Received: 13 December 2018 | Revised: 2 September 2019 | Accepted: 2 September 2019
DOI: 10.1002/bdm.2155

RESEARCH ARTICLE

A systematic review of algorithm aversion in augmented decision making

Jason W. Burton[1] | **Mari-Klara Stein**[2] | **Tina Blegind Jensen**[2]

[1]Department of Psychological Sciences, Birkbeck, University of London, London, UK
[2]Department of Digitalization, Copenhagen Business School, Frederiksberg, Denmark

Correspondence
Jason W. Burton, Department of Psychological Sciences, Birkbeck, University of London, Male Street, London, WC1E 7HX, UK.
Email: jasonwilliamburton@gmail.com

Abstract

Despite abundant literature theorizing societal implications of algorithmic decision making, relatively little is known about the conditions that lead to the acceptance or rejection of algorithmically generated insights by individual users of decision aids. More specifically, recent findings of algorithm aversion—the reluctance of human forecasters to use superior but imperfect algorithms—raise questions about whether joint human-algorithm decision making is feasible in practice. In this paper, we systematically review the topic of algorithm aversion as it appears in 61 peer-reviewed articles between 1950 and 2018 and follow its conceptual trail across disciplines. We categorize and report on the proposed causes and solutions of algorithm aversion in five themes: expectations and expertise, decision autonomy, incentivization, cognitive compatibility, and divergent rationalities. Although each of the presented themes addresses distinct features of an algorithmic decision aid, human users of the decision aid, and/or the decision making environment, apparent interdependencies are highlighted. We conclude that resolving algorithm aversion requires an updated research program with an emphasis on theory integration. We provide a number of empirical questions that can be immediately carried forth by the behavioral decision making community.

KEYWORDS
algorithm aversion, augmented decision making, human-algorithm interaction, systematic review

1 | INTRODUCTION

Algorithms have long been touted as a cognitive cure for the limitations of human judgement and decision making (e.g., Dawes, 1979; Meehl, 1954), and recently, we have witnessed an increasing proportion of both high-stakes and mundane decisions being augmented by algorithmic aids. Yet, in spite of the growing ubiquity of algorithmically augmented decision making, recent research demonstrates the persistence of algorithm aversion, which is the reluctance of human decision makers to use superior but imperfect algorithms (Dietvorst, Simmons, & Massey, 2015). Although explanations for algorithm aversion have indeed been proposed in the past (e.g., Dawes, 1979; Einorn, 1986; Grove & Meehl, 1996; Highhouse, 2008b), the inability to effectively combine human and nonhuman (i.e., algorithmic, statistical, machine, etc.) decision making remains one of the most prominent and perplexing hurdles for the behavioral decision making community.

Amid a growing body of literature, this paper sets out to systematically review the existing research that addresses the central question: Why do people misuse (i.e., under- or over-utilize) algorithmically generated insights in augmented decision making? This review begins with the phenomenon of algorithm aversion and follows its conceptual trail to synthesize a clear account of the cognitive, behavioral, and organizational issues that lead experts and laypeople to inappropriately integrate algorithmic judgement into their own. In addition, we offer practical suggestions that can be mobilized in practice or carried forward in a research agenda within the behavioral decision making community.

In the present work, it must also be noted that various terms are used interchangeably. Although nuanced distinctions are made elsewhere, the focus of this article is on the interaction between human and nonhuman agents in decision making. Hence, the notion of algorithmic decision making is considered an umbrella term for related paradigms like augmented decision making, decision aids, decision support systems, expert systems, decision formulas, computerized aids, and diagnostic aids. Likewise, variations of decision making, judgement, forecasting, and prediction are considered equivalent for the purpose of this review. Beyond this terminology, it is also necessary to distinguish what we consider to be successful from unsuccessful algorithmic decision making. In this paper, we define successful algorithmic decision making as an augmented decision making process where algorithmic insights are utilized accurately and, most importantly, discriminately. This means that a successful human-algorithm augmentation is one in which the human user is able to accurately discern both when and when not to integrate the algorithm's judgement into his or her own decision making. Neither blind neglect nor blind acceptance of algorithmic insights can be considered successful in this view because such decisions signal the absence or failure of interaction between the human and algorithm. There are of course cases where augmentation is inappropriate, where the algorithm should completely automate the decision or where the human user wants to delegate his or her decision autonomy, but the emphasis in this paper is on cases in which agency is shared between human and algorithm.

The remainder of this article proceeds as follows. First, we explain the method employed for searching and coding the relevant literature. Then, we categorize and present the results from existing research in five themes—expectations and expertise, decision control, incentivization, cognitive compatibility, and divergent rationalities—as they relate to causes and solutions for algorithm aversion. Finally, we open a discussion to highlight the connections between themes, draw in relevant theories that are not explicitly addressed in the reviewed literature, and suggest avenues for future research.

2 | METHOD

A systematic review of existing literature was performed to gather and analyze published, peer-reviewed research articles. By way of eight academic search engines and databases, 1,363 abstracts were screened, which resulted in a final set of 61 articles[1] from 36 academic journals deemed eligible for analysis (Appendix, Table 1). The reason behind this drastic reduction from the number of retrieved articles to relevant articles is itself noteworthy, but it can largely be explained by the fact that a majority of researchers' past efforts have been focused on comparing rather than conjoining human and nonhuman decision making (for a review see Kleinmuntz, 1990). It is also evident that the definition of algorithm has evolved with advances in computing and artificial intelligence. For the purpose of this review, a fundamental definition of algorithm was adapted from the Merriam–Webster Dictionary as a mathematical, step-by-step procedure or formula for computation. As such, the literature search was not necessarily limited to digital algorithmic decision aids and also considered the use of basic decision rules or simple paper-and-pencil decision algorithms in collaboration with intuitive judgement under our broad label of nonhuman.

2.1 | Databases

Due to the interdisciplinary nature of augmented, algorithmic decision making, large databases were sought out so as not to narrow the scope of the review unnecessarily. Thus, the literature search was conducted with JSTOR, Wiley Online Library, ScienceDirect, Taylor & Francis Online, SAGE Journals, ACM Digital Library, IEEE Xplore Digital Library, and EBSCOhost cross-database searching. The ACM Digital Library and IEEE Xplore Digital Library were specifically included to account for technical perspectives on algorithmic decision making and the development of modern computational decision aids. The PsycINFO, PsycARTICLES, Business Source Complete, Academic Search Elite, and EconLit databases were included within the EBSCOhost cross-database search. Altogether, these selected databases were deemed sufficient for the purpose of this research due to the incorporation of traditional disciplines (e.g., psychology, management, and human factors) and nontraditional, hybrid disciplines (e.g., behavioral economics, social neuroscience, and information systems). Moreover, the high frequency of repeated articles across databases suggested an adequate level of search rigor.

2.2 | Search terms

Inspired by Dietvorst et al.'s (2015) conceptualization, the selection of search terms began with algorithm aversion. Although this exact terminology appears only sparsely in existing literature, it most accurately encapsulates the phenomenon up for review; that is, the rejection versus acceptance of algorithmically generated insights.

Given the novelty and specificity of algorithm aversion, searches were also performed with the term algorithmic decision making to retrieve articles from the broader research setting. Despite retrieving many irrelevant articles, algorithmic decision making was found to be a shared concept across disciplines.

Upon reading Dietvorst et al. (2015), it was evident that the notion of trust plays a leading role in shaping perceptions of algorithmically generated insights. However, trust is a conversational term with numerous connotations in social contexts and is often interchanged with confidence. Therefore, the review of literature required a narrower, more concrete construct to focus on within human-algorithm interaction in decision making. To satisfy this need, searches were performed with the term “advice utilization” as it is noted that this term is used in research as a proxy behavior for trust; an objective behavioral measure (Prahl & Van Swol, 2017). Because this search

[1]Within the 61 selected references, there is a string of eleven commentary articles from Industrial and Organizational Psychology that are in response to Highhouse's (2008b) target article, and a string of 2 commentary articles from MIS Quarterly in response to Rao et al. (1992).

retrieved a manageable number of articles, no further search qualification was needed.

Delving into the literature on algorithmic decision making and advice utilization provided an in-depth body of evidence for statistical judgement and prediction, but these searches largely overlooked investigations of expert intuition—the main cognitive force contending with rational calculation for decision influence (Patterson, 2017). However, intuition needed to be contextualized in order to narrow down search results. Therefore, searches were performed with intuition and decision aids. As opposed to qualifying intuition with decision making, which retrieved innumerable studies of naturalistic decision making and the like, or with algorithmic, which failed to cast a wide enough net, the additional qualifier of decision aids proved effective in maintaining article relevance to the topic of interaction between human intuition and algorithmic calculation in decision making.

