47. Gerend, MA and Shepherd, JE. Predicting human papillomavirus vaccine uptake in young adult women: comparing the health belief model and theory of planned behavior. *Annals of Behavioral Medicine*, 44, 2 (2012), 171-180.
48. Godin, G; Vézina-Im, L-A; and Naccache, H. Determinants of influenza vaccination among healthcare workers. *Infection Control & Hospital Epidemiology*, 31, 7 (2010), 689-693.
49. Goh, K-Y; Heng, C-S; and Lin, Z. Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content. *Information Systems Research*, 24, 1 (2013), 88-107.
50. Gravelle, TB; Phillips, JB; Reifler, J; and Scotto, TJ. Estimating the size of "anti-vax" and vaccine hesitant populations in the US, UK, and Canada: comparative latent class modeling of vaccine attitudes. *Human Vaccines & Immunotherapeutics*, 18, 1 (2022), 2008214.
51. Gunaratne, J; Zalmanson, L; and Nov, O. The persuasive power of algorithmic and crowdsourced advice. *Journal of Management Information Systems*, 35, 4 (2018), 1092-1120.
52. Gunawan, DD and Huarng, K-H. Viral effects of social network and media on consumers' purchase intention. *Journal of Business Research*, 68, 11 (2015), 2237-2241.
53. Hair, JF; Hult, GTM; Ringle, C; and Sarstedt, M. *A primer on partial least squares structural equation modeling (PLS-SEM), Second Edition*. Thousand Oaks, CA: Sage Publications, 2017.

54. Hamilton, K; Karahalios, K; Sandvig, C; and Eslami, M. A path to understanding the effects of algorithm awareness. Presented at *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, Toronto, Canada, 2014, pp. 631-642.
55. Hao, K. AI is sending people to jail—and getting it wrong. (2019), Date last accessed: August 3, 2020, https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/
56. Hayes, AF. Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Communication Monographs*, 76, 4 (2009), 408-420.
57. Heilweil, R. TikTok offered details about how its most popular feed works. Experts seem unimpressed. (2020), Date last accessed: August 1, 2020, https://www.vox.com/recode/2020/6/23/21296056/tiktok-foryou-algorithm-explained-facebook-news-feed
58. Herath, T and Rao, HR. Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems*, 18, 2 (2009), 106-125.
59. Hong, Y and Furnell, S. Organizational formalization and employee information security behavioral intentions based on an extended TPB model. Presented at *2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, Oxford, UK, 2019, pp. 1-4.
60. Hsieh, P-J. Physicians' acceptance of electronic medical records exchange: An extension of the decomposed TPB model with institutional trust and perceived risk. *International Journal of Medical Informatics*, 84, 1 (2015), 1-14.
61. Hu, Q; Dinev, T; Hart, P; and Cooke, D. Managing employee compliance with information security policies: The critical role of top management and organizational culture. *Decision Sciences*, 43, 4 (2012), 615-660.
62. Hu, Y; Chen, Y; Wang, Y; and Liang, H. Measuring childhood vaccination acceptance of mother in Zhejiang province, East China. *Human Vaccines & Immunotherapeutics*, 15, 2 (2019), 287-294.
63. Huang, GC; Soto, D; Fujimoto, K; and Valente, TW. The interplay of friendship networks and social networking sites: Longitudinal analysis of selection and influence effects on adolescent smoking and alcohol use. *American Journal of Public Health*, 104, 8 (2014), e51-e59.
64. Igoe, KJ. Establishing the Truth: Vaccines, Social Media, and the Spread of Misinformation. (2019), Date last accessed: August 1, 2020, https://www.hsph.harvard.edu/ecpe/vaccines-social-media-spread-misinformation/
65. Jimenez, P and Iyer, GS. Tax compliance in a social setting: The influence of social norms, trust in government, and perceived fairness on taxpayer compliance. *Advances in Accounting*, 34, (2016), 17-26.
66. Jimmieson, NL; Peach, M; and White, KM. Utilizing the theory of planned behavior to inform change management: An investigation of employee intentions to support organizational change. *The Journal of Applied Behavioral Science*, 44, 2 (2008), 237-262.
67. Kang, J and Hustvedt, G. Building trust between consumers and corporations: The role of consumer perceptions of transparency and social responsibility. *Journal of Business Ethics*, 125, 2 (2014), 253-265.
68. Kemper, J and Kolkman, D. Transparent to whom? No algorithmic accountability without a critical audience. *Information, Communication & Society*, 22, 14 (2019), 2081-2096.
69. Khalifa, M and Shen, KN. Drivers for transactional B2C m-commerce adoption: Extended theory of planned behavior. *Journal of Computer Information Systems*, 48, 3 (2008), 111-117.
70. Kiesler, S; Kraut, R; Resnick, P; and Kittur, A. Regulating behavior in online communities. In R.E. Kraut, and P. Resnick (eds.), *Building successful online communities: Evidence-based social design*. Cambridge, MA: MIT Press, 2012, pp. 125-178.
71. Kim, A and Dennis, AR. Says who? The effects of presentation format and source rating on fake news in social media. *MIS Quarterly*, 43, 3 (2019), 1025-1039.
72. Kim, A; Moravec, PL; and Dennis, AR. Combating fake news on social media with source ratings: the effects of user and expert reputation ratings. *Journal of Management Information Systems*, 36, 3

