

# Algorithm Overdependence: How the Use of Algorithmic Recommendation Systems Can Increase Risks to Consumer Well-Being

Sachin Banker and Salil Khetani

## Abstract

Consumers increasingly encounter recommender systems when making consumption decisions of all kinds. While numerous efforts have aimed to improve the quality of algorithm-generated recommendations, evidence has indicated that people often remain averse to superior algorithmic sources of information in favor of their own personal intuitions (a type II problem). The current work highlights an additional (type I) problem associated with the use of recommender systems: algorithm overdependence. Five experiments illustrate that, stemming from a belief that algorithms hold greater domain expertise, consumers surrender to algorithm-generated recommendations even when the recommendations are inferior. Counter to prior findings, this research indicates that consumers frequently depend too much on algorithm-generated recommendations, posing potential harms to their own well-being and leading them to play a role in propagating systemic biases that can influence other users. Given the rapidly expanding application of recommender systems across consumer domains, the authors believe that an appreciation and understanding of these risks is crucial to the effective guidance and development of recommendation systems that support consumer interests.

## Keywords

artificial intelligence, consumer decision making, decision aids, human–computer interaction, recommender systems

Online supplement: <https://doi.org/10.1177/0743915619858057>

Recent developments in information technology are increasingly contributing to the automation of demand: by leveraging individual preference data, recommender systems facilitate both consumer search and choice. Automated recommendations are now ubiquitous in consumer domains as wide ranging as shopping online (Amazon), choosing movies (Netflix), discovering new music (Spotify), selecting financial investments (Wealthfront), and even dating (Tinder). Although recommender systems often aim to assist consumers in filtering information and improving overall decision quality (Häubl and Trifts 2000; West et al. 1999), in this work we identify a dangerous risk: that consumers display overdependence on algorithmic recommendations in a manner that may both reduce their own welfare and propagate biases system-wide. The current research examines how people interact with automated recommender systems to advance the development of tools and policies that may attenuate these risks.

Various methodologies have been applied to recommend new products to consumers based on ratings of other products (e.g., a new movie suggestion based on previously enjoyed movies). Some algorithms implement content-based methods that leverage consumer preference and product attribute data, others use collaborative filtering methods that build on comparisons to peer users, and yet others apply hybrid approaches that combine both methods (Adomavicius and Tuzhilin 2005; Konstan and Riedl 2012; Resnick and Varian 1997). These tools have recently been developed further in marketing settings (Ansari, Essegai, and Kohli 2000; Bodapati 2008; Chung and Rao 2012; Ghose, Ipeiritos, and Li 2012; Lu, Xiao,

---

Sachin Banker (corresponding author) is Assistant Professor of Marketing, Eccles School of Business, University of Utah, USA (email: [sachin.banker@eccles.utah.edu](mailto:sachin.banker@eccles.utah.edu)). Salil Khetani is Assistant Professor of Marketing, School of Business, Computer Science, & Information Technology, Heidelberg University, USA (email: [skhetani@heidelberg.edu](mailto:skhetani@heidelberg.edu)).

and Ding 2016; Ying, Feinberg, and Wedel 2006); by and large, such methodological developments have aimed to yield more accurate recommendations believed to improve consumer outcomes by reducing search costs, increasing decision quality, and delivering greater satisfaction (Häubl and Trifts 2000; Xiao and Benbasat 2007).

Yet despite efforts to improve algorithm quality, evidence suggests that consumers frequently reject algorithm-generated recommendations. People are averse to following recommendations that adopt algorithmic rules, as has been documented across the contexts of employment decisions (Diab et al. 2011), legal decisions (Eastwood, Snook, and Luther 2012), and medical decisions (Shaffer et al. 2013). They are reluctant to use algorithms to assist in management decisions (Fildes and Goodwin 2007), avoid adopting software packages to improve their marketing forecasting performance (Sanders and Manrodt 2003), and fail to apply more accurate algorithmic diagnosis methods in psychiatric contexts (Vrieze and Grove 2009). Furthermore, when provided with inferior product recommendations, consumers display reactance toward recommender systems by actively avoiding these suggestions (Fitzsimons and Lehmann 2004), leading to sustained aversion to algorithms when making subsequent choices (Dietvorst, Simmons, and Massey 2015). Overall, these findings indicate that consumers are often skeptical of algorithmic recommendations, typically preferring to follow their own intuitions (or those of another human) when making decisions.

The current research, in contrast, aims to highlight that consumers often have a tendency to depend too much on algorithm-generated recommendations, so much so that they may even select inferior products and services to their own detriment. We refer to this tendency to rely on algorithmically generated recommendations even when they are inferior as algorithm overdependence. Prior research has conjectured that consumers “surrender to technology” in modern digital environments because they have excessive faith within online information exchanges, leading them to unwittingly give up sensitive personal information (Walker 2016). Positing that users face the most significant privacy vulnerabilities when factors related to the complexity (i.e., user cognitive limitations) and context (i.e., marketplace asymmetries) of the information exchange coincide, the surrendering-to-technology framework has reshaped how policy makers address such risks. We extend the framework by proposing that surrendering-to-technology phenomena are not limited to information exchanges only but also manifest within other digital interactions, particularly when making purchase decisions with the aid of algorithmic recommender systems. Furthermore, we introduce an additional complexity factor to the framework (i.e., limited understanding) relevant to consumer purchase decisions. When aggravating factors related to the complexity and context exist simultaneously, we posit that consumers will exhibit algorithm overdependence by selecting inferior product offerings.

## Accuracy of Recommendation Systems

Although efforts have been made to improve the accuracy of recommender systems, consumers frequently face vulnerable situations in the modern-day digital marketplace, where biased and inferior recommendations are quite prevalent. Consumers often have difficulty articulating their own preferences (e.g., Bettman, Luce, and Payne 1998; Nordgren and Dijksterhuis 2008; Wilson and Schooler 1991), and firms often have difficulty measuring their preferences (Mullainathan and Obermeyer 2017), which can lead to suboptimal recommendations borne out of improper inputs. Marketplace observations have also revealed that several online retailers steer users toward more expensive products by placing more profitable items at the top of suggested purchase lists, even when they are not of high quality (Angwin and Mattu 2016; Hannak et al. 2014; Mikians et al. 2012). Startlingly, additional studies have discovered that recommendations can often be discriminatory, delivering fewer high-income job recommendations to women (Datta, Tschantz, and Datta 2015) and significantly more ads for arrest records on searches for distinctively black names (Sweeney 2013), in spite of the fact that these outcomes were not intended by the algorithm developers. While few regulations currently govern the implementation of recommendation systems, potential consumer vulnerabilities could be addressed with expansions to existing rules governing digital advertising (Federal Trade Commission 2013) and the use of algorithms in automated decision making (Wachter, Mittelstadt, and Floridi 2017).

In orienting research examining algorithm-aided decisions, we suggest that the problems encountered can be understood through the lens of type I and type II errors. That is, in cases of algorithm aversion (e.g., Diab et al. 2011; Dietvorst, Simmons, and Massey 2015; Shaffer et al. 2013), people fail to adopt the recommendations of a superior algorithm and instead favor the use of their own personal intuitions, similar to statistical type II error (false negatives), in which true propositions are incorrectly deemed untrue. Under ideal circumstances, users would indeed adopt recommendations when they are superior (commonly assumed behavior; Adomavicius and Tuzhilin 2005) and avoid recommendations when they are inferior (as in recommendation reactance; Fitzsimons and Lehmann 2004). The current work emphasizes a countervailing issue in which users adopt recommendations even when they reduce consumer welfare. This issue is akin to a type I problem (false positive) in which false propositions are incorrectly deemed true. Improving the efficacy of algorithmic aids will thus require balancing both type I and type II problems (see Table 1).

Although dystopian warnings around the overreliance on artificially intelligent systems may seem distant (e.g., HAL in Stanley Kubrick's *2001: A Space Odyssey*), the budding risks associated with the increased use of recommender systems are deserving of present-day concern. As the adoption and use of recommender systems within online retailing interfaces, interactive agents, and smart devices becomes increasingly widespread, these risks are likely to be magnified. Thus,

**Table I.** Classification of Problems Associated with Algorithm-Driven Recommendations.

	Superior Recommendations	Inferior Recommendations
Avoid Recommendations	<ul style="list-style-type: none"> <li>• Algorithm aversion</li> <li>• Type II problem (false negative)</li> </ul>	<ul style="list-style-type: none"> <li>• Recommendation reactance</li> <li>• Ideal outcome</li> </ul>
Adopt Recommendations	<ul style="list-style-type: none"> <li>• Assumed behavior</li> <li>• Ideal outcome</li> </ul>	<ul style="list-style-type: none"> <li>• Algorithm overdependence</li> <li>• Type I problem (false positive)</li> </ul>

appreciation of both type I and type II problems is crucial to guiding the development of recommender systems in a way that serves to simultaneously maximize consumer decision quality while minimizing potential harms. As we caution in this work, without adequate guidance, the widespread use of recommender systems in automating demand has the potential to harm consumer welfare at a considerably large scale.

## Theoretical Background and Framework

Extant literature regarding interactions with algorithmic decision aids has primarily described phenomena in which consumers display aversion to recommendations. For instance, when experts recommended a low-quality (dominated) choice option, consumers were actually more likely to reject it, selecting the high-quality (nondominated) option 23% more often (Fitzsimons and Lehmann 2004). Thus, existing predictions within the literature suggest that displaying inferior recommendations to users may counterintuitively improve consumer decision quality. We instead posit that consumers are prone to rely too much on algorithms when there is a convergence of complexity factors (limited attention, effort, and understanding) and contextual factors (vulnerabilities, risks, and uncertainties).