2.3 | Inclusion criteria

Inclusion criteria came into play at two stages, setting initial database search parameters and setting conceptual boundaries for abstract screening. Database searches were limited to peer-reviewed journal articles published between 1950 (the decade in which Meehl's, 1954 famous examination of clinical and statistical prediction was published) and 2018, written in English.

To screen abstracts, a hard line was drawn to only incorporate conceptually relevant research. Noting that this review is not interested in the overarching outcomes or ethical debates relating to algorithmic decision making, research that focused on algorithmic governance, fairness, opacity, societal consequences, and so forth was excluded from the core literature selection. To be included in the analysis, research must have addressed the interaction between human and nonhuman agents in decision making with specific attention given to conditions for the under- or over-utilization of algorithmically generated insights. Thus, papers that discussed only intuitive human decision making or only algorithmic decision aids were excluded, whereas papers that discussed intuitive human decision making and algorithmic decision aids were included.

2.4 | Analysis

Analysis began with organizing the body of literature on the basis of target variables. Given that the focal phenomenon, algorithm aversion, arises in human-algorithm interaction, the articles addressed factors pertaining to the human decision maker, the algorithmic tool, and/or the decision making environment (Appendix, Table 1). This categorization, paired with an assessment of methodologies employed (Appendix, Table 2), uncovered basic weak points in the existing literature, such as the absence of robust empirical studies testing alternative algorithmic aid designs.

Coding for themes was performed by labeling underlying theories deployed in each article, as well as considering the discipline and journal it was published in. This strategy allowed for the categorization based on focal variables to be corroborated. For example, Christin (2017) was categorized as an article focusing on the environmental variables influencing human-algorithm interaction, and by noting its use of sociological theory [e.g., Bourdieu's, 1993 theory of fields] and publication in Big Data & Society, the categorization was reiterated with increased confidence. Next, proposed causes for algorithm aversion were drawn out from each article by isolating the variable that, when manipulated, resulted in a main effect on algorithm aversion and the utilization of algorithmic advice. Accompanying proposals for overcoming algorithm aversion were also extracted, however many of these remain untested hypotheses. From this mapping of problems and solutions, an initial set of four themes was distinguished: expectations and expertise, decision autonomy, incentivization, and cognitive compatibility. The fifth theme of divergent rationalities arose on the grounds that algorithmic decision making is not simply human-algorithm interaction, but rather fundamentally rooted in decision making as a cognitive function. From this conclusion, an additional analysis was performed to evaluate where the literature stood in relation to judgement and decision making (JDM) research. To do so, articles were marked on the basis of their supported JDM theories of rationality and categorized on their propensity to favor one school of thought or another; in this case, either the heuristics-and-biases or the fast-and-frugal program (Appendix, Table 3)[2]. For example, Highhouse (2008b) critiques the relevance of Gigerenzer's work and cites Kahneman in support of his position. Such critique illustrates a clear stance on the side of the heuristics-and-biases program that is championed by scholars like Tversky and Kahneman, so said article was categorized accordingly. In many of the analyzed studies, the JDM orientation formed the basis for the research as a whole.

3 | RESULTS

The reviewed literature is characterized by five themes, each of which relate to the key variables in a unique way. Expectations and expertise primarily concerns the human decision maker, decision autonomy centers on the design of the algorithmic aid, incentivization reports on the role of extrinsic incentives present in the decision making environment, cognitive compatibility involves the integration of decision processes between the human and algorithm (i.e., agent to agent), and divergent rationalities explains the issues that arise when the human decision maker and algorithm work toward different decision outcomes due to different interpretations of the environment (i.e., agent to environment). In this section, we present the cause of algorithm aversion (the problem) and how to overcome it (the solution) by synthesizing the message of past research and providing select examples in light of each theme (see Appendix, Table 4 for results summary).

[2]This polarization of JDM orientation to favor either the heuristics-and-biases or the fast-and-frugal program is a relatively recent phenomenon. Older articles tend to remain neutral on the issue or lack any explicit mentions of JDM/rationality theory. The coding sheds light on how the underlying theoretical perspectives have developed over time and influenced the domain of algorithmic decision making so it thus was deemed valuable.

3.1 | Expectations and expertise

3.1.1 | Problem: False expectations

A human decision maker rarely, if ever, confronts an algorithm with a blank slate. Prior to engaging in algorithmic decision making, a human decision maker will have developed expectations as to what an algorithm can do, what an algorithm should do, and how an algorithm functions. These expectations can be the product of firsthand experience with algorithmic aids, experience in the decision domain, or merely secondhand knowledge picked up from peers and media. What manifests from these pre-existing expectations is a paradigm in which human decision makers perceive and respond to advice generated by algorithms differently than advice generated by humans, even if the advice itself is identical. Various mechanisms underlying this difference in response are demonstrated throughout the literature, such as the tendency for humans to seek a social or parasocial relationship with the source of advice (Alexander, Blinder, & Zak, 2018; Önkal, Goodwin, Thomson, Gonul, & Pollock, 2009; Prahl & Van Swol, 2017), the persistent belief that human error is random and repairable whereas algorithmic error is systematic (Dietvorst et al., 2015; Dietvorst, Simmons, & Massey, 2016; Highhouse, 2008b), experts' domain confidence leading to underutilization of seemingly unnecessary algorithmic aids (Arkes, Dawes, & Christensen, 1986; Ashton, Ashton, & Davis, 1994), or a lack of training preventing a human user from properly utilizing an algorithmic aid (Mackay & Elam, 1992; cf. Green & Hughes, 1986). Put simply, the expectations that a human user brings into a human-algorithm interaction influence the way in which he or she utilizes the algorithm.

Within this theme of findings, existing research distinguishes between the effect of specific experience with algorithmic decision aids and the effect of experience with the decision domain. Here, we find that experience with algorithmic decision aids is positively associated with the utilization of algorithmic judgements, whereas domain expertise is negatively associated with utilization (Montazemi, 1991; Whitecotton, 1996). Perhaps this is not so surprising. Take, for example, two individuals who are provided with an algorithmic decision aid and tasked with forecasting the performance of an economic marketplace. The first individual is a trained forecaster who regularly uses algorithms to craft statistically-minded forecasts, and the second individual is a well-established economist who possesses deep knowledge of market theory but no familiarity with algorithmic decision aids. The trained forecaster is likely to feel unconfident in his intuitive ability to interpret the market, is capable of utilizing the algorithm with ease, and is thus more likely to integrate algorithmic judgment into his or her own forecast. On the other hand, the economist, who has a high degree of domain experience, is likely to feel confident without the aid of the algorithm and perceive the effort needed to consult it as unnecessary. In many ways, this problem of expectations can also be linked back to early literature on individual differences (e.g., demographics, statistics/computing experience, profession, etc.) as these play a significant role in determining what information and sentiments an individual attaches to algorithmic decision making.

3.1.2 | Solution: Algorithmic literacy

If false expectations prevent the proper utilization of algorithmic aids, then the solution to algorithm aversion should involve the development of algorithmic literacy among human decision makers. That is, human decision makers need to be trained not only in their professional domain, but also on how to interact with algorithmic tools, how to interpret statistical outputs, and how to appreciate the utility of decision aids (Goodyear et al., 2016; Kuncel, 2008; Lodato, Highhouse, & Brooks, 2011; Sanders & Courtney, 1985; Whitecotton, 1996). Crucially, algorithmic literacy must include the teaching of core statistical concepts like error and uncertainty. For instance, to be algorithmically literate, a decision maker has to be able to tolerate error as inherent to any decision task (Arkes et al., 1986; but also see Einhorn, 1986). Although a good decision aid might be accurate 80% of the time, this success rate is often displayed explicitly, and the 20% chance of inaccuracy is made salient to the user. If this same user is historically accurate 40% of the time when making intuition-based decisions, he or she would undoubtedly benefit from consulting such a decision aid. But, it is likely that his or her success rate is not explicit and the chance of intuition-based error (60% in this case) is concealed. Under such circumstances, an algorithmically illiterate user might interpret the algorithm's 25% chance of making an erroneous judgement as high, when in fact it is far superior to his or her own chance of erring. Indeed, a program of education to overcome algorithm aversion by highlighting such decision making problems may serve to prevent algorithm aversion in the future. On the other hand, algorithmic literacy puts the duty of overcoming algorithm aversion solely on the human decision maker although neglecting variables in the decision-making environment and the design of algorithmic aids. Moreover, the actual impact that algorithmic literacy would have is likely limited because existing studies that do observe algorithm aversion often rely on participants drawn from university programs who are presumably quite algorithmically literate (e.g., Alexander et al., 2018; Dietvorst et al., 2015, 2016; Önkal et al., 2009). Altogether, it seems unlikely that an algorithmic literacy program can suffice as a standalone intervention for solving algorithm aversion.