(2019), 931-968.
73. Kim, B; Barnidge, M; and Kim, Y. The communicative processes of attempted political persuasion in social media environments. *Information Technology & People*, 33, 2 (2019), 813-828.
74. Kim, H; Shin, D-H; and Lee, D. A socio-technical analysis of software policy in Korea: Towards a central role for building ICT ecosystems. *Telecommunications Policy*, 39, 11 (2015), 944-956.
75. Konovsky, MA and Cropanzano, R. Perceived fairness of employee drug testing as a predictor of employee attitudes and job performance. *Journal of Applied Psychology*, 76, 5 (1991), 698-707.
76. Korsgaard, MA; Roberson, L; and Rymph, RD. What motivates fairness? The role of subordinate assertive behavior on manager's interactional fairness. *Journal of Applied Psychology*, 83, 5 (1998), 731.
77. Kramer, AD; Guillory, JE; and Hancock, JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111, 24 (2014), 8788-8790.
78. Krawczyk, AL; Perez, S; Lau, E; Holcroft, CA; Amsel, R; Knäuper, B; and Rosberger, Z. Human papillomavirus vaccination intentions and uptake in college women. *Health Psychology*, 31, 5 (2012), 685.
79. Kuhn, R; Henwood, B; Lawton, A; Kleva, M; Murali, K; King, C; and Gelberg, L. COVID-19 vaccine access and attitudes among people experiencing homelessness from pilot mobile phone survey in Los Angeles, CA. *PloS One*, 16, 7 (2021), e0255246.
80. Lapinski, MK and Rimal, RN. An explication of social norms. *Communication Theory*, 15, 2 (2005), 127-147.
81. Lewandowsky, S; Ecker, UK; Seifert, CM; Schwarz, N; and Cook, J. Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 3 (2012), 106-131.
82. Liao, W; Yuan, YC; and McComas, KA. Communal risk information sharing: Motivations behind voluntary information sharing for reducing interdependent risks in a community. *Communication Research*, 45, 6 (2018), 909-933.
83. Lin, S-P; Tang, T-W; Li, C-H; Wu, C-M; and Lin, H-H. Mediating effect of cooperative norm in predicting organizational citizenship behaviors from procedural justice climate. *Psychological Reports*, 101, 1 (2007), 67-78.
84. Lind, EA and Tyler, TR. *The social psychology of procedural justice*. New York: Plenum Press, 1988.
85. Lohr, S. Facial Recognition Is Accurate, if You're a White Guy. August 3, 2020 (2018), Date last accessed: August 3, 2020, https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html
86. Lois, G and Wessa, M. Honest mistake or perhaps not: The role of descriptive and injunctive norms on the magnitude of dishonesty. *Journal of Behavioral Decision Making*, Forthcoming, DOI: 10:1002/bdm.2196 (2020), 1-15.
87. Lowry, PB; D'Arcy, J; Hammer, B; and Moody, GD. "Cargo Cult" science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels. *The Journal of Strategic Information Systems*, 25, 3 (2016), 232-240.
88. Lowry, PB and Gaskin, J. Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication*, 57, 2 (2014), 123-146.
89. Maxwell, S. Rule-based price fairness and its effect on willingness to purchase. *Journal of Economic Psychology*, 23, 2 (2002), 191-212.
90. McEachan, RRC; Conner, M; Taylor, NJ; and Lawton, RJ. Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. *Health Psychology Review*, 5, 2 (2011), 97-144.
91. McLaughlin, C and Vitak, J. Norm evolution and violation on Facebook. *New Media & Society*, 14,

2 (2012), 299-315.
92. McNeal, GS. Facebook Manipulated User News Feeds To Create Emotional Responses. (2014), Date last accessed: August 3, 2020, https://www.forbes.com/sites/gregorymcneal/2014/06/28/facebook-manipulated-user-news-feeds-to-create-emotional-contagion/#2dd3999639dc
93. Meyer, R. Everything We Know About Facebook's Secret Mood Manipulation Experiment. (2014), Date last accessed: August 3, 2020, https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/
94. Min, SJ. From algorithmic disengagement to algorithmic activism: Charting social media users' responses to news filtering algorithms. *Telematics and Informatics*, 43, 1 (2019), 101251.
95. Mohr, D; Cuijpers, P; and Lehman, K. Supportive accountability: A model for providing human support to enhance adherence to eHealth interventions. *Journal of Medical Internet Research*, 13, 1 (2011), e30.
96. Moravec, PL; Kim, A; and Dennis, AR. Appealing to Sense and Sensibility: System 1 and System 2 Interventions for Fake News on Social Media. *Information Systems Research*, 31, 3 (2020), 987-1006.
97. Nash, J. Exploring how social media platforms influence fashion consumer decisions in the UK retail sector. *Journal of Fashion Marketing and Management:*, 23, 1 (2019), 82-103.
98. Nilashi, M; Jannach, D; Ibrahim, Ob; Esfahani, MD; and Ahmadi, H. Recommendation quality, transparency, and website quality for trust-building in recommendation agents. *Electronic Commerce Research and Applications*, 19, 1 (2016), 70-84.
99. Oppenheimer, DM; Meyvis, T; and Davidenko, N. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45, 4 (2009), 867-872.
100. Park, HS; Klein, KA; Smith, S; and Martell, D. Separating subjective norms, university descriptive and injunctive norms, and US descriptive and injunctive norms for drinking behavior intentions. *Health Communication*, 24, 8 (2009), 746-751.
101. Pavlou, PA and Fygenson, M. Understanding and predicting electronic commerce adoption: An extension of the theory of planned behavior. *MIS Quarterly*, 30, 1 (2006), 115-143.
102. Peng, DX and Lai, F. Using partial least squares in operations management research: A practical guideline and summary of past research. *Journal of Operations Management*, 30, 6 (2012), 467-480.
103. Pluviano, S; Watt, C; and Della Sala, S. Misinformation lingers in memory: Failure of three pro-vaccination strategies. *PLoS One*, 12, 7 (2017), 1-15.
104. Podsakoff, PM; MacKenzie, SB; Lee, J-Y; and Podsakoff, NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 5 (2003), 879-903.
105. Posey, C; Roberts, TL; and Lowry, PB. The impact of organizational commitment on insiders' motivation to protect organizational information assets. *Journal of Management Information Systems*, 32, 4 (2015), 179-214.
106. Powers, E. My news feed is filtered? Awareness of news personalization among college students. *Digital Journalism*, 5, 10 (2017), 1315-1335.
107. Puri, N; Coomes, EA; Haghbayan, H; and Gunaratne, K. Social media and vaccine hesitancy: New updates for the era of COVID-19 and globalized infectious diseases. *Human Vaccines & Immunotherapeutics*, Forthcoming, (2020), 1-8.
108. Rader, E and Gray, R. Understanding user beliefs about algorithmic curation in the Facebook news feed. Presented at *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, Seoul, South Korea, 2015, pp. 173-182.
109. Rezaei, R; Safa, L; Damalas, CA; and Ganjkhanloo, MM. Drivers of farmers' intention to use integrated pest management: Integrating theory of planned behavior and norm activation model.