### Complexity Challenges in the Use of Algorithmic Recommender Systems

Given the increasingly complex decision environments that consumers operate within, various limitations in the human capacities to be able to attend to, process, and understand large amounts of information can expose consumers to vulnerabilities that may reduce welfare. We expand on these complexity factors next.

*Limits of attention.* Shopping for products online is becoming an increasingly overwhelming undertaking. The expanding range of product offerings, number of customizable attributes, and array of different retailers, ratings, and reviews one may sort through can make choosing difficult endeavor for consumers (Dellaert and Stremersch 2005). Because consumers selectively allocate limited attentional resources, they may be unable to fully evaluate every product even when information

is easily accessible. Just as people may ignore a gorilla walking across a basketball court when focused on a task (Simons and Chabris 1999), they may also overlook important details when simply seeking a product that meets a need. Moreover, managing many different goals simultaneously can impose time pressures that magnify attentional biases (Louro, Pieters, and Zeelenberg 2007). As a consequence, people may rely on summary information and recommendations when making purchasing decisions online.

*Limits of cognitive effort and capacity.* Even when people are capable of devoting full attention to a task, they can be limited by their ability to fully process information. Because making sense of information requires effortful cognitive operations that carry an aversive phenomenology, individuals often behave as if they are distributing a limited mental resource (Baumeister, Tice, and Vohs 2018; Kurzban et al. 2013). Consumers are more likely to defer their decisions, experience decreased satisfaction and confidence, and face greater regret when processing choice options requires greater effort (Chernev, Böckenholt, and Goodman 2015). Thus, people who are unwilling to expend effort are more inclined to rely on shortcuts such as recommendations when making decisions (Banker et al. 2017).

*Limits of understanding.* As products become more complex through the advent of more advanced electronics, specialized medical procedures, and esoteric financial products, consumers more frequently lack expertise and understanding of their decision domains. Consequently, they may develop inaccurate mental models around how they believe a product or service works and may not understand what features they truly should care about. People who are more experienced within product categories do not need to expend as much effort to evaluate product information and make decisions, can make finer discriminations between options, and are able to process more information more deeply relative to those who are inexperienced (Alba and Hutchinson 1987); however, those who lack expertise in a domain can gain it through the repeated exposure to products and attributes. Greater expertise within a product domain reduces perceived complexity (Dellaert and Stremersch 2005), susceptibility to framing effects (Coupey, Irwin, and Payne 1998), and the tendency to construct preferences on the fly (Bettman, Luce, and Payne 1998). Thus, low-expertise consumers may feel as though they do not understand complex product decisions and therefore have a greater need to rely on decision aids such as recommender systems.

### Contextual Challenges in the Use of Algorithmic Recommender Systems

In addition to the human limitations encountered when making complex decisions, marketplace characteristics also contribute to the overdependence on algorithms. We expand on these contextual factors next.

**Vulnerabilities.** Recommender system algorithms can expose consumers to vulnerable situations in which suggestions may not be aligned with consumer interests through several mechanisms. For example, data supplied to recommender systems are subject to considerable measurement error due to user inability to accurately articulate preferences (e.g., Bettman, Luce, and Payne 1998; Wilson and Schooler 1991) that can lead to faulty recommendations. Even when developers aim to maximize accuracy, a recommender system may continue to generate biased recommendations due to various sources of mismeasurement (Mullainathan and Obermeyer 2017). Assuming the data supplied by the user are unbiased, collaborative filtering algorithms may still deliver biased recommendations because they tend to push people toward popular products and avoid items with limited historical data (Fleder and Hosanagar 2009), and such systems are also easily manipulated by fake profiles and purchases (Mobasher et al. 2007). Furthermore, because firms frequently have conflicting incentives to increase sales, ad exposure, user engagement, and achieve other strategic goals, these motives can bias the presentation of products on retailer websites (Angwin and Mattu 2016; Hannak et al. 2014; Mikians et al. 2012). Customers are typically not informed of the additional factors beyond accuracy that led to certain items being top-ranked or recommended.

**Risks and uncertainties.** The increasing complexity of products and services that consumers purchase contributes to risks and uncertainties in online shopping decisions. Due to informational asymmetries, while firms typically have a complete understanding of products being sold, consumers often do not possess the same level of detail (e.g., regarding quality or other unmentioned unobservable product attributes). Consumers may thus rely on recommendation systems to manage such uncertainties. Complex algorithms that operate on large amounts of data can create an impression of certainty and expertise, which consumers often seek out prior to making decisions (Burghardt et al. 2017; Paul, Hong, and Chi 2012). However, this imbalance in information can expose consumers to risks of manipulation.

### *The Emergence of Algorithm Overdependence*

Complexity factors (limited attention, effort, or understanding) and contextual factors (vulnerabilities, risks, and uncertainties) can jointly compound individual tendencies to rely on algorithms during decision making. Specifically, when users possess limited understanding of the product domain and recommender systems convey a high degree of certainty, consumers may surrender to whatever the recommendation agent chooses, without regard for the quality of the recommendation. However, if users instead believe that they have greater expertise than the algorithm (with respect to the product domain and their self-knowledge of their own preferences), they will be less inclined to heed the recommendation. Thus, we posit the following hypothesis: when complexity and contextual factors are aligned, consumers will display algorithm overdependence.

This alignment occurs when consumers encounter contexts in which they perceive algorithms to reflect more certain sources of information relative to their own limited understanding of the product domain. Overdependence specifically occurs in situations of vulnerability, wherein recommendations are antithetical to consumer interests.

For example, consider the search for an online tax preparation tool, such as TurboTax. Although lower-income Americans are eligible to file their taxes for free, Google search engine recommendations for free tax filing have instead been shown to link consumers to paid versions of tax preparation software (Elliott 2019; Elliott and Waldron 2019). Because consumers may hold the belief that search algorithms have greater expertise than themselves regarding the vast number of pages online, people may surrender to the recommendations, convinced they were linked to the correct page, and in turn end up paying extraneous fees. It is only when consumers believe to have greater expertise than the search engine algorithm itself (i.e., being certain that free tax filing should be available to them) would they be able to overcome the erroneous recommendations.

## **Overview of the Experiments**

We present a series of studies that aim to understand the impact of recommendations on consumer decision quality, the reasons for consumer adoption of these recommendations, and ways in which interventions may reduce risks of algorithm overdependence. Experiment 1 first assesses how inferior recommendations provided by algorithm-driven recommender systems influence consumer decisions. Experiments 2a and 2b subsequently evaluate different factors that may influence consumer selection of such recommended options, including attention and effort. Experiment 3 then illustrates that algorithm overdependence occurs when complexity and contextual factors are aligned, specifically when consumers perceive algorithms to hold comparatively greater expertise within the product domain. Building on the proposed framework, Experiment 4 finally assesses the efficacy of a consumer education-based intervention approach that may reduce risks to consumers.

## **Experiment 1**

As an initial assessment, we first examined the impact of algorithmic product recommendations on consumer purchasing decisions. While literature on recommendation reactance (Fitzsimons and Lehmann 2004) predicted that inferior recommendations would counterintuitively increase decision quality, the algorithm overdependence framework instead predicted that they would decrease decision quality.

### *Method*

**Participants.** A total of 157 participants (55 women;  $M_{\text{age}} = 24$  years,  $SD = 6.0$ ) completed the study for partial course credit.

**Procedure.** This study involved participants making purchasing decisions with the presence or absence of automated product recommendations. We varied the presence and the nature of recommendations such that people were assigned to one of three conditions: no recommendations, nondominated recommendations, or dominated recommendations. The overall task sequence involved first reading general information about portable chargers (the product domain), sharing preferences to train the recommendation algorithm, and subsequently making a purchase decision from a set of portable chargers (in which the presence and nature of the recommendations was varied).

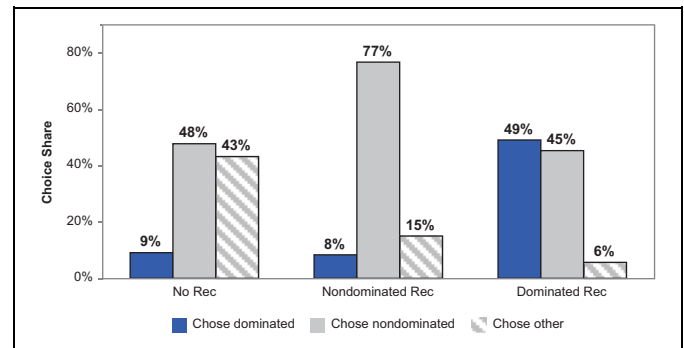
In conditions with recommendations, participants shared information on their preferences and habits (e.g., phone usage, traveling tendencies) to inform the recommendation algorithm. They viewed a brief waiting screen while the recommendations were ostensibly being generated. Unbeknownst to participants, the recommended options were randomly determined to be either dominated (inferior) or nondominated (superior) options regardless of the preference information they shared. This approach allowed us to assess decision quality by observing the selection of decoy options in the choice set (Huber, Payne, and Puto 1982), consistent with prior literature (Diehl 2005; Fitzsimons and Lehmann 2004; Häubl and Trifts 2000; Swaminathan 2003). A no-recommendation control condition served as a baseline comparison.