3.2 | Decision autonomy

3.2.1 | Problem: Lack of decision control

For a human decision maker to act on an algorithm's judgement, he or she must feel in control and confident enough to place trust in it (Colarelli & Thompson, 2008; Scherer, Zikmund-Fisher, Witteman, & Fagerlin, 2015). This feeling of control can come from a real understanding of the algorithm's performance, but it can also come from adjustments to the algorithmic decision making process that have little or no bearing on the actual functioning of the algorithm (e.g., changing the interface of information presentation without changing the way the algorithm analyzes information). Muir (1987) points out that trust in a decision aid is calibrated according to predictability, dependability, technical competence, reciprocity, and morality (i.e., an understanding that the aid is decent and is there to help rather than deceive or usurp). Along similar lines, Scherer et al. (2015) demonstrate that human decision makers often expect deliberation, a slow and effortful consideration of evidence, in high-stakes scenarios despite empirical findings suggesting that deliberation does not necessarily equate to better decision making. Regarding algorithmic decision aids, studies such as these highlight the need for affording real or perceived decision control to the human user in order to satisfy his or her psychological needs and self-interest (Colarelli & Thompson, 2008). In fact, this conclusion corraborates Dietvorst et al.'s (2015, 2016) finding that trust in an algorithm degrades quickly upon seeing it err, but that it can be equally quick to restore by allowing the human decision maker to modify the algorithm's judgment, even under constraints. Here, algorithm aversion appears to manifest itself in augmented decision-making systems that fail to address human users' psychological need for agency, autonomy, and control.

3.2.2 | Solution: Human-in-the-loop decision making

In large part, the recent findings of Dietvorst et al. (2016) are a rework of an old concept: human-in-the-loop decision making. Essentially, this entails an augmented decision making system in which the human user semisupervises the algorithm by having opportunities to intervene, provide input, and have the final say. As described in the reviewed literature, such decision making systems can take shape in a variety of ways, such as interactive support systems (Lim & O'Connor, 1996), human-automation systems (Patterson, 2017), engaged systems (Pagano et al., 2016), constructive dialog in expert systems (Eining, Jones, & Loebbecke, 1997), judgmental systems (Prahl & Van Swol, 2017), or procedural presentation in interfaces (Lamberti & Wallace, 1990). Nevertheless, Dietvorst et al. (2016) highlight an important new consideration: that people are relatively insensitive to the amount by which they can modify the imperfect algorithm's forecasts as long as they are able to incorporate their own input and participate in the ultimate decision (p. 1161). This suggests that even an illusion of autonomy will remedy algorithm aversion, and that augmented decision making systems need to include a kind of behavioral packaging or set of credibility factors that might be peripheral to decision performance, but central to overcoming algorithm aversion (Landsbergen, Coursey, Loveless, & Shangraw, 1997). Human-in-the-loop decision making can thus be viewed both as a principle of esthetic design and a principle of functionality. However, such added features are likely to impose costs by decreasing the speed at which users can extract the necessary decision information. Therefore, the viability of this solution is likely restricted to domains in which decision makers are given adequate time to collaborate with an algorithmic aid.

3.3 | Incentivization

3.3.1 | Problem: Lack of incentivization

Existing research points out that organizational and social structures favor the expert intuiter over a cold algorithmic decision maker and incentivize accordingly (e.g., Alexander et al., 2018; Brown, 2015; Eastwood, Snook, & Luther, 2012; Highhouse, 2008b; Klimoski & Jones, 2008; Kuncel, 2008; Önkal et al., 2009). Brown (2015) and Hafenbrädl, Waeger, Marewski, and Gigerenzer (2016) argue that augmented decision making requires extra motivation because it involves combining multiple judgements rather than the acceptance of a single calculation. This means that the successful implementation of algorithmic decision making requires motivating, or incentivizing, human decision makers to utilize algorithmic aids in order to balance the costs of effort with the benefits of decision performance (Christin, 2017). Throughout the literature, two types of incentives are prevalent: economic (e.g., monetary incentives for making accurate decisions) and social (e.g., abiding by social norms; maintaining reputation among peers and colleagues).

Let us first consider economic incentives. Given the robustness of research that demonstrates benefits of utilizing algorithms in decision making, one would expect human decision makers to readily incorporate algorithmic insights to make more accurate decisions, especially if they are offered monetary rewards for doing so. Paradoxically, however, economic incentives for decision performance have been shown to decrease the utilization of an algorithmic aid (Arkes et al., 1986). This finding highlights the nuances that come with incentivizing decision makers based on their performance. For example, if a decision maker is incentivized to make the best decision (relative to peers on a case-by-case competition basis) rather than a good decision (relative to one's own performance in the long run), then he or she would need to find a way to gain a unique advantage over competitors. If all competitors have access to the same or similar algorithmic aids, then the decision maker would put him or herself at a disadvantage by utilizing the algorithmic judgement because this would mean simply mirroring, rather than surpassing, the performance of other decision makers. As such, the backfire effect of economic incentives for decision performance, such as that used by Arkes et al. (1986), can in fact be considered to be the outcome of putatively rational behavior. However, conflicting results exist in which competitive cash rewards did not lead to algorithm aversion (e.g., Prahl & Van Swol, 2017), as well as experiments where algorithm aversion persists in the absence of competitive economic incentivization (e.g., Önkal et al., 2009). Reconciling this body of research is surely important, yet existing work only provides speculative explanations. For instance, Prahl and Van Swol (2017) suggest that their experiment included performance feedback and consistent message characteristics (i.e., only the source description was varied), whereas the set up in Önkal et al. (2009) did not include performance feedback and manipulated message characteristics across source types. Understanding how these factors interact with economic incentivization is an empirical question that deserves attention.

These contrasts lead us to the second source of extrinsic motivation for utilizing algorithmic aids: social incentives. It is widely accepted that decision making is inextricable from the social setting in which it takes place. There are often various stakeholders that each hold expectations and ideas of what constitutes a good decision, which may not necessarily include probabilistic accuracy. Because of this, decision makers are motivated to conform to social norms, be it a professional maintaining an aura of omniscience in front of clients/patients (Arkes, Shaffer, & Medow, 2007; Eastwood et al., 2012) or an employee seeking the support of management (Sanders & Courtney, 1985). In fact, information about others' algorithm utilization (i.e., social incentivization) is shown to have a greater influence than information about the algorithm itself (i.e., algorithmic literacy) on decision makers' engagement and performance with algorithmic aids (Alexander et al., 2018). Though this too is a nuanced finding. In comparing the influence of social information (i.e., information about social norms) with that of algorithm-related information, Alexander et al. (2018) used statistical information ("The algorithm is 75% accurate." p. 281) as their algorithm-related information. This assumes participants understand the need to tolerate probabilistic error, whereas using functional algorithm-related information might have had a different effect if, say, it was to explain how the algorithm works in lay-person terms. Nevertheless, it does seem reasonable that knowledge of social norms can serve a human user in augmented decision making by reducing the cognitive strain imposed on the user. When provided with a new algorithmic tool, for instance, a user has to make a judgement of the tool's reliability: How consistent is it? Whose interests has it been programmed to abide by? In receiving information about how others have appraised the tool's reliability, the user is able to effectively crowdsource these reliability judgements and focus on the decision task at hand (Alexander et al., 2018).

3.3.2 | Solution: Behavioral design

Although the literature around incentivization of algorithm utilization is riddled with inconsistencies, it is fair to assume that motivating human decision makers to heed algorithmic judgement requires consciously framing the decision context. This means approaching algorithm aversion as a project of behavior change in which hardwired organizational routines and social norms pose as major obstacles. A number of suggestions have been made along these lines: Choragwicka and Janta (2008) suggest framing the benefits of algorithm utilization in relatable terminology, Alexander et al. (2018) propose manipulating the perceived social consensus, and Fisher (2008), Klimoski and Jones (2008), and Kuncel (2008) advocate for localized reward schemes that apply to specific decision-making roles in organizations. Each of these holds promise, and the most effective incentivization program is likely to vary by environment. For this reason, it seems likely that the implementation of successful algorithmic decision necessitates context-specific behavioral design. Much like the popular use of behavioral economics for steering healthy eating habits or financial saving, the utilization of algorithmic decision aids could plausibly be improved with a program of transparent nudges (Bruns, Kantorowicz-Reznichenko, Klement, Jonsson, & Rahali, 2018; Thaler & Sunstein, 2008) and boosts (Hertwig & Grüne-Yanoff, 2017) that remedy human decision makers' motivational deficiencies without impinging on their autonomy. However, it should be noted that such an approach to resolving algorithm aversion may not necessarily be sustainable. Nudges have been critiqued for diverting efforts from more substantive solutions (e.g., Hagmann, Ho, & Loewenstein, 2019). It will, thus, be important to not allow a quick fix like behavioral design to crowd out support for developing more costly but more impactful solutions.