Journal of Environmental Management, 236, 1 (2019), 328-339.
110. Sarstedt, M and Cheah, J-H. Partial least squares structural equation modeling using SmartPLS: A software review. *Journal of Marketing Analytics*, 7, 3 (2019), 196-202.
111. Saurwein, F and Spencer-Smith, C. Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe. *Digital Journalism*, 1 (2020), 1-22.
112. Schmidt, AL; Zollo, F; Scala, A; Betsch, C; and Quattrociocchi, W. Polarization of the vaccination debate on Facebook. *Vaccine*, 36, 25 (2018), 3606-3612.
113. Shin, D and Park, YJ. Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, 98, 1 (2019), 277-284.
114. Shrout, PE and Bolger, N. Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 4 (2002), 422-445.
115. Sinha, R and Swearingen, K. The role of transparency in recommender systems. Presented at *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, Minneapolis, MN, 2002, pp. 830-831.
116. Sobel, ME. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 1 (1982), 290-312.
117. Stecula, DA; Kuru, O; and Jamieson, KH. How trust in experts and media use affect acceptance of common anti-vaccination claims. *Harvard Kennedy School Misinformation Review*, 1, 1 (2020), https://doi.org/10.37016/mr-2020-007.
118. Steelman, ZR; Hammer, BI; and Limayem, M. Data collection in the digital age: Innovative alternatives to student samples. *MIS Quarterly*, 38, 2 (2014), 355-378.
119. Stout, N; Dennis, AR; and Wells, TM. The buck stops there: The impact of perceived accountability and control on the intention to delegate to software agents. *AIS Transactions on Human-Computer Interaction*, 6, 1 (2014), 1-15.
120. Straub, D; Boudreau, M-C; and Gefen, D. Validation guidelines for IS positivist research. *Communications of the Association for Information Systems*, 13, 1 (2004), 380-426.
121. Thorson, K; Cotter, K; Medeiros, M; and Pak, C. Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society*, Forthcoming, DOI: 10.1080/1369118X.2019.1642934 (2020), 1-18.
122. Vance, A; Lowry, PB; and Eggett, D. Using accountability to reduce access policy violations in information systems. *Journal of Management Information Systems*, 29, 4 (2013), 263-290.
123. Vance, A; Lowry, PB; and Eggett, DL. Increasing accountability through the user interface design artifacts: A new approach to addressing the problem of access-policy violations. *MIS Quarterly*, 39, 2 (2015), 345-366.
124. Wakefield, J. TED 2018: Alphabet firm's tools to combat extremism. (2018), Date last accessed: January 14, 2021, https://www.bbc.com/news/technology-43760213
125. Walker, I; Wong, NK; and Kretzschmar, K. Relative deprivation and attribution: From grievance to action. In I. Walker, and H.J. Smith (eds.), *Relative deprivation: Specification, development, and integration*. New York, NY: Cambridge University Press, 2002, pp. 351-374.
126. Wan, J; Lu, Y; Wang, B; and Zhao, L. How attachment influences users' willingness to donate to content creators in social media: A socio-technical systems perspective. *Information & Management*, 54, 7 (2017), 837-850.
127. Wang, W and Benbasat, I. Empirical Assessment of Alternative Designs for Enhancing Different Types of Trusting Beliefs in Online Recommendation Agents. *Journal of Management Information Systems*, 33, 3 (2016), 744-775.
128. Wang, W and Wang, M. Effects of Sponsorship Disclosure on Perceived Integrity of Biased Recommendation Agents: Psychological Contract Violation and Knowledge-Based Trust Perspectives. *Information Systems Research*, 30, 2 (2019), 507-522.
129. Wheeler, M and Buttenheim, AM. Parental vaccine concerns, information source, and choice of alternative immunization schedules. *Human Vaccines & Immunotherapeutics*, 9, 8 (2013), 1782-1789.