Subsequently, participants made purchase decisions by selecting a product from a set of six portable chargers. The two nondominated recommendations (Option A and Option B) were identical to the dominated options (Option A' and Option B') in all dimensions, including color and design, except that they either were cheaper or had greater charging capacity; for attribute details, see Table A1. In the control condition, none of the choice options were labeled as recommended. The order of all choice options was randomized.

After making a choice, participants' shared their level of expertise and understanding of the product domain by indicating their decision confidence on a 0%–100% scale (Tsai and McGill 2011). We also asked participants to evaluate the recommendation with regard to its quality, perceived complexity, and personalization (for details, see the Web Appendix).

## Results

**Algorithm overdependence.** Confirming our predictions, participants selected inferior options more frequently after receiving a dominated recommendation (49%) relative to receiving a nondominated recommendation (8.3%) or no recommendation at all (9.1%;  $\chi^2(2, 157) = 33.1, p < .001$ ). Thus, dominated recommendations indeed lowered decision quality (see Figure 1). The delivery of nondominated recommendations did not lower decision quality relative to the no-recommendation control ( $\chi^2 < 1$ ). These findings indicate that participants were more than five times as likely to make inferior decisions when interacting with a biased recommendation system, relative to both the baseline of having no recommendations at all ( $\chi^2(1, 97) = 16.2, p < .001$ ) or when interacting with a superior



**Figure 1.** Percentage of participants choosing dominated, nondominated, and other options in Experiment 1.

recommendation system ( $\chi^2(1, 113) = 21.4, p < .001$ ). Indeed, these participants would have been better off if they had not relied on a recommender system at all.

**Confidence.** Confidence in the decision did not vary significantly across conditions, averaging 74% in the nondominated condition, 73% in the control condition, and 69% in the dominated condition ( $F(2, 153) = 1.16, p = .31$ ). However, participants who had chosen the dominated options were those who reported significantly lower levels of confidence compared with those who did not choose the dominated options (65% vs. 74%;  $t(154) = 2.28, p = .024$ ).

## Discussion

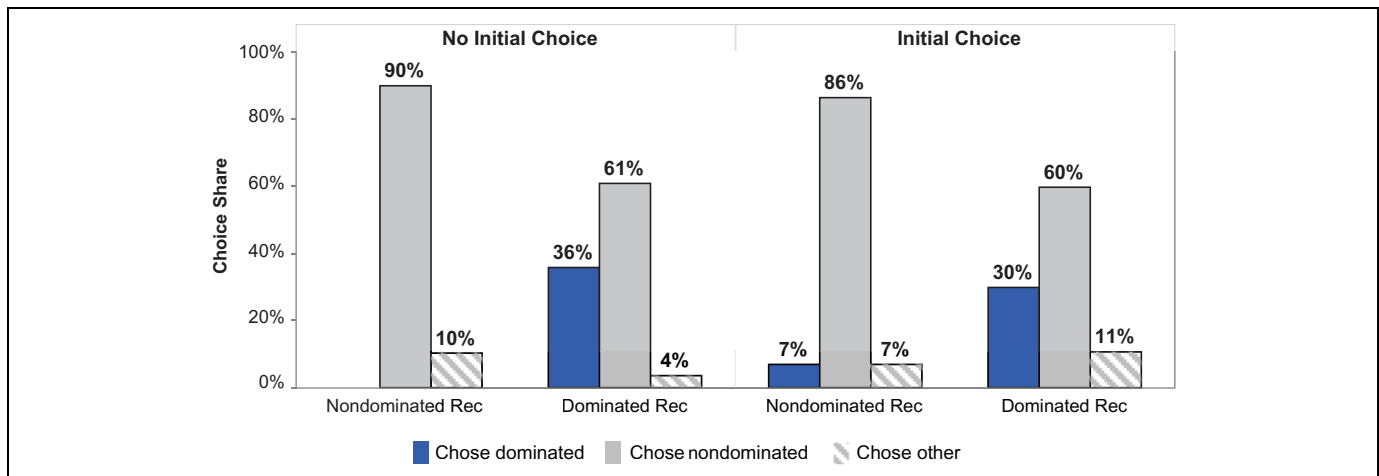
Experiment 1 demonstrated that, counter to recommendation reactance predictions, consumers do in fact rely on algorithm-generated recommendations even when they are inferior. Relative to making decisions without the assistance of a recommender system, the presence of a biased recommendation increased risks of consumer welfare loss by more than a factor of five.

## Overview of Experiments 2a and 2b

Experiments 2a and 2b aimed to evaluate the role of different complexity factors (limited attention, effort, and understanding) in driving algorithm overdependence. Our findings suggest that consumers exhibit algorithm overdependence even when attentional factors are not at play (Experiment 2a) and when choice options require comparatively little processing effort (Experiment 2b). While we anticipate that limits of attention and limits of effort can in many cases exacerbate algorithm overdependence, the current evidence suggests that limited attention and effort are not necessary for the overreliance on algorithms. Our data lend support to the idea that limits of understanding within a product domain can be sufficient to yield algorithm overdependence.

## Experiment 2a

Experiment 2a examined whether algorithm overdependence persisted when attentional limitations were eliminated. In



**Figure 2.** Percentage of participants choosing dominated, nondominated, and other options in each condition in Experiment 2a.

addition, the current experiment addresses the concern that individuals in Experiment 1 may have selected the more expensive, dominated options because the higher price may signal higher quality. Experiment 2a holds the price of products constant across choice options and only manipulates other product features.

### Method

**Participants.** A total of 143 participants (58 women;  $M_{\text{age}} = 24$  years,  $SD = 9.2$ ) completed the study for partial course credit.

**Procedure.** We implemented a  $2 \times 2$  between-participants design varying (1) whether participants made an initial selection before receiving the recommendation and (2) whether the recommended option was a dominated or nondominated product. Participants in this study shopped for a pair of wireless headphones and selected a product from a set of five options that varied on several dimensions (sound quality, dynamic range, noise cancellation, comfort, battery life, and style). Expert reviewer ratings were provided for each attribute for all headphones, and participants were asked to choose the pair they were most interested in receiving. The nondominated option (Option A) was identical in all dimensions to the dominated option (Option A'), except that the nondominated headphones had higher ratings on both sound quality and noise cancellation (see Table A2). Choices were real; the selections for two random participants were implemented at the end of the study.

All participants were first given information about headphones and their various features prior to making a decision. Participants in a condition asking for an initial choice saw all five headphones, made a selection, and shared their level of confidence. Subsequently, participants in all conditions shared their preferences regarding headphone features to inform the recommendation algorithm. A recommendation was generated ostensibly based on this information while participants viewed a waiting screen. All participants then made their final decision

and shared their choice confidence. Finally, participants evaluated the recommendation with regard to its quality, perceived complexity, and personalization.

### Results

**Algorithm overdependence.** We submitted the choice of the dominated option to a logistic regression on the recommendation condition, the initial choice condition, and their interaction. Consistent with algorithm overdependence predictions, the results revealed a significant main effect of the recommendation condition ( $\chi^2(1, 143) = 4.77, p = .029$ ), such that dominated recommendations significantly decreased decision quality (32% vs. 2.9% chose dominated options; see Figure 2). However, we also observed that the main effect of the initial choice condition and the interaction effect were not significant ( $\chi^2\text{s} < 1$ ), suggesting that individuals displayed algorithm overdependence even when they had already evaluated all options in their initial choice. Consumers exposed to inferior recommendations were more than four times as likely to experience lowered decision quality.

**Switching behavior.** After making an initial choice, participants were significantly more likely to switch their selection when they encountered a dominated recommendation (38%) rather than a nondominated recommendation (14%;  $\chi^2(1, 76) = 5.63, p = .018$ ). Notably, participants had selected the nondominated option with equal frequency (77% vs. 83%;  $\chi^2 < 1$ ) during their initial selection, before their interaction with the recommender system; thus, greater switching was not due to differential identification of nondominated options in the initial choice. Rather, in spite of the majority of participants correctly identifying the nondominated choice option during their initial evaluation, participants were significantly less likely to select the nondominated option when they were presented with an inferior recommendation (60% vs. 88%;  $\chi^2(1, 76) = 6.50, p = .011$ ). That is, participants who had initially selected higher-quality options elected to switch their choice to an inferior option that lowered

decision quality when it was recommended by the algorithm; consumers were susceptible to this loss in welfare even when they had indicated preference for superior alternatives only moments earlier.

**Confidence.** In line with the idea that algorithm overdependence occurs when people perceive algorithms as having greater expertise relative to themselves, participants who chose the dominated recommendation also reported significantly lower levels of confidence compared with participants who did not ( $M = 69\%$  vs.  $M = 84\%$ ;  $t(73) = 3.04$ ,  $p < .01$ ).

## Discussion

Experiment 2a demonstrated that biased recommender systems can lower decision quality even when consumers had already evaluated the options and had a choice in mind prior to their interaction with the algorithm. Thus, attentional limitations are not necessary for the emergence of algorithm overdependence. Moreover, we observed that participants actually switched their personally selected choices from high-quality options to inferior options when they interacted with a biased recommender system, counter to algorithm aversion findings.

## Experiment 2b

Experiment 2b next examined the role of limited effort in contributing to algorithm overdependence. In this study, we assessed whether algorithm overdependence persisted when varying the effort required to evaluate choice options. We would expect to observe better-quality decisions when participants make choices involving easy-to-evaluate options, if limited effort is a necessary factor underlying algorithm overdependence.