3.4 | Cognitive compatibility

3.5 | Problem: Combatting intuition

Algorithmic decision making inherently requires the integration of two decision processes: that of the human decision maker and that of the algorithmic aid. Both decision processes need to be mapped and understood transparently enough for them to be overlaid, lest they simply run in parallel and confront each other at the ultimate decision. For this reason, cognitive compatibility—the recognition and alignment of human and algorithmic decision processes—is crucial for successful augmented decision making. Without cognitive compatibility, algorithmic aids simply combat rather than engage human intuition.

Efforts to compatibly match decision aids to decision makers were made in early literature by exploring the role of decision makers' cognitive style or decision style (e.g., Alavi & Henderson, 1981; Benbasat & Taylor, 1978; Benbasat & Taylor, 1981; Er, 1988; Rao, Jacob, & Lin, 1992). Primarily, this line of research aimed to characterize the nature of human information processing so that the role of computer systems that support human decisions might be better understood (Robey & Taggart, 1982, p. 63). This largely entailed classifying decision makers on continuums, such as heuristic versus analytic, and comparing these decision makers' performance with various decision aids. Research here showed that decision makers' cognitive style predicted how they search, organize, and analyze data (Benbasat & Taylor, 1978; Moldafsky & Kwon, 1994). However, this research was also subject to critique that pointed out the malleability of cognitive style under situational pressures and the potential for decision makers' predispositions or biases to be exacerbated if their intuitive thinking is conformed to, rather than complemented (Huber, 1992; Robey, 1992). Although this body of work has grown dated, it is nonetheless important as it highlighted the necessity of modeling human intuitive processes for algorithmic augmentation to be plausible.

More recently, researchers have moved beyond the concept of cognitive style in favor of identifying specific heuristics and biases in human cognition that prevent decision makers from utilizing decision aids effectively. That is, although much attention is given to the opaque, black-boxed nature of algorithms (Christin, 2017; Dietvorst et al., 2015; Eastwood et al., 2012), research suggests that human decision making operates through a black box of its own: intuition. For instance, decision aiding naturally expects a decision maker to adapt his intuition and/or deliberate analyses, but to do so, he would have to understand descriptively the mental processes underlying his unaided intuitive choice well enough to prescribe how to practically transform that intuition into the ideal judgement (Brown, 2015, p. 217). However, research shows that people display persistent, unrecognized overconfidence (e.g., Arkes et al., 1986; Brown & Jones, 1998; Eining et al., 1997; Sieck & Arkes, 2005) and conservatism (e.g., Lim & O'Connor, 1996). These biases need to be considered as integral parts of intuitive decision making, not random miscues, and should thus be accommodated in order to achieve cognitive compatibility and resolve algorithm aversion.

3.5.1 | Solution: Engaging intuition

Using normative theories of how decision making ought to take place as the basis for designing decision aids requires valid models of the descriptive decision processes that people actually use to navigate information (Brown, 2015). This is a serious research agenda in its own right, but the ability to bridge the paradoxical relationship of intuition and rationality is needed so that algorithmic decision aids can pick up people where they stand and make improvements to the decision process that people already follow. That is, in place of requiring people to learn a new process from scratch, one can develop prescriptive aids for intuitive and effective use (Hafenbrädl et al., 2016, p. 217). This means that overcoming algorithm aversion requires carefully examining the subconscious processes that lead up to an intuitive decision to identify the criterion used by human decision makers for gathering and evaluating information under environmental restrictions (Mullins & Rogers, 2008; Thayer, 2008). In doing so, decision making can be broken down into a multistep procedure and the potential for integrating algorithmic judgement increases with each discrete step. By adding transparency on both sides of human-algorithm interaction, the agent-to-agent alignment of decision processes will inherently afford more opportunity for interaction, trust building, and confidence calibration, regardless of the decision task or structure of the environment. However, pushing for algorithms to be transparent often comes as a trade-off with the performance of the algorithm. For example, although the decision tree model advocated by Hafenbrädl et al. (2016) is indeed transparent and interpretable for a user, this type of algorithm is only suited to aid in binary classification tasks. On the other hand, neural networks can be trained to aid in wide-ranging decision tasks, but these are the prototypical black boxes that tend to give rise to algorithm aversion.

3.6 | Divergent rationalities

3.6.1 | Problem: Conflicting concepts of rationality

The algorithmic decision making literature has largely ignored the plurality of views of how people make decisions in the real world. A significant proportion of existing research that addresses algorithm aversion and augmented decision making has uncritically adopted the view of the heuristics-and-biases program (e.g., Kahneman, 2003, 2011; Kahneman, Slovic, & Tversky, 1982; Appendix, Table 3), which originates in research cataloging the many cognitive illusions that result from human decision makers' inability to perform rational calculations. Through this perspective, algorithmic decision aids are understood to be a kind of cognitive fix for the natural limitations of human thinking, with the ultimate aim of pushing back the bounds of rationality. This can be conspicuously seen in research that measures decision performance by comparing descriptive results with normative optimality, which has in many ways been considered the gold standard for decision analysis (e.g., Kahn & Baron, 1995; Lim & O'Connor, 1996; Sieck & Arkes, 2005; cf. Brown & Vari, 1992; Hafenbrädl et al., 2016). Undoubtedly, the heuristics-and-biases program's view of decision making and rationality has added to the algorithmic decision making literature by identifying individuals' cognitive and motivational deficiencies that could benefit from complementary augmentation. But, by relying on one theory of decision making, algorithmic decision making researchers have restricted themselves where other views, namely that of fast-and-frugal heuristics (e.g., Arkes, Gigerenzer, & Hertwig, 2016; Gigerenzer, Todd, & ABC Research Group, 1999; Hertwig, Hoffrage, & ABC Research Group, 2013), offer value. The fast-and-frugal perspective emphasizes the role of simple heuristics—formal search, stopping, and decision rules that human decision makers deploy under uncertainty—that improve decision making and inference. Crucially, this view defines good decision making as ecologically rational[3] rather than focusing on axioms of traditional rational choice (e.g., internal coherence or transitivity). Here, the structure of the decision task and the informational environment in which decision making takes place is shown to dictate the targeted decision outcome that a human decision maker aims toward.

The effects of task structure on decision aid utilization has been subject to significant investigation in older literature (e.g., Benbasat & Taylor, 1978; Carey & Kacmar, 2003; Er, 1988; Green & Hughes, 1986; Kahn & Baron, 1995; Sage, 1981; Sanders & Courtney, 1985). For the most part, this research tended to support the notion that human decision makers were more likely to seek the advice of decision aids in unstructured decision tasks, but that most decision aids were more suited to structured problems. Although the explicit focus on task structure has seemingly dwindled in recent years, this literature is reminiscent of the fast-and-frugal view that human decision makers' decision strategies are contingent on the statistical structures available in the environment (e.g., What calculable risks and alternatives are known?). Put simply, human decision makers often operate in a world of uncertainty (where alternatives, consequences, and probabilities are unknown and optimization is unfeasible) whereas algorithms operate in a world of risk (where probabilities are known or calculable and optimization should be the objective; Hafenbrädl et al., 2016). The best decision strategy under risk is often not the best decision under uncertainty. So, when a human decision maker or an algorithmic aid is unable to reconcile its view of what constitutes a good decision under the specific circumstances of a given task (i.e., the environment) with the other, algorithm aversion is observed.

[3]Ecological rationality is a practical account that claims the rationality of a decision is contingent on the environment in which it occurs. Ecological rationality violates rational choice theory's normative criteria and measures decision models on their predictive power under uncertainty (instead of data fitting), competitive model testing (instead of null hypothesis testing), and real-world validity (instead of internal coherence) (Todd & Gigerenzer, 2007).