130. WHO. Ten threats to global health in 2019. (2019), Date last accessed: July 21, 2019, https://www.who.int/emergencies/ten-threats-to-global-health-in-2019
131. Wiyeh, AB; Cooper, S; Jaca, A; Mavundza, E; Ndwandwe, D; and Wiysonge, CS. Social media and HPV vaccination: Unsolicited public comments on a Facebook post by the Western Cape Department of Health provide insights into determinants of vaccine hesitancy in South Africa. *Vaccine*, 37, 43 (2019), 6317-6323.
132. Xu, H; Dinev, T; Smith, J; and Hart, P. Information privacy concerns: Linking individual perceptions with institutional privacy assurances. *Journal of the Association for Information Systems*, 12, 12 (2011), 798-824.
133. Yang, K and Stoyanovich, J. Measuring fairness in ranked outputs. Presented at *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, Chicago, IL, 2017, pp. 1-6.
134. Yoon, C. Theory of planned behavior and ethics theory in digital piracy: An integrated model. *Journal of Business Ethics*, 100, 3 (2011), 405-417.
135. Zahedi, FM; Abbasi, A; and Chen, Y. Fake-website detection tools: Identifying elements that promote individuals' use and enhance their performance. *Journal of the Association for Information Systems*, 16, 6 (2015), 448-484.
136. Zhou, L; Wang, W; Xu, JD; Liu, T; and Gu, J. Perceived information transparency in B2C e-commerce: An empirical investigation. *Information & Management*, 55, 7 (2018), 912-927.
137. Zipprich, J; Winter, K; Hacker, J; Xia, D; Watt, J; and Harriman, K. Measles outbreak—California, December 2014–February 2015. *MMWR. Morbidity and Mortality Weekly Report*, 64, 6 (2015), 153-154.
138. Zoogah, DB. Why should I be left behind? Employees' perceived relative deprivation and participation in development activities. *Journal of Applied Psychology*, 95, 1 (2010), 159-173.

---

[i] For example, the description of the Association of Computing Machinery's conference on Fairness, Accountability, and Transparency (https://facctconference.org/) states:

> *Algorithmic systems are being adopted in a growing number of contexts, fueled by big data. These systems filter, sort, score, recommend, personalize, and otherwise shape human experience, increasingly making or informing decisions with major impact on access to, e.g., credit, insurance, healthcare, parole, social security, and immigration. Although these systems may bring myriad benefits, they also contain inherent risks, such as codifying and entrenching biases; reducing accountability, and hindering due process; they also increase the information asymmetry between individuals whose data feed into these systems and big players capable of inferring potentially relevant information.*

[ii] https://www.fatml.org/

How Facebook's Newsfeed Algorithm Shapes Childhood Vaccine Hesitancy: An Algorithmic Fairness, Accountability, Transparency (FAT) Perspective

Note to editors and reviewers: These are supplementary online appendices for review purposes, not for print publication with the main article.

Online Appendix A. Construct Definitions and Measurement Items

Table A.1. Construct Definitions

Construct (Source)	Definition
Perceived fairness of Facebook algorithm (PFFA); adapted from Kamdar et al. [13]	Users perceptions that the logic the Facebook newsfeed algorithm uses to make content display decisions is fair and just [21, 25].
Accountability of Facebook algorithm (AFA); adapted from [24, p. 347]	Refers to users' perceptions of the ability to hold some entity responsible for the quality of the decisions made by the algorithm [7, 14, 21].
Transparency of Facebook algorithm (TFAB); adapted from Zhou et al. [26, p. 914]	Refers to users' understandings of how the algorithm works, or how much they think they know about how the algorithm makes decisions [21].
Parental attitude about vaccination (negative) (VHR); adapted from [1, p. 188]	Attitude is a personal "latent disposition or tendency to respond with some degree of" unfavorableness toward childhood vaccination [8, p. 76].
Facebook anti-vaccination descriptive norms (DOSN); adapted from Göckeritz et al. [10]	Refer to what a person believes that the referent group to actually be doing with respect to the behavior (e.g., their Facebook social connections are actually declining to vaccinate their children) [6].
OSN Injunctive Anti-Vaccine Norms (IOSN); adapted from Göckeritz et al. [10, p. 515]	Refer to what a person believes that the referent group (e.g., Facebook social connections) thinks they should do with respect to a particular behavior (e.g., not vaccinate their children) [6].
Perceived behavioral control over vaccination (PBC); adapted from Fishbein and Ajzen [8]	Captures "the degree to which an individual actually has control over performing the behavior" [8, p. 64].
Parental vaccination intention (INT); adapted from TPB intention Bock et al. [4, p. 90]	The individual's readiness to perform a determined behavior (e.g., vaccinate his or her children) [8].

Table A.2. Measurement Item Details and Sources

Construct (Source)	Construct indicator	Item	Mean	Std. Dev.
Perceived fairness of Facebook algorithm (PFFA); adapted from Kamdar et al. [13]	**Prompt:** Please read carefully and indicate how much you agree with each of the following statements about Facebook's newsfeed algorithm. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	PFFA1	Facebook's newsfeed algorithm uses fair logic to determine what content I see.	5.238	1.397
	PFFA2	Facebook's newsfeed algorithm's procedures and guidelines for determining what content I see are very fair.	5.210	1.443
	PFFA3	The rules that Facebook's newsfeed algorithm uses to make content decisions are fair.	5.214	1.387
	PFFA4	I can count on Facebook's newsfeed algorithm to use fair methods of deciding what content I see.	5.220	1.419
Perceived accountability for Facebook algorithm (AFA); adapted from Vance et al. [24]	**Prompt:** Please read carefully and indicate how much you agree with each of the following statements about Facebook's newsfeed algorithm. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	AFA1	Facebook Inc. can be held accountable for the content its newsfeed algorithm displays to users.	5.455	1.197
	AFA2	Facebook Inc. can be held accountable for all the content display decisions its newsfeed algorithm makes.	5.436	1.219
	AFA3	I believe that Facebook Inc. can be held accountable for the decisions its newsfeed algorithm makes regarding what content is displayed to users.	5.481	1.215
Perceived transparency of Facebook algorithm (TFAB); adapted from Zhou et al. [26]	**Prompt:** Please read carefully and indicate how much you agree with each of the following statements about Facebook's newsfeed algorithm. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	TFAB1	I know how Facebook's newsfeed algorithm determines which content I see in my newsfeed.	5.366	1.334
	TFAB2	I understand the way Facebook's newsfeed algorithm decides which content is displayed in my newsfeed.	5.257	1.334
	TFAB3	I have a clear idea of how the Facebook newsfeed algorithm works.	5.301	1.357
	TFAB4	I am familiar with the process Facebook's newsfeed algorithm uses to determine the content I see in my newsfeed.	5.251	1.301
Negative attitude toward vaccination (VHR); adapted from Askelson et al. [2]	**Prompt:** Please read carefully and indicate how much you agree with each of the following statements about childhood vaccinations. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	VHR1	I believe that childhood vaccinations are unnecessary.	4.624	1.966
	VHR2	I believe that vaccinating children is not a good idea.	4.511	1.969
	VHR3	I believe that vaccines are not beneficial for children.	4.578	1.926
Facebook anti-	**Prompt:** Please read the following carefully and indicate how much you agree with each of the following statements,			