## Method

**Participants.** A total of 533 participants located in the United States (311 women;  $M_{\text{age}} = 38$  years,  $SD = 13$ ) completed the study on Amazon Mechanical Turk (MTurk) in exchange for a small monetary reward.

**Procedure.** We implemented a  $2 \times 3$  design varying both recommendation quality and decision difficulty. Participants received dominated recommendations, nondominated recommendations, or no recommendations. In addition, to manipulate decision difficulty we varied the number of attributes associated with each choice as either high (nine attributes) or low (two attributes), following prior work (Chernev 2003; Hoch, Bradlow, and Wansink 1999).

Participants shopped for a portable charger in a paradigm identical to Experiment 1. After providing preference information and waiting for the system to generate recommendations, participants then made a selection from six available portable chargers. As before, the nondominated options (Options A and B) were identical to the dominated options (Options A' and B') in all dimensions except that they were either cheaper or had

greater charging capacity. The dominated attributes were held constant across both high-decision-difficulty and low-decision-difficulty conditions. For attribute details, see Table A3. Participants received recommendations for the nondominated options, the dominated options, or no recommendations at all.

Prior to the main study, participants also indicated their level of expertise and understanding with portable chargers on a two-item scale (1 = "not at all familiar/knowledgeable," and 7 = "very familiar/knowledgeable";  $\alpha = .86$ ). After making a decision, participants rated their confidence and completed a three-item decision difficulty scale ( $\alpha = .90$ ; Diehl and Poynor 2010). Finally, participants evaluated the recommendation with respect to quality, perceived complexity, personalization, and trust.

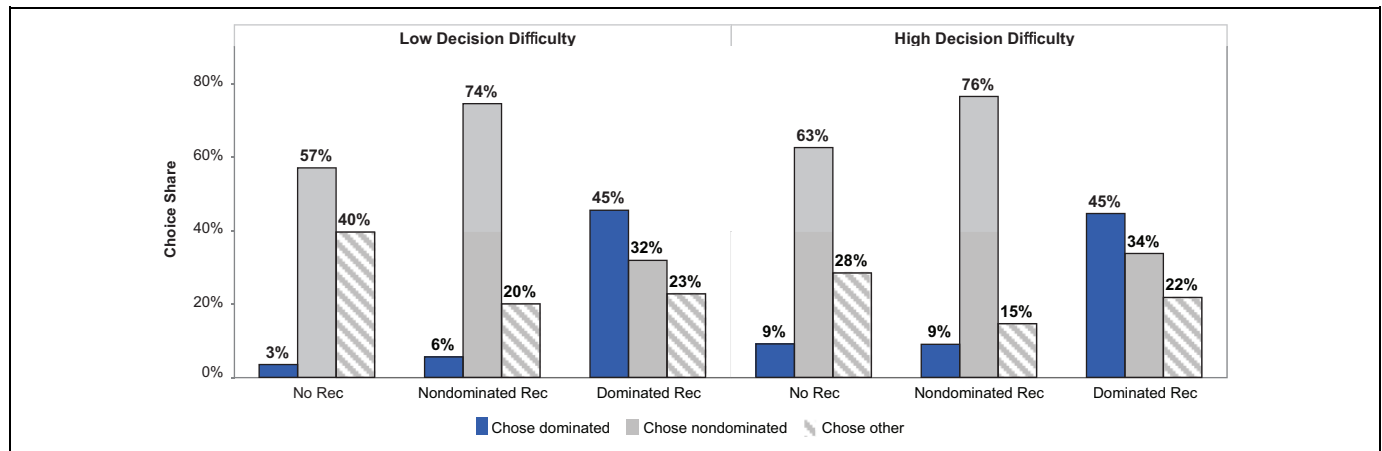
## Results

**Decision difficulty.** As a manipulation check, we confirmed that participants in the nine-attribute condition ( $M = 3.16$ ,  $SD = 1.81$ ) indeed reported experiencing significantly greater difficulty than participants in the two-attribute condition ( $M = 2.85$ ,  $SD = 1.80$ ;  $t(531) = 1.97$ ,  $p = .049$ ). Consistent with the notion that evaluating products with nine attributes requires more effort than evaluating those with two attributes, participants spent significantly more time making decisions in the high-difficulty condition compared with the low-difficulty condition ( $M = 80$  seconds vs.  $M = 42$  seconds;  $t(531) = 6.90$ ,  $p < .001$ ).

**Algorithm overdependence.** Regardless of the decision difficulty, participants selected inferior options significantly more often (thus exhibiting lowered decision quality) when presented with dominated recommendations (45%) rather than nondominated recommendations (7.3%;  $\chi^2(1, 359) = 71.9$ ,  $p < .001$ ) or no recommendations at all (6.3%;  $\chi^2(1, 354) = 75.9$ ,  $p < .001$ ). Again, for these participants, use of the recommender system reduced welfare; exposure to inferior recommendations led to more than a sevenfold increase in risks of lowered decision quality. We observed an identical pattern of significance within each of the decision difficulty conditions (see Figure 3).

**Confidence and consumer expertise.** We again found that the participants who selected the dominated recommendations were those who displayed lower confidence in their decision (74%) relative to those who did not choose them (79%;  $t(178) = 2.05$ ,  $p = .042$ ). In addition, consistent with the idea algorithm overdependence occurs when consumers perceive recommender systems to hold greater expertise relative to consumers themselves, we found that the participants who had selected inferior options were those who expressed lower levels of expertise ( $M = 3.88$ ) within the product domain compared with those who did not choose inferior options ( $M = 4.23$ ;  $t(357) = 1.69$ ,  $p = .091$ ).





**Figure 3.** Percentage of participants choosing dominated, nondominated, and other options in Experiment 2b.

### Discussion

Experiment 2b revealed that participants displayed algorithm overdependence with similar frequencies regardless of whether the decisions they were facing involved comparatively high or low levels of effort. Cognitive effort limitations do not appear to be a necessary factor required for consumer overdependence on algorithms.

### Experiment 3

Experiment 3 examined the interplay between the human-centered complexity factors and marketplace-centered contextual factors involved in driving algorithm overdependence. Specifically, we hypothesized that when consumers perceive algorithms to have greater expertise within a product domain relative to consumers themselves, they will be more inclined to surrender to recommendation systems even when the suggestions are of poor quality.

### Method

**Participants.** A total of 120 participants (46 women;  $M_{\text{age}} = 25$ ,  $SD = 11$ ) completed the study for partial course credit.

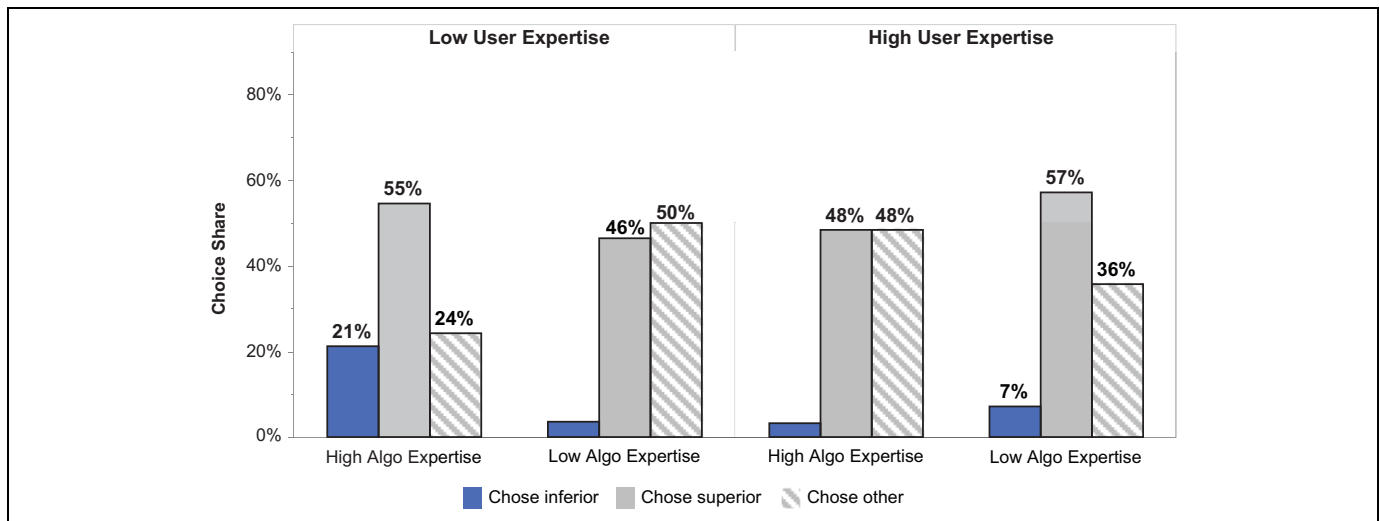
**Procedure.** We implemented a  $2 \times 2$  between-participants design varying both the participants' and the algorithm's domain expertise to assess how the convergence of complexity factors (limits of understanding) and contextual factors (risks and uncertainties) influenced consumer tendencies to adopt inferior recommendations. Participants were asked to consider a situation in which they needed to seek medical treatment. After injuring a muscle and experiencing continued pain, they were told that they now sought to determine which of several possible remedies to pursue. Before participants received a recommendation, we manipulated their domain expertise. Building on prior literature that has established that repeated decision making in a domain develops consumer expertise (Alba and Hutchinson 1987), individuals in the high-user-expertise condition made a series of decisions involving the

choice options under consideration. Specifically, they made ten choices that included all pairwise combinations of the five choice options, allowing users to clearly establish their preferences over these options. A validation study confirmed that this procedure increased participants' perceived domain expertise ( $N = 50$  MTurk participants;  $t(48) = 4.12$ ,  $p < .001$ ). The choice options included five different potential treatment options that the participant could pursue: see a physician costing \$50, see a physiotherapist costing \$120, see a nurse practitioner costing \$10, see a chiropractor costing \$60, or purchase an herbal treatment seen on *The Dr. Oz Show* for \$30. The order of the choices was randomized. Participants in the low-user-expertise condition moved on directly to the next stage of the study after reading about the scenario without making any pairwise choices over the medical treatment options.