3.6.2 | Solution: Aiding ecological rationality

Perhaps as a result of the focus on the heuristics-and-biases program's theory of rationality, algorithmic decision making faces a conspicuous research-practice gap where complexity is seemingly valued more than real-world suitability (Brown, 2015; Hafenbrädl et al., 2016). Considering the time, accessibility, and cost-efficiency restrictions of real-world decision making, designing algorithmic aids to help achieve ecological rationality rather than probabilistic optimality presents an opportunity for human and algorithmic problem-solving strategies to further complement one another. Patterson (2017) points out that regardless of one's view on decision making and rationality, there is consensual agreement that intuition is the leading force in human cognition. Rather than directly contesting it, algorithmic decision making systems would thus benefit from being modeled to suit the range of rationalities that intuitive thinking abides by (Sage, 1981; Westin, Borst, & Hilburn, 2016). For researchers and practitioners alike, this means ensuring that algorithmic decision aids work for the human decision maker and not vice versa. Understanding alternative decision making theories, like that of fast-and-frugal heuristics, not as descriptions of irrational behavior but as a valid conceptualization of real-world cognition, will serve to both advance applied decision making research and inform the design of augmented decision making systems that are less prone to algorithm aversion. Thus, we advocate for accepting the plurality of decision making and rationality theory and exploiting its variety for the betterment of algorithmic decision making. Importantly, this is not a suggestion to uncritically cater to human users' concepts of rationality, which could plausibly lead to incorporating unfavorable biases into augmented decision outcomes. Instead, aiding for ecological rationality entails the identification of which models perform best under different constraints, or at different points on the risk-uncertainty continuum, and ensuring that algorithmic aids are not unduly wedded to a single normative theory.

4 | DISCUSSION

The objective of this review has been to curate the existing research that explains why human and nonhuman (i.e., algorithmic, statistical, machine, etc.) decision making is so difficult to merge, particularly in the context of algorithmic decision making's growing ubiquity and recent findings of algorithm aversion. Although we have identified expectations and expertise, decision autonomy, incentivization, cognitive compatibility, and divergent rationalities as distinctive themes, it is important to remember they are not to be taken as independent of one another. Mapping the interdependencies between the mechanisms underlying algorithm aversion and recognizing the value of theory integration (as opposed to theory generation), seems necessary when addressing an interdisciplinary topic like algorithm aversion. In the following paragraphs, we return to each of the five themes in order to make such interdependencies clear, to link the reviewed literature to relevant theories and hypotheses that were not explicitly addressed, and to suggest avenues for future research.

At the forefront of any effort to remedy algorithm aversion has to be attention to the expectations and expertise that human decision makers inevitably carry into human-algorithm interactions. This is something that every one of us is susceptible to. We hold prior beliefs about how decisions should be made, what variables carry weight, and what outcomes are right and wrong under specific conditions. These beliefs influence the way we interact with a decision aid and the degree to which we update our beliefs when provided with algorithmically generated insights, regardless of whether our prior beliefs are accurate or not. This means that if such beliefs go unaddressed, essentially any advances in algorithm design or changes to the decision making environment can be subverted. Indeed, algorithmic literacy should be encouraged, and it will likely develop naturally, albeit slowly, as exposure to algorithms increases in volume and variety. However, a concept that has not been deeply explored in algorithmic decision making despite seeming relevance is theory of mind, which broadly refers to an agent's ability to impute mental states, intentions, and beliefs to itself and others (Premack & Woodruff, 1978). Although research around theory of mind has predominantly been reserved to developmental psychology domains, its applicability to artificial intelligence and robotics has been recognized in recent work (e.g., Rabinowitz et al., 2018; Winfield, 2018). For example, theory of mind explains that humans rely on high-level models of others for daily social reasoning: We infer what others are thinking in order to communicate and cooperate better. Despite the fact that these models do not include references to the neural mechanisms at play in others' brains, they are extremely efficient in everyday life. In regard to algorithmic decision making, this suggests that human-algorithm coordination needs not human agents who grasp the code behind the algorithmic aid, but rather a high-level model of its purpose and perception. This idea is perhaps an obvious one, but it is one that does not appear in the reviewed literature. Although a theory of (algorithmic) mind (cf. theory of machine, Logg, Minson, & Moore, 2019) naturally applies to the idea of algorithmic literacy, its connections to other themes are also apparent: Does an accurate internal model of an algorithm's perceptions moderate the degree to which a human user feels a need for control, the degree to which a user requires extrinsic incentivization, the degree to which a user is capable of integrating an algorithm's decision process, or the degree to which a user is able to align with an algorithm's rational decision outcome?

In the reviewed literature, decision autonomy is mostly portrayed as some form of post hoc deliberation where the human user of an algorithmic aid is granted opportunities to edit that aid's judgement (e.g., Dietvorst et al., 2016). Yet, there are other ways of distributing autonomy between a human and algorithm in decision making. In fact, it is in the original work of Meehl (1954), Einhorn (1972), and Dawes (1979) that the idea of shared decision autonomy, between statistical and clinical methods of judgement, was first substantiated with empirical data (also see Sawyer, 1966). As Camerer (1981) explains, the experiments that pitted clinical and statistical judgement against one another lead to the conclusion that human decision makers are in fact quite good at collecting data (i.e., providing the input for a model), but are bad at combining it. Conversely, algorithms are good at combining data (i.e., calculating the output based on a model), but are bad at collecting it. Bootstrapping models mobilize this finding by allowing human decision makers to intuitively gather and encode information and then having this human-collected information put into the empirically established relationships of a regression algorithm that ultimately pull human users up by their proverbial bootstraps (Camerer, 1981, p. 411). Practically speaking, bootstrapping models are indeed a version of human-in-the-loop decision making because human agents play a direct part in constructing the algorithmic model, which seems to be a plausible remedy for algorithm aversion driven by a lack of decision control. Moreover, bootstrapping models could also feature in building cognitive compatibility because they inherently break down the decision making process into delegable, comprehensible steps. To our knowledge, however, no study has looked at how bootstrapping models might fare in terms of algorithm aversion. Do human decision makers feel a greater degree of autonomy with bootstrapping models as compared with models with predetermined inputs (e.g., the model used in Dietvorst et al., 2015)? Could the reduction in algorithm aversion caused by allowing human users to modify the algorithmic aid's output (e.g., Dietvorst et al., 2016) be furthered by allowing them to also modify the algorithmic aid's input?

Intrinsically dependent on the social or organizational decision making environment, the role of incentivization in algorithm aversion is perhaps the most distinctive. But given the apparently high degree of domain-specificity of incentivization's effects, making confident proposals for motivating human decision makers to heed an algorithm's advice is difficult. Nevertheless, related research on algorithms in social contexts suggests certain principles, like justifiability and interpretability (e.g., Brkan, 2019; Goodman & Flaxman, 2016), could be key to introducing algorithmic decision making in traditionally human environments. With consideration of the social pressures placed upon individuals (and their peers), particularly in organizational settings, it seems that algorithmic aids need not only to be accurate, but also to be understood by the humans using them (Yeomans, Shah, Mullainathan, & Kleinberg, 2019). For example, in medical decision making, or even routine operational business decisions, the human tasked with making the decision is held accountable. If a decision maker does not understand how an algorithm came to its conclusion, then utilizing it may jeopardize that decision maker's ability to justify a decision with implicated stakeholders. In fact, this functional role of justifiability as an incentive for decision making performance has been a long withstanding topic (e.g., Ashton, 1990, 1992; Tetlock, 1985), which indeed appears sporadically throughout the reviewed literature (e.g., Brown & Jones, 1998; Eining et al., 1997; Landsbergen et al., 1997; Scherer et al., 2015; Sieck & Arkes, 2005; Swinney, 1999). What these inquiries suggest is that the designers of augmented decision making systems need to approach the human-algorithm interaction not as a one-to-one relationship, but rather as a political relationship, because most important decisions are not, after all, the product of isolated information processors; they are the product of intensive interactions among members of groups (Tetlock, 1985, p. 298). Although an ecologically valid experiment on algorithm aversion in social contexts is difficult to imagine, the themes presented in this review generate some reasonable hypotheses: Might participatory, human-in-the-loop decision making systems be more justifiable for the average user? Could a more interpretable algorithmic decision process be better at dissolving users' false expectations?