Construct (Source)	Construct indicator	Item	Mean	Std. Dev.
vaccination descriptive norms (DOSN); adapted from Göckeritz et al. [10]	only in respect to your opinions about your Facebook social network connections. Facebook social network connections refer to any entity whose posts you see in your Facebook newsfeed; for example, Facebook friends, people in Facebook groups to which you belong, or celebrities or organizations you have followed. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	DOSN1	My Facebook social connections vaccinate their children. (R)	5.404	1.226
	DOSN2	The people I am connected to on Facebook vaccinate their children. (R)	5.444	1.193
	DOSN3	The social connections I have on Facebook get their children vaccinated. (R)	5.376	1.219
	DOSN4	My Facebook social connections have their children vaccinated. (R)	5.420	1.239
Facebook anti-vaccination injunctive norms (IOSN) adapted from Göckeritz et al. [10]	**Prompt:** Please read the following carefully and indicate how much you agree with each of the following statements, only in respect to your opinions about your Facebook social network connections. Facebook social network connections refer to any entity whose posts you see in your Facebook newsfeed; for example, Facebook friends, people in Facebook groups to which you belong, or celebrities or organizations you have followed. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	IOSN1	My Facebook social connections approve of people who vaccinate their children. (R)	5.424	1.210
	IOSN2	The social connections I have on Facebook approve of people who have their children vaccinated. (R)	5.406	1.241
	IOSN3	The social connections I have on Facebook approve of people who vaccinate their children. (R)	5.408	1.217
Perceived behavioral control over vaccination (PBC); adapted from Fishbein and Ajzen [8]	**Prompt:** Please read carefully and indicate how much you agree with each of the following statements about childhood vaccinations. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	PBC1	I have control over whether my children are vaccinated.	5.644	1.117
	PBC2	I feel in complete control over whether I vaccinate my children.	5.590	1.157
	PBC3	Whether or not I vaccinate my children is completely up to me.	5.523	1.206
	PBC4	It is mostly up to me whether I vaccinate my children.	5.689	1.112
	PBC5	There are not factors outside of my control that determine whether I vaccinate my children.		
Intention to vaccinate (INT); adapted from TPB intention Bock et al. [4]	**Prompt:** Please read carefully and indicate how much you agree with each of the following statements about childhood vaccinations. Likert-type scales from 1 = "Strongly disagree" to 7 = "Strongly agree"			
	INT1	I intend to vaccinate my children.	5.642	1.254
	INT2	I want my children to receive all their recommended vaccines.	5.616	1.239
	INT3	I will try to comply with all the vaccination guidelines for my children.	5.549	1.280
Attention traps	AT1	Please select "Disagree" for this question.		
	AT2	If two plus three is equal to five, then select "Somewhat agree" for this question.		

Online Appendix B. Factorial Validity and Reliability

Convergent and Discriminant Validity to Establish Factorial Validity

Convergent and discriminant validity are interrelated concepts that should be confirmed prior to testing the structural model. Convergent validity "is the extent to which a measure correlates positively with alternative measures of the same construct" [12, p. 112]. Convergent validity techniques are used to establish that all items intended to measure a construct "converge, or show significant, high correlations with one another" [22, p. 391]. Discriminant validity "is the extent to which a construct is truly distinct from other constructs by empirical standards" [12, p. 115]. Discriminant validity techniques are used to establish that items for a construct "differ from those [items] that are not believed to make up the construct" [22, p. 389]. In our pre-analysis, we used two techniques to establish convergent and two techniques to establish discriminant validity.

First, we examined the outer model loadings to establish convergent validity; see Table B.1. Reasonably high outer loadings for the items are desired to establish convergent validity. Hair et al. [12] suggest that to establish convergent validity all outer loadings should be statistically significant with outer loading higher than or equal to 0.70. In Table B.1, we present the outer loading, the t-values, and the level of significance for each item. All items in our model had statistically significant outer loadings above or equal to 0.70, except PBC5 that has a loading of 0.673. For large sample sizes, loadings > 0.50 are acceptable [11] and PBC is a control, therefore we retain the item. Second, we examined the cross-loading matrix generated by SmartPLS, provided in Table B.2. The recommendation is that all items for a latent variable should load highest on that latent variable and have no problematic cross-loadings (difference between the two highest loadings should be > 0.100), which also helps establish discriminant validity. We found that our items load highest on the expected latent variable and there are no problematic cross-loadings. Both techniques indicate the convergent validity of our instrumentation.

To establish discriminant validity, we again examined the cross-loading matrix provided in Table B.2. The recommendation is that the difference between the loading on the primary latent variable and any other loading should be > 0.100 [15]. Table B.2 shows that these conditions are met, which establishes discriminant validity. We used the Fornell-Larcker criterion as our second technique for establishing discriminant validity [9]. The Fornell-Larcker criterion compares the square root of the average variance extracted (AVE) of each latent variable with the latent variable correlations. The recommendation is that the square root of each latent variable's AVE "should be greater than its highest correlation with any other [latent variable]" [12, p. 116], which is the case for all latent variables in our model. These results indicate the discriminant validity of our instrumentation.