Next, participants were informed that they were to be testing a recommender system aimed at assisting patients with medical decisions and shared their preferences regarding different medical treatments (related to cost, familiarity, distance, etc.). Participants were told that the recommender system was being developed in conjunction with the medical school and computer science department. To vary the perceived expertise of the algorithm, we described it as either a highly developed system that integrated hundreds of user responses over an extended period of time (high algorithm expertise) or as a new tool in very early stages of testing that was being developed on the basis of feedback (low algorithm expertise).

After submitting their preferences, participants viewed a waiting screen in which recommendations were ostensibly being generated. They were next shown the five choice options, in which two of the options were recommended by the algorithm. Unbeknownst to the participants, the recommended options were held constant (as both "regular physician" and "physiotherapist"). We elected to study this medical treatment context to evaluate whether algorithm overdependence can also occur within consequential domains in which choice options may not have clearly quantifiable attributes, a common feature of many daily decisions. Because the choice attributes are not quantified, we cannot identify inferior recommendations in the





**Figure 4.** Percentage of participants choosing inferior, superior, and other options in Experiment 3.

strict (dominated vs. nondominated) manner that we applied in previous studies. However, we are still able to identify inferior options in line with their established efficacy (as shared by medical professionals) and their perceived efficacy (as shared by the participants themselves). In situations of continued pain, it is advised by the Mayo Clinic and Cleveland Clinic that patients see a physician. We thus treated choice of the more expensive physiotherapist visit as the inferior option; for convenience we refer to the physician as the superior option. A pretest study ( $N = 151$  students drawn from the same population) also confirmed that lay perceptions coincided with this fact, as the physiotherapist option was chosen with the least frequency (4%, tied with the herbal treatment seen on *The Dr. Oz Show*) when participants were asked to pursue one of the five treatment plans without any recommendation present.

Participants were asked to decide which medical treatment plan they would prefer. They then indicated their level of confidence, evaluated the recommendation as in previous studies, and shared their perceptions of the algorithm's relative knowledge.

## Results

**Algorithm overdependence.** We submitted choice of the recommended option to a logistic regression on the user expertise condition, algorithm expertise condition, and their interaction. The main effects of user expertise and recommender expertise were not significant ( $\chi^2s < 1$ ). However, the results revealed a significant interaction effect ( $\chi^2(1, 120) = 4.78, p = .029$ ). When participants had low domain expertise, they were indeed more inclined to adopt the recommendations of high-expertise algorithms compared with low-expertise algorithms (76% vs. 50%;  $\chi^2(1, 61) = 4.39, p = .036$ ). However, when participants had high domain expertise, the recommendations generated by a high-expertise algorithm were no more likely to be chosen than those generated by a low-expertise

algorithm (52% vs. 64%;  $\chi^2(1, 59) < 1$ ). These results are consistent with the idea that people rely on recommendations when they perceive the algorithm as holding greater domain expertise than themselves.

In addition, we submitted choice of the inferior option to a logistic regression on the user expertise condition, recommender expertise condition, and their interaction. The main effects of user expertise and recommender expertise were again not significant ( $\chi^2s < 1$ ). However, the results revealed a marginal interaction effect ( $\chi^2(1, 120) = 3.27, p = .071$ ). In line with predictions, when participants had low domain expertise, they selected dominated options significantly more frequently when recommended by a high-expertise algorithm rather than a low-expertise algorithm (21% vs. 4%;  $\chi^2(1, 61) = 4.67, p = .031$ ). When participants had high domain expertise themselves, they were no more likely to choose an inferior recommendation from a high-expertise algorithm than from a low-expertise algorithm (3% vs. 7%;  $\chi^2 < 1$ ; see Figure 4).

**Choice confidence.** Again, participants who selected inferior recommendations were those who expressed lower levels of confidence in their decision (71%) relative to participants who did not select inferior recommendations (80%;  $t(118) = 2.24, p = .027$ ).

## Discussion

Experiment 3 illustrated that when participants perceive themselves as holding comparatively less experience in a particular domain, they are more inclined to adopt recommendations generated by the algorithm, even when such recommendations are inferior options. Overall, our evidence indicates that algorithm overdependence does not require limited attention or limited effort (Experiments 2a and 2b) and can occur simply when consumers see algorithms as holding comparatively greater levels of domain expertise.

## Experiment 4

Findings from the prior studies suggest that algorithm overdependence occurs when there is a convergence between human-centered complexity factors and marketplace-centered contextual factors, specifically when people perceive recommender systems as holding greater domain expertise than they possess themselves. Drawing on this understanding, we sought to develop an intervention approach to mitigate associated risks. While type II problems related to algorithm aversion can be addressed by increasing consumer utilization of algorithm-driven decision aids, such interventions will also have the undesirable consequence of exacerbating type I problems. Thus, addressing risks associated with algorithm overdependence requires an opposing intervention approach.

Aiming to target both contextual and complexity factors involved in driving algorithm overdependence, we applied a consumer education intervention focused on alerting people to the prevalence of misinformation that one may encounter online. This intervention served to increase consumer awareness of vulnerabilities by warning individuals about existence of false information they may encounter online. In addition, the intervention aimed to alert consumers to aggravating factors such as their own limits of understanding that can lead people to believe this false information. By cautioning consumers about the vulnerabilities they may face when falling for false information online, consumers may consequently have less faith in the recommender system and be more motivated to direct attention and effort toward making careful decisions rather than surrendering to the technology.

We selected this information-based approach because it presents a simple, low-cost, and easily adaptable strategy that can be implemented at a much larger scale. To understand the extent to which a broad consumer education approach may be able to insulate individuals against the range of different recommendation algorithms consumers may encounter online, we also examined how the intervention influenced behavior in two different website contexts. Specifically, we analyzed the efficacy of the intervention within both suspicious website environments (where consumers have considerable basis to be skeptical of recommendations) and nonsuspicious website environments (where consumers have less basis for skepticism).

### Method

**Participants.** A total of 200 participants located in the United States (123 women;  $M_{\text{age}} = 39$  years,  $SD = 13$ ) completed the study through MTurk in exchange for a small monetary reward.

**Procedure.** We applied a  $2 \times 2$  between-participants design varying the presence of the consumer education intervention (treatment vs. control) and the web environment (suspicious vs. nonsuspicious website). As in Experiment 1, the study asked participants to shop for a portable charger. Procedures were

similar to Experiment 1 apart for a few differences that enabled us to assess the efficacy of the consumer education intervention. First, prior to receiving product recommendations, participants were randomly assigned into the intervention treatment or a no-intervention control. In the control condition, participants did not receive any additional information and simply moved on directly to the recommendations, as in Experiment 1. However, in the treatment condition, participants were presented with an article informing readers of the vulnerabilities associated with the widespread presence of misinformation and the frequency with which consumers believe this false information. The article began as follows:

#### **Algorithms and machine learning are increasingly being used to make more decisions – how should this be approached?**

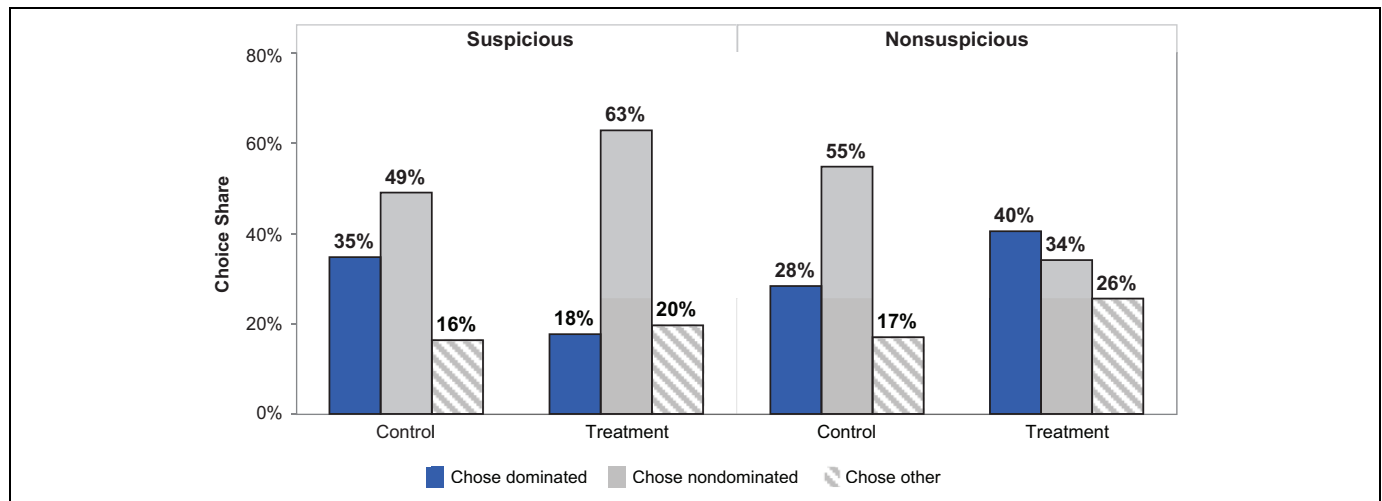
This week, a BuzzFeed survey found that three in four American adults who see fake-news headlines believe them. It's not hard to see why: A website peddling made-up news stories can easily look nearly as polished as The New York Times, and it's impossible to keep up with the sheer volume of information published online every minute. And when people believe fake news stories, bad things can happen.