Naturally, the influence of decision autonomy and social incentivization on algorithm aversion depends on cognitive compatibility between human and algorithm. Distributing autonomy across certain steps of a decision making process and making interpretable, justifiable decisions that suit the social context requires that the human user of an algorithmic aid is able to recognize when and why the algorithm's process overlaps or diverges from his or her own. The more stages of a human decision making process that can be engaged by an algorithm, the more opportunities there are for that algorithm's judgement to be integrated. Based on this conclusion, it is clear that both the designers of algorithmic aids and the humans that are supposed to utilize the aids need some knowledge of decision processes. Simon (1977, p. 77) famously breaks down decision making into three essential steps: intelligence (searching for information and identifying alternatives), design (calculating the consequences of alternatives), and choice (evaluating and selecting an alternative). Although pointing out that each of these steps could be approached as decisions in themselves, Simon's (1977) simple model has significantly influenced the development of decision aids by translating seemingly complex, holistic processes into programmable increments (Pomerol & Adam, 2006). Traditionally, different types of decision aids targeted different steps of the decision making process; however, the rise in machine learning means that algorithms are increasingly able to take over decision processes in their entirety. This progress is undoubtedly exciting for anyone interested in artificial intelligence, but it poses serious challenges for the prospect of cognitively compatible, augmented decision making. As this review describes, algorithm aversion can manifest when human and algorithmic decision processes run in parallel, largely because the lack of interaction lends itself toward poor confidence calibration on behalf of the human user (Muir, 1987; Sieck & Arkes, 2005). Put simply, if algorithms dictate the whole of a decision process and only propose an ultimate choice, this is debatably more along the lines of automation rather than augmentation. Whether or not, human decision makers are willing to interact with contemporary algorithms at each step of a decision process remains to be seen. But, if algorithms can aid a wider range of aspects in the decision making process, then it seems plausible that they could be properly utilized by a wider range of human decision makers and in a wider range of decision making environments. This line of inquiry too can be readily translated into empirical questions: What steps of decision making are human users most reluctant/willing to delegate to an algorithmic aid? What types of decision tasks benefit most from (in terms of performance and algorithm aversion) algorithmically automated intelligence, design, and choice?

Implied theories of decision making and rationality underlie each of the previous four themes. Broadly speaking, these theories provide the normative logic upon which a decision can be evaluated as good or bad. If one person considers a rational decision to be one that adheres to traditional principles of internal consistency (e.g., transitivity or additivity of probabilities), and another person considers a rational choice to be one that has the best external performance (e.g., timeliness, justifiability, and cost-efficiency), then these two people are aiming toward fundamentally different decision goals (Gigerenzer, 2001). As was previously mentioned, there are in fact a plurality of views on decision making and rationality that people employ in the real world. Because algorithmic aids inherently rely on some programmable decision making ideal, the underlying theory has significant ramifications for how the aids can be used in practice. Although the rather dismal view of human decision making capabilities put forth by the heuristics-and-biases program has been widely presupposed in the reviewed literature, it is not difficult to find or imagine examples of algorithms that are founded on other theories of rationality. For example, where the heuristics-and-biases program has a regression model, the fast-and-frugal view has a signal detection-style decision tree. These fast-and-frugal trees (Hafenbrädl et al., 2016; Phillips, Neth, Woike, & Gaissmaier, 2017) are especially relevant to the algorithm aversion discussion not only because they allow the human decision maker to dictate the external measures upon which an augmented decision will be judged, but also because they are transparent. This in turn suggests that human users could interpret, justify, control, and interact with a fast-and-frugal decision aid, which touches on virtually all the drivers of algorithm aversion covered in this review. Once again, however, this is an empirical question that remains unanswered. Although we have lumped together various algorithmic models under the label of nonhuman in this review, how might various models compare in terms of algorithm aversion? Like competitive model testing and out-of-sample prediction provide alternative methodological principles to null hypothesis testing and data fitting, respectively, could adding algorithm aversion to the arsenal of model metrics be the next step for augmented decision making?

Needless to say, addressing algorithm aversion is a research venture that is well-informed by rich existing literature, but overall, it is a venture that has failed to translate into satisfactory practice. Like the clinical versus statistical prediction debate before it, the discourse around algorithmic decision making has been primarily concerned with comparing human and nonhuman decision makers, rather than addressing the practical issues that prevent combining the best aspects of the two. Perhaps due to the need for improved communication across disciplines, the existing literature also struggles to define its key constructs. Although the use of varying terminologies is not inherently bad, a lack of clarity can cloud important concepts, and at times this leads to misperceptions where reconcilable findings are presented as empirical contradictions, and vice versa. With recent work suggesting that humans display algorithm appreciation (Logg et al., 2019) rather than algorithm aversion, the need for clarity seems especially pressing.

5 | CONCLUSION

Research shows that augmented, human-algorithm decision systems outperform both lone human and lone algorithm decision makers (e.g., Einhorn, 1972; Kleinmuntz, 1990). More recently, however, fervent calls against opaque algocracy have occupied the limelight (e.g., Danaher, 2016; O'Neil, 2016; Pasquale, 2015). Making the human-algorithm relationship work thus seems to be in everyone's best interest, but practical solutions for algorithm aversion have yet to take shape. Despite significant advances in our understanding of the neural mechanisms underlying advice taking, our cognitive decision processes and limitations, the computing capabilities of algorithms, and the perceptions of algorithms in organizational settings, the links between such findings remain sparse. This review has highlighted the range of perspectives one can take in appraising algorithm aversion in augmented decision making, and as such, it seems that for real-world progress to be made there needs to be at least an equal emphasis on theory integration as there has been on theory generation. Given digitalization's and datafication's rapid expansion into evermore aspects of everyday life, there should be no lack of impetus to build cooperative relationships with the algorithms that help us make sense of a quantified society.

ORCID

Jason W. Burton https://orcid.org/0000-0002-6797-2299

REFERENCES

Alavi, M., & Henderson, J. C. (1981). An evolutionary strategy for implementing a decision support system. *Management Science*, *27*(11), 1309–1323. https://doi.org/10.1287/mnsc.27.11.1309

Alexander, V., Blinder, C., & Zak, P. J. (2018). Why trust an algorithm? Performance, cognition, and neurophysiology. *Computers in Human Behavior.*, *89*, 279–288. https://doi.org/10.1016/j.chb.2018.07.026

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes*, *37*, 93–110. https://doi.org/10.1016/0749-5978(86)90046-4

Arkes, H. R., Gigerenzer, G., & Hertwig, R. (2016). How bad is incoherence? *American Psychological Association*, *3*(1), 20–39. https://doi.org/10.1037/dec0000043

Arkes, H. R., Shaffer, V. A., & Medow, M. A. (2007). Patients Derogate Physicians Who Use a Computer-Assisted Diagnostic Aid. *Medical Decision Making*, *27*, 189–202. https://doi.org/10.1177/0272989X06297391

Ashton, A. H., Ashton, R. H., & Davis, M. N. (1994). White-collar robotics: Levering managerial decision making. *California Management Review*, *37*(I), 83–110.

Ashton, R. H. (1990). Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *Journal of Accounting Research*, *28*, 148–180. https://doi.org/10.2307/2491253

Ashton, R. H. (1992). Effects of justification and a mechanical aid on judgment performance. *Organizational Behavior and Human Decision Processes*, *52*(2), 292–306. https://doi.org/10.1016/0749-5978(92)90040-E

Benbasat, I., & Taylor, R. N. (1978). The impact of cognitive styles on information system design. *MIS Quarterly*, *2*(2), 43–54. https://doi.org/10.2307/248940

Benbasat, I., & Taylor, R. N. (1981). Behavioral aspects of information processing for the design of management information systems. *IEEE Transactions on Systems, Man, and Cybernetics*, *12*, 439–450. https://doi.org/10.1109/TSMC.1982.4308848

Bourdieu, P. (1993). *The Field of Cultural Production*. New York, NY: Columbia University Press.

Brkan, M. (2019). Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond. *International Journal of Law and Information Technology*, *27*(2), 91–121.

Brown, D. L., & Jones, D. R. (1998). Factors that influence reliance on decision aids: A model and an experiment. *Journal of Information Systems*, *12*(2), 75–94.

Brown, R., & Vari, A. (1992). Towards a research agenda for prescriptive decision science: The normative tempered by the descriptive. *Acta Psychologica*, *80*, 33–47. https://doi.org/10.1016/0001-6918(92)90039-G

Brown, R. V. (2015). Decision science as a by-product of decision-aiding: A practitioner's perspective. *Journal of Applied Research in Memory and Cognition*, *4*, 212–220. https://doi.org/10.1016/j.jarmac.2015.07.005

Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M. L., & Rahali, B. (2018). Can nudges be transparent and yet effective? *Journal of Economic Psychology*, *65*, 41–58. https://doi.org/10.1016/j.joep.2018.02.002

Camerer, C. (1981). General conditions for the success of bootstrapping models. *Organizational Behavior and Human Performance*, *27*(3), 411–422. https://doi.org/10.1016/0030-5073(81)90031-3

Carey, J. M., & Kacmar, C. J. (2003). Toward a general theoretical model of computerbased factors that affect managerial decision making. *Journal of Managerial Issues*, *15*(4), 430–449.