Establishing Reliabilities

Reliability statistics provide an evaluation of "the extent to which the respondent can answer the same questions or close approximations the same way each time" [23, p. 151]. We obtained the Cronbach's alpha (CA) and the construct reliability (CR) statistic for each latent variable of our model from SmartPLS; statistics are shown in Table B.3. The recommendations are that the CA and CR should be greater than 0.700 and greater than the AVE (which should be > 0.500) [5, 9]. Table B.3 shows that all our CAs and CRs are greater than 0.700 and greater than the corresponding AVEs. All of the AVEs are greater than 0.500. These statistics indicate acceptable reliability of our instrumentation.

Establishing the Lack of Common Methods Bias

Our research methods were designed to counteract common method bias (CMB). Specifically, we ensured participant anonymity and randomized the presentation of the items to the participants, as recommended by the literature **[3, 16, 19]**. The Qualtrics™ survey platform was used to distribute the survey, which enabled

us to share an anonymous link, prevented us from collecting personal information from respondents, and allowed us to randomize items.

Additionally, we conducted two statistical checks to detect the presence of CMB. First, we examined the latent variable correlation matrix shown in Table B.3. The recommendation is that latent variable correlations should not be unreasonably high (> 0.90) [17]. An examination of the constructs correlations Table B.3 shows that all latent variable correlations are < 0.90. Second, we employed the common latent factor (CLF) technique, which is the current standard to test for CMB [20]. The CLF technique is intended to "measure the influence of a common latent method factor on each individual indicator in the model versus the influence of each indicator's corresponding construct [19]" [15, Appendix p. 3]. To perform this check, first, a common latent method factor is created that includes all indicators. Then, separate constructs are created for each indicator. The common latent method factor is connected to each of the single indicator constructs and the original latent variable is also connected to each of the single-item constructs. The model is run, and the outer loadings for each item on the common latent method factor and the outer loadings on the original latent constructs are obtained; Table B.4 shows all these loadings. The average of the outer loadings and the variance explained should be substantially higher for the original latent constructs than for the CLF loadings. Our check indicates that CMB is unlikely to be an issue to our research methods given that the average loading on the original constructs is 0.831 and explains 69.283% of the variance, which is substantially higher than the average loading of 0.258 and the 36.001% of the variance explained by the CLF.

Checking for Multicollinearity

The variance inflation factors (VIFs) are commonly used statistics to check for multicollinearity. According to current guidance, a VIF value greater than 5.0 indicates potential multicollinearity problems because it suggests that greater than 80% of the variance is accounted for by the remaining items in the model **[12]**. Table B.5 shows the VIFs for all of our items. The highest VIF is 2.967, which is below the acceptable level and indicates multicollinearity issues are not present in our instrumentation.

Summary of Pre-analysis Validation

The pre-analyses provided in this appendix demonstrate (1) evidence of convergent and discriminant validity, (2) adequate scale reliabilities, (3) that common method bias is unlikely to be an issue in our study, and (4) no multicollinearity issues. Our pre-analyses results illustrate that our model meets the standard anticipated for SEM-PLS-based analysis [5, 12, 15, 18].

Table B.1. Outer Model Loadings to Establish Convergent Validity

Latent construct	Items	Outer loading	t-statistic
Perceived fairness of Facebook algorithm (PFFA); adapted from Kamdar et al. [13]	PFFA1	0.877	64.604***
	PFFA2	0.895	74.115***
	PFFA3	0.878	66.706***
	PFFA4	0.863	48.998***
Perceived accountability for Facebook algorithm (AFA); adapted from Vance et al. [24]	AFA1	0.798	33.944***
	AFA2	0.825	38.724***
	AFA3	0.797	28.955***
Perceived transparency of Facebook algorithm (TFAB); adapted from Zhou et al. [26]	TFAB1	0.810	35.767***
	TFAB2	0.822	37.486***
	TFAB3	0.846	50.198***
	TFAB4	0.829	44.654***
Negative attitude toward vaccination (VHR); adapted from Askelson et al. [2]	VHR1	0.916	74.971***
	VHR2	0.919	104.248***
	VHR3	0.913	85.532***
Facebook anti-vaccination descriptive (DIOSN); adapted from Göckeritz et al. [10]	DIOSN1	0.814	38.571***
	DIOSN2	0.814	34.455***
	DIOSN3	0.836	45.409***
	DIOSN4	0.826	38.404***
Facebook anti-vaccination injunctive norms (IOSN) adapted from Göckeritz et al. [10]	IOSN1	0.855	52.872***
	IOSN2	0.859	52.594***
	IOSN3	0.856	54.406***
Perceived behavioral control over vaccination (PBC); adapted from Fishbein and Ajzen [8]	PBC1	0.784	33.913***
	PBC2	0.732	24.478***
	PBC3	0.734	25.344***
	PBC4	0.764	29.099***
	PBC5	0.673	15.261***
Intention to vaccinate (INT); adapted from TPB intention Bock et al. [4]	INT1	0.850	49.865***
	INT2	0.824	33.070***
	INT3	0.862	53.990***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, (n/s) = not significant.