After the treatment, all participants subsequently made decisions with the assistance of a recommendation system, which was incorporated within either a suspicious or nonsuspicious web environment. The nonsuspicious website presented participants with the same stimuli as those shown in Experiment 1, a standard website with no additional embellishments. However, the suspicious website featured several banner ads and dated design elements such as flashing images that provided users with some basis to question the accuracy of the recommendations. A validation study ( $N = 34$  MTurk participants) confirmed that, relative to the nonsuspicious control website, participants viewed the suspicious website as possessing lower levels of expertise ( $t(32) = 4.24, p < .01$ ) and as being more likely to take advantage of users ( $t(32) = 3.29, p < .01$ ). Thus, the suspicious context evoked a greater sense of vulnerability and uncertainty.

Unbeknownst to the participants and regardless of their individual preferences, all of the recommendations shown were for the dominated choice options only (i.e., Option A' and Option B'). We thus measured susceptibility to algorithm overdependence by observing the frequency with which participants selected these inferior recommendations across conditions. Finally, participants evaluated the recommender system on trust, knowledge, and expertise.

### Results

**Algorithm overdependence.** We submitted choice of dominated options to a logistic regression on the intervention condition, web environment condition, and their interaction. The main effects of the intervention ( $\chi^2 < 1$ ) and web environment ( $\chi^2(1, 200) = 1.82$ ) were not significant. Thus, the consumer education intervention did not lead participants to avoid



**Figure 5.** Percentage of participants choosing dominated, nondominated, and other options in Experiment 4.

inferior recommendations in general; similarly, interacting with a suspicious website also did not lead participants to avoid inferior recommendations in general. Instead, we observed a significant interaction effect ( $\chi^2(1, 200) = 5.17, p = .023$ ). The consumer education treatment insulated users against algorithm overdependence when participants encountered a suspicious website by reducing choice of the inferior options relative to the no-intervention control condition (18% vs. 35%;  $\chi^2(1, 100) = 3.82, p = .051$ ). In fact, the intervention had the positive effect of significantly increasing selection of superior choice options relative to the no-intervention control (63% vs. 49%;  $\chi^2(1, 100) = 3.82, p = .037$ ). When participants instead encountered a standard website in which there was comparatively less basis for suspicion, the intervention had no effect on participants' adoption of inferior recommendations ( $\chi^2(1, 100) = 1.63, p > .20$ ). Thus, the intervention served to effectively reduce risks related to algorithm overdependence, but only when individuals faced relatively suspicious web environments (see Figure 5).

**Perceived expertise.** Consistent with the idea that consumers are more susceptible to adopting inferior recommendations when they perceive the algorithm as holding comparatively greater expertise, we found that participants who encountered the suspicious website evaluated it as having significantly lower levels of expertise than the standard website ( $M = 3.83$  vs.  $4.38, t(198) = 2.70, p < .01$ ). They expressed similar sentiments regarding the suspicious (vs. standard) website when evaluating the extent to which the recommender system knew what was best for them ( $M = 3.50$  vs.  $M = 4.04; t(198) = 2.44, p = .016$ ).

### Discussion

The findings of this study provide insights regarding the efficacy of a low-cost and scalable consumer education intervention approach that addresses the contextual and complexity factors involved in driving algorithm overdependence.

Specifically, these results demonstrate that a broad-based consumer education approach may be effective in reducing algorithm overdependence, but only when consumers have some basis for developing skepticism of the recommender system. While the current consumer education intervention may mitigate algorithm overdependence in web contexts where consumers can easily infer potential vulnerabilities and uncertainties, additional intervention strategies may be required to minimize algorithm overdependence on websites that consumers perceive as being more reliable. Consequently, one way to minimize welfare loss could be to apply a two-pronged approach that targets both consumers and firms. For example, the incorporation of warnings and disclosures in online shopping portals may serve to alert consumers to the vulnerabilities and uncertainties they face, enabling individuals to more easily determine when they should avoid recommendations even when they are not experts in the product domain.

### General Discussion

As recommender systems are increasingly implemented within a range of consumer settings, the risks they pose to consumer welfare have growing importance. Indeed, recommender systems automate many elements of consumers' search and choice processes to ease decision making, but they can simultaneously expose users to greater vulnerability by facilitating the choice of inferior products and services.

When making decisions with the assistance of an algorithmic aid, consumers often avoid the superior recommendations in favor of their own intuitions, as prior research has established. Consequently, consumers can incur welfare loss by making suboptimal decisions due to algorithm aversion. However, as we aim to emphasize in this work, policy approaches that aim to assist consumers by singularly increasing reliance on algorithmic aids are misguided. The current research illustrates that in many situations consumers in fact display overdependence on algorithmic recommendations by adopting

inferior recommendations even when they reduce welfare. We suggest that this pair of issues associated with the use of algorithmic assistants can be likened to type I and type II errors. That is, policy makers must appreciate and balance type I false positive problems (i.e., algorithm overdependence) against type II false negative problems (i.e., algorithm aversion) when monitoring the rapid application of recommender systems to digital marketplaces.

### *Theoretical Implications*

The current work offers evidence that documents the prevalence and robustness of algorithm overdependence, organizes and evaluates the factors that give rise to this behavior, and presents a concrete intervention approach that can mitigate welfare loss associated with the use of biased recommender systems. These findings fit within a broader conceptualization of the surrendering-to-technology framework (Walker 2016), expanding the domain of digital interactions on which such phenomena have previously been analyzed, from privacy vulnerabilities surrounding information sharing to consumer welfare vulnerabilities linked to purchase and consumption decisions.

In building on the surrendering-to-technology framework, we characterize an additional complexity factor (i.e., limits of understanding) that can contribute to consumer susceptibilities to harm in an even wider range of digital interactions. Inaccurate mental models regarding how complex products and algorithms work not only spur algorithm overdependence but may also play a role in other problems linked to sharing and consuming information. For example, even when users are applying full attention and effort, they may hold inaccurate mental models around how digital social media applications work (e.g., believing that when content automatically disappears or is deleted from a profile, it is no longer accessible to the company), which can lead to misplaced faith in the technology. In addition, when people consume information online, a limited understanding of complex topics such as vaccinations or genetically modified foods can contribute to consumer susceptibility to misinformation, even when individuals devote full attention and effort toward evaluating news articles.

The current studies examined online shopping settings in which consumers commonly interact with recommendation systems (such as when buying headphones or portable chargers<sup>1</sup>). In settings that call for greater levels of involvement (such as important medical treatment or financial investment decisions), consumers may generally devote greater attention and effort to their decisions. However, because complexity and contextual factors may continue to align, it is not clear that decisions demanding greater involvement necessarily lessen

concerns of algorithm overdependence. If typical consumers have a limited understanding of highly specialized medical and financial domains and perceive algorithmic aids as possessing a high level of certainty (particularly if provided by a trusted hospital or banking institution), susceptibility to algorithm overdependence may indeed persist in these environments. In fact, as our data (Experiments 2a and 2b) illustrate, consumers continue to display algorithm overdependence even when attention and effort are not limiting factors.

In distinguishing between problems of false negatives (type II issues) and false positives (type I issues), we hope to highlight the dual nature of the issues associated with assisting consumers in making better decisions with algorithmic aids. While much of the existing literature has focused on addressing problems of algorithm aversion, researchers must consider both issues when assessing the efficacy of a new algorithm or an intervention that aims to improve consumer decisions. For instance, rather than aiming to maximize the adoption of algorithmic sources of information alone, it would be more appropriate to maximize a composite score that accounts for both type I and type II errors. Signal detection theory offers various measures that have been applied to assess performance of classification algorithms, and similar metrics could be appropriated to assess consumers' ability to correctly "classify" and respond to both superior and inferior recommendations (e.g., an area under the curve [AUC] measure; Hernández-Orallo, Flach, and Ferri 2012).

### *Policy Implications*

It has been well-documented that the recommendations one encounters online can frequently be biased. They often favor mainstream best sellers (Fleder and Hosanagar 2009), push more expensive items with higher margins (Angwin and Mattu 2016; Hannak et al. 2014; Mikians et al. 2012), and may even play a role in perpetuating social inequalities (Datta, Tschantz, and Datta 2015; Sweeney 2013). Our data indicate that people are susceptible to these inferior recommendations even when they are presented in plain sight, as we primarily operationalized inferior options as those that verifiably cost more or offered fewer benefits than other options in the choice set. However, in the marketplace, inferior recommendations can often be obscured. Without full information on product attributes and consumers' true utility functions, both watchdogs and consumers themselves cannot accurately assess whether adopting the recommendations they are presented with may lead to a welfare loss. Each of these elements may indeed be obscured as exact attribute information is often uncharacterized (e.g., quality or other unobservables) and because consumers frequently do not know or may not even hold a clear set of true preferences (e.g., when making decisions that involve complex trade-offs over multiple attributes). Thus, we anticipate that the risks of exploitation associated with algorithm overdependence apply to a much broader set of situations in which inferior recommendations are not presented in plain sight but are instead more subtle in nature.