Carrigan, N., Gardner, P. H., Conner, M., & Maule, J. (2004). The impact of structuring information in a patient decision aid. *Psychology & Health*, *19*(4), 457–477. https://doi.org/10.1080/08870440310001652641

Choragwicka, B., & Janta, B. (2008). Why is it so hard to apply professional selection methods in business practice? *Industrial and Organizational Psychology*, *1*(3), 355–358. https://doi.org/10.1111/j.1754-9434.2008.00062.x

Christin, A. (2017). Algorithms in practice: Comparing web journalism and criminal justice. *Big Data & Society*, *4*, 1–14. https://doi.org/10.1177/2053951717718855

Colarelli, S. M., & Thompson, M. (2008). Stubborn reliance on human nature in employee selection: Statistical decision aids are evolutionarily novel. *Industrial and Organizational Psychology*, *1*(3), 374–351. https://doi.org/10.1111/j.1754-9434.2008.00060.x

Danaher, J. (2016). The threat of algocracy: Reality, resistance and accommodation. *Philosophy and Technology*, *29*(3), 245–268. https://doi.org/10.1007/s13347-015-0211-1

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*(7), 571–582. https://doi.org/10.1037/0003-066X.34.7.571

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2016). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Eastwood, J., Snook, B., & Luther, K. (2012). What people want from their professionals: Attitudes toward decision-making strategies. *Journal of Behavioral Decision Making*, *25*, 458–468. https://doi.org/10.1002/bdm.741

Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational Behavior and Human Performance*, *7*(1), 86–1972. https://doi.org/10.1016/0030-5073(72)90009-8

Einhorn, H. J. (1986). Accepting error to make less error. *Personality Assessment*, *50*(3), 387–395. https://doi.org/10.1207/s15327752jpa5003

Eining, M. M., Jones, D. R., & Loebbecke, J. K. (1997). Reliance on decision aids: An examination of auditors' assessment of management fraud. *Auditing: A Journal of Practice & Theory*, *16*(2), 1–19.

Einorn, H. J. (1986). Accepting error to make less error. *Journal of Personality Assessment*, *50*(3), 387–395. https://doi.org/10.1207/s15327752jpa5003_8

Er, M. C. (1988). Decision Support Systems: A Summary, Problems, and Future Trends. *Decision Support Systems*, *4*, 355–363. https://doi.org/10.1016/0167-9236(88)90022-X

Fisher, C. D. (2008). Why don't they learn? *Industrial and Organizational Psychology*, *1*(3), 364–366. https://doi.org/10.1111/j.1754-9434.2008.00065.x

Gigerenzer, G. (2001). Decision making: Nonrational theories. *International Encyclopedia of the Social and Behavioral Sciences*, *5*, 3304–3309. https://doi.org/10.1016/B978-0-08-097086-8.26017-0

Gigerenzer, G., Todd, P. M., & ABC Research Group (1999). Fast and frugal heuristics: The Adaptive toolbox. *Simple Heuristics That Make Us Smart.*, *7*, 93–104. https://doi.org/10.1177/1354067X0171006

Goodman, B., & Flaxman, S. (2016). European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". *ArXiv Preprint*, *38*, 1–9. Retrieved from. https://doi.org/10.1609/aimag.v38i3.2741

Goodyear, K., Parasuraman, R., Chernyak, S., de Visser, E., Madhavan, P., Deshpande, G., & Krueger, F. (2017). An fMRI and effective connectivity study investigating miss errors during advice utilization from human and machine agents. *Social Neuroscience*, *12*(5), 570–581. https://doi.org/10.1080/17470919.2016.1205131

Goodyear, K., Parasuraman, R., Chernyak, S., Madhavan, P., Deshpande, G., & Krueger, F. (2016). Advice Taking from Humans and Machines: An fMRI and effective connectivity study. *Frontiers in Human Neuroscience*, *10*(542), 1–15. https://doi.org/10.3389/fnhum.2016.00542

Green, G. I., & Hughes, C. T. (1986). Effects of decision support systems training and cognitive style on decision process attributes. *Journal of Management Information Systems*, *3*(2), 83–93. https://doi.org/10.1080/07421222.1986.11517764

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, *2*, 293–323. https://doi.org/10.1037/1076-8971.2.2.293

Hafenbrädl, S., Waeger, D., Marewski, J. N., & Gigerenzer, G. (2016). Applied decision making with fast-and-frugal heuristics. *Journal of Applied Research in Memory and Cognition*, *5*, 215–231. https://doi.org/10.1016/j.jarmac.2016.04.011

Hagmann, D., Ho, E. H., & Loewenstein, G. (2019). Nudging out support for a carbon tax. *Nature Climate Change*, *9*(6), 484–489. https://doi.org/10.1038/s41558-019-0474-0

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and Boosting: Steering or Empowering Good Decisions. *Perspectives on Psychological Science*, *12*(6), 973–986. https://doi.org/10.1177/1745691617702496

Hertwig, R., Hoffrage, U., & ABC Research Group (2013). *Simple heuristics in a social world*. New York, NY: Oxford University Press.

Highhouse, S. (2008a). Facts are stubborn things. *Industrial and Organizational Psychology*, *1*(3), 373–376. https://doi.org/10.1111/j.1754-9434.2008.00069.x

Highhouse, S. (2008b). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, *1*(3), 333–342. https://doi.org/10.1111/j.1754-9434.2008.00058.x

Huber, G. P. (1992). Response to Rao, et al: How to deal with cognitive style. *MIS Quarterly*, *16*(2), 153–154. https://doi.org/10.2307/249572

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, *61*, 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Kahn, B. E., & Baron, J. (1995). An exploratory study of choice rules favored for high-stakes decisions. *Journal of consumer Psychology*, *4*(4), 305–328.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Macmillan.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9780511809477

Kleinmuntz, B. (1990). Why we still use our heads instead of formulas: Toward an integrative approach. *Psychological Bulletin*, *107*(3), 296–310. https://doi.org/10.1037/0033-2909.107.3.296

Klimoski, R., & Jones, R. G. (2008). Intuiting the selection context. *Industrial and Organizational Psychology*, *1*(3), 352–354. https://doi.org/10.1111/j.1754-9434.2008.00061.x

Kuncel, N. R. (2008). Some New (and Old) Suggestions for improving personnel selection. *Industrial and Organizational Psychology*, *1*(3), 343–346. https://doi.org/10.1111/j.1754-9434.2008.00059.x

Lamberti, D. M., & Wallace, W. A. (1990). Expert systems intelligent interface assessment of design: An empirical knowledge presentation in expert systems1. *MIS Quarterly*, *14*(3), 279–311. https://doi.org/10.2307/248891

Landsbergen, D., Coursey, D. H., Loveless, S., & Shangraw, R. F. (1997). Decision quality, confidence, and commitment with expert systems: An experimental study. *Journal of Public Administration Research and Theory*, *7*(1), 131–157. https://doi.org/10.1093/oxfordjournals.jpart.a024336

Lim, J. S., & O'Connor, M. (1996). Judgmental forecasting with interactive forecasting support systems. *Decision Support Systems*, *16*, 339–357. https://doi.org/10.1016/0167-9236(95)00009-7

Lodato, M. A., Highhouse, S., & Brooks, M. E. (2011). Predicting professional preferences for intuition-based hiring. *Journal of Managerial Psychology*, *26*(5), 352–365. https://doi.org/10.1108/02683941111138985

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Mackay, J. M., & Elam, J. J. (1992). A comparative study of how experts and novices use a decision aid to solve problems in complex knowledge domains. *Information Systems Research*, *3*(2), 150–172. https://doi.org/10.1287/isre.3.2.150

Martin, S. L. (2008). Managers also overrely on tests. *Industrial and Organizational Psychology*, *1*(3), 359–360. https://doi.org/10.1111/j.1754-9434.2008.00063.x

Meehl, P. E. (1954). *Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence*.