Table B.2. Cross-Loadings

Items	PFFA	AFA	TFAB	VHR	DOSN	IOSN	INT	PBC
PFFA1	0.877	0.442	0.578	0.392	-0.465	-0.440	0.257	0.296
PFFA2	0.895	0.425	0.625	0.425	-0.482	-0.479	0.251	0.299
PFFA3	0.878	0.406	0.627	0.420	-0.471	-0.466	0.224	0.307
PFFA4	0.864	0.392	0.624	0.451	-0.436	-0.425	0.214	0.300
AFA1	0.349	0.798	0.444	0.198	-0.388	-0.385	0.304	0.376
AFA2	0.383	0.827	0.431	0.227	-0.410	-0.388	0.289	0.415
AFA3	0.417	0.799	0.401	0.194	-0.390	-0.396	0.280	0.397
TFAB1	0.543	0.427	0.811	0.371	-0.430	-0.414	0.223	0.308
TFAB2	0.554	0.422	0.823	0.380	-0.421	-0.411	0.244	0.332
TFAB3	0.605	0.455	0.846	0.389	-0.470	-0.437	0.267	0.391
TFAB4	0.606	0.437	0.829	0.373	-0.444	-0.442	0.267	0.324
VHR1	0.420	0.219	0.401	0.916	-0.167	-0.159	-0.175	0.063
VHR2	0.468	0.258	0.430	0.918	-0.191	-0.183	-0.139	0.073
VHR3	0.430	0.223	0.424	0.913	-0.174	-0.199	-0.150	0.077
DOSN1	-0.399	-0.376	-0.426	-0.147	0.814	0.621	-0.476	-0.417
DOSN2	-0.435	-0.402	-0.405	-0.156	0.814	0.662	-0.458	-0.388
DOSN3	-0.460	-0.398	-0.436	-0.149	0.836	0.645	-0.528	-0.403
DOSN4	-0.442	-0.436	-0.487	-0.185	0.827	0.665	-0.505	-0.375
IOSN5	-0.424	-0.384	-0.418	-0.147	0.676	0.855	-0.526	-0.428
IOSN6	-0.440	-0.430	-0.460	-0.196	0.689	0.858	-0.537	-0.435
IOSN7	-0.460	-0.423	-0.445	-0.161	0.657	0.855	-0.492	-0.386
INT1	0.235	0.296	0.227	-0.162	-0.523	-0.500	0.850	0.425
INT2	0.226	0.332	0.283	-0.118	-0.469	-0.505	0.826	0.493
INT3	0.223	0.288	0.260	-0.148	-0.526	-0.533	0.863	0.448
PBC1	0.230	0.361	0.246	0.008	-0.377	-0.390	0.447	0.785
PBC2	0.269	0.318	0.311	0.000	-0.325	-0.353	0.401	0.733
PBC3	0.201	0.373	0.283	0.064	-0.330	-0.314	0.379	0.735
PBC4	0.221	0.404	0.320	0.025	-0.375	-0.362	0.432	0.763
PBC5	0.379	0.360	0.387	0.244	-0.377	-0.388	0.305	0.674

Table B.3. Fornell-Larcker Criterion and Reliability Statistics

Latent construct	CA	CR	AVE	PFFA	AFA	TFAB	VHR	DOSN	IOSN	INT	PBC
PFFA	0.902	0.931	0.772	**0.879**							
AFA	0.734	0.850	0.653	0.474	**0.808**						
TFAB	0.846	0.897	0.684	0.698	0.527	**0.827**					
VHR	0.904	0.940	0.839	0.480	0.256	0.457	**0.916**				
DOSN	0.841	0.893	0.677	-0.528	-0.491	-0.534	-0.194	**0.823**			
IOSN	0.818	0.892	0.733	-0.515	-0.482	-0.515	-0.197	0.788	**0.856**		
INT	0.802	0.883	0.716	0.269	0.360	0.303	-0.169	-0.598	-0.606	**0.846**	
PBC	0.793	0.857	0.546	0.342	0.490	0.410	0.078	-0.480	-0.486	0.537	**0.739**

Note: Bolded, underlined values represent the square root of the AVEs

Table B.4. Common Latent Method Factor Analysis

Items	Substantive factor loadings (s)	Variance explained (s^2)	Method factor loading (m)	Variance explained (m^2)
PFFA1	0.878	0.771	0.686	0.471
PFFA2	0.894	0.799	0.711	0.506
PFFA3	0.877	0.769	0.698	0.487
PFFA4	0.865	0.748	0.678	0.460
AFA1	0.796	0.634	0.553	0.306
AFA2	0.828	0.686	0.576	0.332
AFA3	0.800	0.640	0.563	0.317
TFAB1	0.812	0.659	0.637	0.406
TFAB2	0.826	0.682	0.647	0.419
TFAB3	0.844	0.712	0.698	0.487
TFAB4	0.827	0.684	0.674	0.454
VHR1	0.919	0.845	0.358	0.128
VHR2	0.915	0.837	0.397	0.158
VHR3	0.914	0.835	0.379	0.144
DOSN1	0.818	0.669	-0.662	0.438
DOSN2	0.819	0.671	-0.671	0.450
DOSN3	0.833	0.694	-0.693	0.480
DOSN4	0.822	0.676	-0.703	0.494
IOSN1	0.857	0.734	-0.685	0.469
IOSN6	0.854	0.729	-0.715	0.511
DIOSN7	0.857	0.734	-0.692	0.479
INT1	0.850	0.723	0.503	0.253
INT2	0.826	0.682	0.522	0.272
INT3	0.862	0.743	0.520	0.270
PBC1	0.773	0.598	0.492	0.242
PBC2	0.730	0.533	0.479	0.229
PBC3	0.739	0.546	0.462	0.213
PBC4	0.753	0.567	0.504	0.254
PBC5	0.701	0.491	0.557	0.310
Averages:	0.831	69.283%	0.258	36.001%