<sup>1</sup> Consumers (n = 530 MTurk participants) also selected inferior granola bars more often when shown a dominated (13%) rather than nondominated (3.4%;  $\chi^2(1, 530) = 16.4, p < .001$ ) recommendation, adopting an experimental design from Fitzsimons and Lehman's (2004) Experiment 1.

Moreover, algorithm overdependence has the potential to spread throughout networks of users, snowballing through dynamic amplification. Because collaborative filtering recommendation algorithms often connect users to one other, when one user adopts an inferior recommendation, the item becomes more likely to be presented to similar users. Collaborative filtering algorithms typically function by matching a user to others who share closely overlapping past purchase histories (indicating similar preferences) and then recommending items purchased by similar others that have not been purchased by the target user. Thus, if customer A uses her favorite online retailing platform to purchase a new laptop and selects an overpriced alternative presented by the recommender system, when a similar user, customer B, searches for a new laptop himself, he will be more likely to see the inferior alternative within his own recommendations. This process perpetuates and can subsequently affect all similar users. Thus, inferior recommendations have the potential to propagate throughout the network and can consequently present risks to consumer welfare for a very wide range of users.

While few policies have currently been implemented to safeguard consumers from faulty or malicious algorithms, new regulations imposed by the General Data Protection Regulation in the European Union have begun to offer consumer protections that ensure one's rights to nondiscrimination and explanation. Such efforts indeed help pressure firms to avoid overtly exploiting consumers, though we remain concerned about the potential for more subtle ways in which consumers may be vulnerable to algorithm overdependence (e.g., when they hold less information than firms about hard-to-assess product attributes like quality). In the United States, the Federal Trade Commission (FTC) has provided guidelines regarding product endorsements by bloggers, online reviewers, and advertising agencies that require parties to both (1) substantiate claims and (2) declare any endorsement compensation. Widening each of these FTC guidelines to apply to automated forms of advertisement could protect consumer interests when interacting with recommender systems.

Extending existing FTC regulations around advertising substantiation and disclosure practices (FTC 2013) offers one avenue for the oversight of recommendation systems. Because recommendations make implicit claims (i.e., promising a better product or experience for consumers), such claims may fit under the purview of existing advertising-related policy. For example, just as advertising claims such as "removes tough stains" requires substantiation, regulators may also require substantiation for recommendation claims. To validate the efficacy of a recommender system, online retailers may commission independent studies similar to methods in which companies validate advertising claims. If consumers are informed while shopping online that 90% of consumers were more satisfied with their purchases when using Retailer X's recommender system (vs. 50% at Retailer Y), such information can allow consumers to better manage uncertainties by correspondingly allocating cognitive resources and seeking out other sources of information as necessary.

Because few rules currently govern recommendation systems directly, consumers often are not able to ascertain what recommendation labels truly mean. When labels such as "recommended" or "Amazon's choice" are applied to products, providing users with access to information around how and why such items were selected could also enable consumers to more easily determine whether they should adopt or avoid recommendations. Requiring disclosure of the reasons behind why the recommendations are believed to be beneficial to users (e.g., similarity to other products purchased, fit with known aspects of the user's preferences) may allow consumers to better assess the relevance of recommendations when making decisions. For example, consider selecting a new health insurance plan. While consumers may not have a full understanding of the types of services covered by the plan, they are likely to be aware of their basic needs. When a recommendation informs consumers that the plan is suggested "based on their interest in access to more eyecare providers," users are able to verify whether the criteria behind the recommendation are in line with their interests, dealing with these limits of understanding. Some firms already offer a degree of transparency when providing recommendations (e.g., Netflix explaining the reasoning behind movie suggestions), and recent methodological work facilitates the ability to automatically share reasoning behind recommendations with users (Datta, Sen, and Zick 2016).

Because recommender systems can be viewed as a form of automated endorsements, extensions to FTC endorsement guidelines (FTC 2017) could offer an additional avenue for providing oversight. Currently, users are typically not informed when automated recommendations suggest products based in part on profitability or other firm-related strategic interests. By extending endorsement guidelines to prohibit the use of certain types of information (i.e., profit margins or other aspects that do not align with consumer interests) being used when making automated recommendations, the frequency with which consumers encounter vulnerabilities when interacting with recommender systems could be greatly reduced. Regulators could also offer badges or certifications (similar to Energy Star labels) to companies that refrain from incorporating strategic or profit-related variables when delivering recommendations to consumers. These badges would enable consumers to understand when they ought to be more or less suspicious of the recommendations provided.

Common endorsement guidelines require some form of disclosure when providing paid endorsements through advertisements on blogs, YouTube, and social media sites. Similar disclosure requirements that inform consumers when recommendations are being delivered for reasons that do not align with consumer interests would allow users to better identify situations in which adopting recommendations may incur welfare loss. For example, required warnings informing online shoppers that "Recommended products are not necessarily better value. Use at your own risk." may alert people to the uncertainty associated with the recommendations and allow them to

identify situations in which they might need to seek out independent sources of information. These disclosures may be particularly valuable to consumers when product recommendations are biased by a firm's strategic objectives, such as the promotion of higher-margin items. However, a trickier challenge for policy makers may be to protect consumers from recommender systems that are not overtly malicious, yet still have the potential to provide poor quality recommendations (e.g., due to poor estimation of preferences or poor implementation of recommendation algorithms).

### Concluding Remarks

Because algorithmic decision aids are increasingly being applied within the marketplace to automate consumer search and choice processes, they present potentially large risks that must be understood and addressed. Effective recommender systems should be designed to leverage the strengths and counteract the inherent weaknesses of their users. As recommender systems continue to develop, we are hopeful that, with a deeper appreciation for the risks that they entail, these decision support tools will indeed aid and improve consumer decisions.

## Appendix

**Table A1.** Summary of the Choice Options in Experiments 1 and 4.

	Option A	Option A'	Option B	Option B'	Option C	Option D
Price	\$14.99	\$14.99	\$24.99	\$30.99	\$39.99	\$44.99
No. of charge cycles	4	3.5	6.5	6.5	7	10
Size	Medium	Medium	Small	Small	Very Small	Very Small
Amazon review	4.2/5	4.2/5	4.7/5	4.7/5	3.9/5	4.6/5
Charging speed	Medium	Medium	High	High	High	Very High
No. of USB slots	2	2	2	2	1	4
Warranty	24 months	24 months	18 months	18 months	18 months	12 months
Plug type	USB	USB	USB	USB	USB	USB

**Table A2.** Summary of the Choice Options in Experiment 2a.

Attribute	Option A	Option A'	Option B	Option C	Option D
Sound quality	4.9	4.6	4.2	4.6	3.3
Range	4.6	4.6	3.3	5.0	3.1
Noise cancellation	4.3	3.8	4.7	2.9	4.1
Comfort	4.6	4.6	4.1	4.2	4.9
Battery life	4.7	4.7	4.5	4.6	4.6
Style	4.5	4.5	5.0	3.8	3.9

**Table A3.** Summary of the Choice Options in Experiment 2b.

	Option A	Option A'	Option B	Option B'	Option C	Option D
Price	\$14.99	\$14.99	\$24.99	\$30.99	\$39.99	\$44.99
No. of charge cycles	4	3.5	6.5	6.5	7	10
Size	Medium	Medium	Small	Small	Very Small	Very Small
Amazon review	4.2/5	4.2/5	4.7/5	4.7/5	3.9/5	4.6/5
Charging speed	Medium	Medium	High	High	High	Very High
No. of USB slots	2	2	2	2	1	4
Warranty	24 months	24 months	18 months	18 months	18 months	12 months
Ruggedness	High	High	High	High	High	Low
Weight	Medium	Medium	Light	Light	Heavy	Medium
Plug type	USB	USB	USB	USB	USB	USB

Notes: Only italicized attributes were displayed in the low-decision-difficulty conditions.



## Editorial Team

Kristen Walker, George Milne, and Bruce Weinberg served as guest editors for this article.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## References