Table B.5. Collinearity Statistics

Latent construct	Items	VIF
Perceived fairness of Facebook algorithm (PFFA); adapted from Kamdar et al. [13]	PFFA1	2.575
	PFFA2	2.826
	PFFA3	2.529
	PFFA4	2.375
Perceived accountability for Facebook algorithm (AFA); adapted from Vance et al. [24]	AFA1	1.418
	AFA2	1.523
	AFA3	1.432
Perceived transparency of Facebook algorithm (TFAB); adapted from Zhou et al. [26]	TFAB1	1.788
	TFAB2	1.886
	TFAB3	2.005
	TFAB4	1.871
Negative attitude toward vaccination (VHR); adapted from	VHR1	2.967
	VHR2	2.875

Latent construct	Items	VIF
Askelson et al. [2]	VHR3	2.840
Facebook anti-vaccination descriptive norms (DIOSN); adapted from Göckeritz et al. [10]	DOSN1	1.832
	DOSN2	1.854
	DOSN3	1.930
	DOSN4	1.846
Facebook anti-vaccination injunctive norms (IOSN) adapted from Göckeritz et al. [10]	IOSN1	1.828
	IOSN2	1.793
	IOSN3	1.827
Intention to vaccinate (INT); adapted from TPB intention Bock et al. [4]	INT1	1.765
	INT2	1.616
	INT3	1.834
Perceived behavioral control over vaccination (PBC); adapted from Fishbein and Ajzen [8]	PBC1	1.605
	PBC2	1.490
	PBC3	1.528
	PBC4	1.554
	PBC5	1.402

References for Supplementary Appendices

1. Ajzen, I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 2 (1991), 179-211.
2. Askelson, NM; Campo, S; Lowe, JB; Smith, S; Dennis, LK; and Andsager, J. Using the theory of planned behavior to predict mothers' intentions to vaccinate their daughters against HPV. *The Journal of School Nursing*, 26, 3 (2010), 194-202.
3. Bagozzi, RP. Measurement and meaning in information systems and organizational research: Methodological and philosophical foundations. *MIS Quarterly*, 35, 2 (2011), 261-292.
4. Bock, G-W; Zmud, RW; Kim, Y-G; and Lee, J-N. Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Quarterly*, 29, 1 (2005), 87-111.
5. Chin, WW. The partial least squares approach for structural equation modeling. *Modern methods for business research.*, *Methodology for business and management.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 1998, pp. 295-336.
6. Cialdini, RB; Reno, RR; and Kallgren, CA. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, 6 (1990), 1015-1026.
7. Diakopoulos, N. Accountability in algorithmic decision making. *Communications of the ACM*, 59, 2 (2016), 56-62.
8. Fishbein, M and Ajzen, I. *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press, 2010.
9. Fornell, C and Larcker, DF. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 1 (1981), 39-50.
10. Göckeritz, S; Schultz, PW; Rendón, T; Cialdini, RB; Goldstein, NJ; and Griskevicius, V. Descriptive normative beliefs and conservation behavior: The moderating roles of personal involvement and injunctive normative beliefs. *European Journal of Social Psychology*, 40, 3 (2010), 514-523.
11. Hair, JF; Black, WC; Babin, BJ; Anderson, RE; and Tatham, RL. *Multivariate Data Analysis*. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
12. Hair, JF; Hult, GTM; Ringle, C; and Sarstedt, M. *A primer on partial least squares structural equation modeling (PLS-SEM), Second Edition*. Thousand Oaks, CA: Sage Publications, 2017.

13. Kamdar, D; McAllister, DJ; and Turban, DB. "All in a day's work": How follower individual differences and justice perceptions predict OCB role definitions and behavior. *Journal of Applied Psychology*, 91, 4 (2006), 841-855.
14. Kemper, J and Kolkman, D. Transparent to whom? No algorithmic accountability without a critical audience. *Information, Communication & Society*, 22, 14 (2019), 2081-2096.
15. Lowry, PB and Gaskin, J. Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication*, 57, 2 (2014), 123-146.
16. MacKenzie, SB; Podsakoff, PM; and Podsakoff, NP. Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35, 2 (2011), 293-334.
17. Pavlou, PA; Liang, H; and Xue, Y. Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective. *MIS Quarterly*, 31, 1 (2007), 105-136.
18. Peng, DX and Lai, F. Using partial least squares in operations management research: A practical guideline and summary of past research. *Journal of Operations Management*, 30, 6 (2012), 467-480.
19. Podsakoff, PM; MacKenzie, SB; Lee, J-Y; and Podsakoff, NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 5 (2003), 879-903.
20. Schwarz, A; Rizzuto, T; Carraher-Wolverton, C; Roldán, JL; and Barrera-Barrera, R. Examining the Impact and Detection of the Urban Legend of Common Method Bias. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, 48, 1 (2017), 93-119.
21. Shin, D and Park, YJ. Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, 98, 1 (2019), 277-284.
22. Straub, D; Boudreau, M-C; and Gefen, D. Validation guidelines for IS positivist research. *Communications of the Association for Information Systems*, 13, 1 (2004), 380-426.
23. Straub, DW. Validating instruments in MIS research. *MIS Quarterly*, 13, 2 (1989), 147-169.
24. Vance, A; Lowry, PB; and Eggett, DL. Increasing accountability through the user interface design artifacts: A new approach to addressing the problem of access-policy violations. *MIS Quarterly*, 39, 2 (2015), 345-366.
25. Yang, K and Stoyanovich, J. Measuring fairness in ranked outputs. Presented at *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, Chicago, IL, 2017, pp. 1-6.
26. Zhou, L; Wang, W; Xu, JD; Liu, T; and Gu, J. Perceived information transparency in B2C e-commerce: An empirical investigation. *Information & Management*, 55, 7 (2018), 912-927.