- Adomavicius, Gediminas and Alexander Tuzhilin (2005), "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering*, 17 (6), 734–49.
- Alba, Joseph W. and J. Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (4), 411–54.
- Angwin, Julia and Surya Mattu (2016), "Amazon Says It Puts Customers First. But Its Pricing Algorithm Doesn't," *ProPublica* (September 20), <https://www.propublica.org/article/amazon-says-it-puts-customers-first-but-its-pricing-algorithm-doesnt>.
- Ansari, Asim, Skander Essegai, and Rajeev Kohli (2000), "Internet Recommendation Systems," *Journal of Marketing Research*, 37 (3), 363–75.
- Banker, Sachin, Sarah E. Ainsworth, Roy F. Baumeister, Dan Ariely, and Kathleen D. Vohs (2017), "The Sticky Anchor Hypothesis: Ego Depletion Increases Susceptibility to Situational Cues," *Journal of Behavioral Decision Making*, 30 (5), 1027–40.
- Baumeister, Roy F., Dianne M. Tice, and Kathleen D. Vohs (2018), "The Strength Model of Self-Regulation: Conclusions From the Second Decade of Willpower Research," *Perspectives on Psychological Science*, 13 (2), 141–45.
- Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (3), 187–217.
- Bodapati, Anand V. (2008), "Recommendation Systems with Purchase Data," *Journal of Marketing Research*, 45 (1), 77–93.
- Burghardt, Keith, Emanuel F. Alsina, Michelle Girvan, William Rand, and Kristina Lerman (2017), "The Myopia of Crowds: Cognitive Load and Collective Evaluation of Answers on Stack Exchange," *PloS One*, 12 (3), e0173610.
- Chernev, Alexander (2003), "When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice," *Journal of Consumer Research*, 30 (2), 170–83.
- Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman (2015), "Choice Overload: A Conceptual Review and Meta-Analysis," *Journal of Consumer Psychology*, 25 (2), 333–58.
- Chung, Jaihak and Vithala R. Rao (2012), "A General Consumer Preference Model for Experience Products: Application to Internet Recommendation Services," *Journal of Marketing Research*, 49 (3), 289–305.
- Coupey, Eloise, Julie R. Irwin, and John W. Payne (1998), "Product Category Familiarity and Preference Construction," *Journal of Consumer Research*, 24 (4), 459–68.
- Datta, Amit, Michael Carl Tschantz, and Anupam Datta (2015), "Automated Experiments On Ad Privacy Settings," *Proceedings on Privacy Enhancing Technologies*, 2015 (1), 92–112.
- Datta, Anupam, Shayak Sen, and Yair Zick (2016), "Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems," in *2016 IEEE Symposium on Security and Privacy (SP)*. New York, IEEE, 598–617.
- Dellaert, Benedict G.C. and Stefan Stremersch (2005), "Marketing Mass-Customized Products: Striking a Balance Between Utility and Complexity," *Journal of Marketing Research*, 42 (2), 219–27.
- Diab, Dalia L., Shuang-Yueh Pui, Maya Yankelevich, and Scott Highhouse (2011), "Lay Perceptions of Selection Decision Aids in US and Non-US Samples," *International Journal of Selection and Assessment*, 19 (2), 209–16.
- Diehl, Kristin (2005), "When Two Rights Make a Wrong: Searching Too Much in Ordered Environments," *Journal of Marketing Research*, 42 (3), 313–22.
- Diehl, Kristin and Cait Poyner (2010), "Great Expectations?! Assortment Size, Expectations, and Satisfaction," *Journal of Marketing Research*, 47 (2), 312–22.
- Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey (2015), "Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err," *Journal of Experimental Psychology: General*, 144 (1), 114–26.
- Eastwood, Joseph, Brent Snook, and Kirk Luther (2012), "What People Want from Their Professionals: Attitudes Toward Decision-Making Strategies," *Journal of Behavioral Decision Making*, 25 (5), 458–68.
- Elliott, Justin (2019), "TurboTax Deliberately Hid Its Free File Page From Search Engines," *ProPublica* (April 26), <https://www.propublica.org/article/turbotax-deliberately-hides-its-free-file-page-from-search-engines>.
- Elliott, Justin and Lucas Waldron (2019), "Here's How TurboTax Just Tricked You into Paying to File Your Taxes," *ProPublica* (April 22), <https://www.propublica.org/article/turbotax-just-tricked-you-into-paying-to-file-your-taxes>.
- Federal Trade Commission (2013), "*.com Disclosures: How to Make Effective Disclosures in Digital Advertising*," research report (March), <https://www.ftc.gov/sites/default/files/attachments/press-releases/ftc-staff-revises-online-advertising-disclosure-guidelines/130312dotcomdisclosures.pdf>.
- Federal Trade Commission (2017), "*The FTC's Endorsement Guides: What People Are Asking*," (September), <https://www.ftc.gov/tips-advice/business-center/guidance/ftcs-endorsement-guides-what-people-are-asking>.
- Fildes, Robert and Paul Goodwin (2007), "Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting," *Interfaces*, 37 (6), 570–76.
- Fitzsimons, Gavan J. and Donald R. Lehmann (2004), "Reactance to Recommendations: When Unsolicited Advice Yields Contrary Responses," *Marketing Science*, 23 (1), 82–94.

- Fleder, Daniel and Kartik Hosanagar (2009), "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity," *Management Science*, 55 (5), 697–712.
- Ghose, Anindya, Panagiotis G. Ipeirotis, and Beibei Li (2012), "Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content," *Marketing Science*, 31 (3), 493–520.
- Hannak, Aniko, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson (2014), "Measuring Price Discrimination and Steering on E-Commerce Web Sites," in *Proceedings of the 2014 Conference on Internet Measurement Conference, IMC '14*. New York: Association for Computing Machinery, 305–18.
- Häubl, Gerald and Valerie Trifts (2000), "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids," *Marketing Science*, 19 (1), 4–21.
- Hernández-Orallo, José, Peter Flach, and Cèsar Ferri (2012), "A Unified View of Performance Metrics: Translating Threshold Choice into Expected Classification Loss," *Journal of Machine Learning Research*, 13 (1), 2813–69.
- Hoch, Stephen J., Eric T. Bradlow, and Brian Wansink (1999), "The Variety of an Assortment," *Marketing Science*, 18 (4), 527–46.
- Huber, Joel, John W. Payne, and Christopher Puto (1982), "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9 (1), 90–98.
- Konstan, Joseph A. and John Riedl (2012), "Recommender Systems: From Algorithms to User Experience," *User Modeling and User-Adapted Interaction*, 22 (1/2), 101–23.
- Kurzban, Robert, Angela Duckworth, Joseph W. Kable, and Justus Myers (2013), "An Opportunity Cost Model of Subjective Effort and Task Performance," *Behavioral and Brain Sciences*, 36 (6), 661–79.
- Louro, Maria J., Rik Pieters, and Marcel Zeelenberg (2007), "Dynamics of Multiple-Goal Pursuit," *Journal of Personality and Social Psychology*, 93 (2), 174–93.
- Lu, Shasha, Li Xiao, and Min Ding (2016), "A Video-Based Automated Recommender (VAR) System for Garments," *Marketing Science*, 35 (3), 484–510.
- Mikians, Jakub, László Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris (2012), "Detecting Price and Search Discrimination on the Internet," in *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*. New York: Association for Computing Machinery, 79–84.
- Mobasher, Bamshad, Robin Burke, Runa Bhaumik, and Chad Williams (2007), "Toward Trustworthy Recommender Systems: An Analysis of Attack Models and Algorithm Robustness," *ACM Transactions on Internet Technology*, 7 (4), 23.
- Mullainathan, Sendhil and Ziad Obermeyer (2017), "Does Machine Learning Automate Moral Hazard and Error?" *American Economic Review*, 107 (5), 476–80.
- Nordgren, Loran F. and A.P. Dijksterhuis (2008), "The Devil Is in the Deliberation: Thinking Too Much Reduces Preference Consistency," *Journal of Consumer Research*, 36 (1), 39–46.
- Paul, Sharoda A., Lichan Hong, and Ed H. Chi (2012), "Who Is Authoritative? Understanding Reputation Mechanisms in Quora," presented at the Collective Intelligence Conference, arXiv:1204.3724.
- Resnick, Paul and Hal R. Varian (1997), "Recommender Systems," *Communications of the ACM*, 40 (3), 56–58.
- Sanders, Nada R. and Karl B. Manrodt (2003), "Forecasting Software in Practice: Use, Satisfaction, and Performance," *Interfaces*, 33 (5), 90–93.
- Shaffer, Victoria A., C. Adam Probst, Edgar C. Merkle, Hal R. Arkes, and Mitchell A. Medow (2013), "Why Do Patients Derogate Physicians Who Use a Computer-Based Diagnostic Support System?" *Medical Decision Making*, 33 (1), 108–18.
- Simons, Daniel J. and Christopher F. Chabris (1999), "Gorillas in Our Midst: Sustained Inattention Blindness for Dynamic Events," *Perception*, 28 (9), 1059–74.
- Swaminathan, Vanitha (2003), "The Impact of Recommendation Agents on Consumer Evaluation and Choice: The Moderating Role of Category Risk, Product Complexity, and Consumer Knowledge," *Journal of Consumer Psychology*, 13 (1), 93–101.
- Sweeney, Latanya (2013), "Discrimination in Online Ad Delivery," *Queue*, 11 (3), 10.
- Tsai, Claire I. and Ann L. McGill (2011), "No Pain, No Gain? How Fluency and Construal Level Affect Consumer Confidence," *Journal of Consumer Research*, 37 (5), 807–21.
- Vrieze, Scott I. and William M. Grove (2009), "Survey on the Use of Clinical and Mechanical Prediction Methods in Clinical Psychology," *Professional Psychology: Research and Practice*, 40 (5), 525–31.
- Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi (2017), "Why a Right to Explanation of Automated Decision-Making Does not Exist in the General Data Protection Regulation," *International Data Privacy Law*, 7 (2), 76–99.
- Walker, Kristen L. (2016), "Surrendering Information Through the Looking Glass: Transparency, Trust, and Protection," *Journal of Public Policy & Marketing*, 35 (1), 144–58.
- West, Patricia M., Dan Ariely, Steve Bellman, Eric Bradlow, Joel Huber, Eric Johnson, et al. (1999), "Agents to the Rescue?" *Marketing Letters*, 10 (3), 285–300.
- Wilson, Timothy D. and Jonathan W. Schooler (1991), "Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions," *Journal of Personality and Social Psychology*, 60 (2), 181.
- Xiao, Bo and Izak Benbasat (2007), "E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact," *MIS Quarterly*, 31 (1), 137–209.
- Ying, Yuanping, Fred Feinberg, and Michel Wedel (2006), "Leveraging Missing Ratings to Improve Online Recommendation Systems," *Journal of Marketing Research*, 43 (3), 355–65.