Mosier, K. L., & Fischer, U. M. (2010). Judgment and decision making by individuals and teams: issues, models, and applications. *Reviews of human factors and ergonomics*, *6*(1), 198–256. https://doi.org/10.1518/155723410X12849346788822

Moldafsky, N., & Kwon, I.-W. (1994). Attributes affecting computer-aided decision making — a literature survey. *Computers in Human Behavior*, *10*(3), 299–323. https://doi.org/10.1016/0747-5632(94)90057-4

Montazemi, A. R. (1991). The impact of experience on the design of user interface. *International Journal of Man-Machine Studies*, *34*, 731–749. https://doi.org/10.1016/0020-7373(91)90022-Y

Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies*, *27*, 527–539. https://doi.org/0020-7373/87/050527, https://doi.org/10.1016/S0020-7373(87)80013-5

Mullins, M. E., & Rogers, C. (2008). Reliance on intuition and faculty hiring. *Industrial and Organizational Psychology*, *1*(January), 370–371. https://doi.org/10.1111/j.1754-9434.2008.00067.x

O'Brien, J. (2008). Interviewer resistance to structure. *Industrial and Organizational Psychology*, *1*(3), 367–369. https://doi.org/10.1111/j.1754-9434.2008.00066.x

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown Publishers.

Önkal, D., Goodwin, P., Thomson, M., Gonul, S., & Pollock, A. (2009). The Relative Influence of Advice From Human Experts and Statistical Methods on Forecast Adjustments. *Journal of Behavioral Decision Making*, *22*, 390–409. https://doi.org/10.1002/bdm.637

Pagano, T. C., Pappenberger, F., Wood, A. W., Ramos, M.-H., Persson, A., & Anderson, B. (2016). Automation and human expertise in operational river forecasting. *Wiley Interdisciplinary Reviews: Water*, *3*(5), 692–705. https://doi.org/10.1002/wat2.1163

Pasquale, F. (2015). *The Black Box Society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Patterson, R. E. (2017). Intuitive cognition and models of human-automation interaction. *Human Factors*, *59*(1), 101–115. https://doi.org/10.1177/0018720816659796

Phillips, J. M., & Gully, S. M. (2008). The role of perceptions versus reality in managers' choice of selection decision aids. *Industrial and Organizational Psychology*, *1*(3), 361–363. https://doi.org/10.1111/j.1754-9434.2008.00064.x

Phillips, N. D., Neth, H., Woike, J. K., & Gaissmaier, W. (2017). FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees. *Judgment and Decision Making*, *12*(4), 344–368. Retrieved from. http://journal.sjdm.org/17/17217/jdm17217.pdf

Pomerol, J.-C., & Adam, F. (2006). On the legacy of Herbert Simon and his contribution to decision-making support systems and artificial intelligence. In *Intelligent Decision-making Support Systems* (pp. 25–43). London: Springer. https://doi.org/10.1007/1-84628-231-4_2

Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, *36*, 691–702. https://doi.org/10.1002/for.2464

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, *4*, 515–526. https://doi.org/10.1017/S0140525X00076512

Rabinowitz, N. C., Perbet, F., Song, H. F., Zhang, C., Eslami, S. M. A., & Botvinick, M. (2018). machine theory of mind. *ArXiv Preprint*. Retrieved from http://arxiv.org/abs/1802.07740

Rao, H. R., Jacob, V. S., & Lin, F. (1992). Hemispheric specialization, cognitive differences, and their implications for the design of decision support systems. *MIS Quarterly*, *16*(2), 145–151. https://doi.org/10.2307/249570

Robey, D. (1992). Response to Rao, et al: More ado about cognitive style and DSS design. *MIS Quarterly*, *16*(2), 151–153. https://doi.org/10.2307/249571

Robey, D., & Taggart, W. (1982). Human information processing in information and decision support systems. *MIS Quarterly*, *6*(2), 61–73. https://doi.org/10.2307/249283

Sage, A. P. (1981). Behavioral and organizational considerations in the design of information systems and processes for planning and decision support. *IEEE Transactions on Systems, Man, and Cybernetics*, *11*(6), 640–678. https://doi.org/10.1109/TSMC.1981.4308761

Sanders, G. L., & Courtney, J. F. (1985). A field study of organizational factors influencing dss success. *MIS Quarterly*, *9*(1), 77–93. https://doi.org/0.1016/0378-7206(84)90042-9

Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, *66*(3), 178–200. https://doi.org/10.1037/h0023624

Scherer, L. D., Zikmund-Fisher, B. J., Witteman, H. O., & Fagerlin, A. (2015). Trust in deliberation: the consequences of deliberative decision strategies for medical decisions. *Health Psychology*, *34*(11), 1090–1099. Retrieved from. https://doi.org/10.1037/hea0000203.supp

Sieck, W. R., & Arkes, H. A. L. R. (2005). The recalcitrance of overconfidence and its contribution to decision aid neglect. *Journal of Behavioral Decision Making*, *53*, *18*, 29–53. https://doi.org/10.1002/bdm.486

Simon, H. (1977). *The new science of management decision making*. New York: Harper and Row.

Sutherland, S. C., Harteveld, C., & Young, M. E. (2016). Effects of the advisor and environment on requesting and complying with automated advice. *ACM Transactions on Interactive Intelligent Systems.*, *6*, 1–36. https://doi.org/10.1145/2905370

Swinney, L. (1999). Consideration of the social context of auditors reliance on expert system output during evaluation of loan loss reserves. *International Journal of Intelligent Systems in Accounting, Finance & Management*, *8*, 199–213. https://doi.org/10.1002/(sici)1099-1174(199909)8:3<199::aid-isaf160>3.0.co;2-a

Tetlock, P. (1985). Accountability: The neglected social context of judgment and choice. *Research in Organizational Behavior*, *7*, 297–332. doi.org/0-89232-497-X

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Thayer, P. W. (2008). That's not the only problem. *Industrial and Organizational Psychology*, *1*(3), 372. https://doi.org/10.1111/j.1754-9434.2008.00068.x

Todd, P. M., & Gigerenzer, G. (2007). Environments That Makes Us Smart: Ecological Rationality. *Current Directions in Psychological Science*, *16*(3), 167–172. https://doi.org/10.1111/j.1467-8721.2007.00497.x

Westin, C., Borst, C., & Hilburn, B. (2016). Strategic conformance: Overcoming acceptance issues of decision aiding automation? *IEEE Transactions on Human-Machine Systems*, *46*(1), 41–52. https://doi.org/10.1109/THMS.2015.2482480

Whitecotton, S. M. (1996). The effects of experience and confidence on decision aid reliance: A causal model. *Behavioral Research in Accounting*, *8*, 194–216.

Winfield, A. F. T. (2018). Experiments in artificial theory of mind: From safety to story-telling. *Frontiers in Robotics and AI*, *5*(75), 1–13. https://doi.org/10.3389/frobt.2018.00075

Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making*, 1–12. https://doi.org/10.1002/bdm.2118

AUTHOR BIOGRAPHY

Jason W. Burton is a PhD candidate at the Centre for Cognition, Computation, & Modelling, situated in the Department of Psychological Sciences at Birkbeck, University of London. His research revolves around the topic of human rationality, particularly in the context of contemporary media, technology, and politics.

How to cite this article: Burton JW, Stein M-K, Jensen TB. A systematic review of algorithm aversion in augmented decision making. *J Behav Dec Making*. 2020;33:220–239. https://doi.org/10.1002/bdm.2155

APPENDIX

TABLE 1 Reference list with journals and methodologies

#	Reference	Journal	Methodology
1	Alavi and Henderson	Management Science	E
2	Alexander et al. (2018)	Computers in Human Behavior	E
3	Arkes et al. (1986)	Organizational Behavior and Human Decision Processes	E
4	Arkes et al. (2007)	Medical Decision Making	E
5	Ashton et al. (1994)	California Management Review	C
6	Benbasat and Taylor (1978)	MIS Quarterly	C
7	Benbasat & Taylor (1981)	IEEE Transactions on Systems, Man, and Cybernetics	C
8	Brown and Jones (1998)	J. of Information Systems	C
9	Brown and Vari (1992)	Acta Psychologica	C
10	Brown (2015)	J. of Applied Research in Memory and Cognition	C
11	Carey and Kacmar (2003)	J. of Managerial Studies	E
12	Carrigan et al. (2004)	Psychology and Health	E
13	Choragwicka and Janta (2008)	Industrial and Organizational Psychology	C
14	Christin (2017)	Big Data & Society	F
15	Colarelli and Thompson (2008)	Industrial and Organizational Psychology	C
16	Dietvorst et al. (2015)	J. of Experimental Psychology: General	E
17	Dietvorst et al. (2016)	Management Science	E
18	Eastwood et al. (2012)	J. of Behavioral Decision Making	E

(Continues)

TABLE 1 (Continued)

#	Reference	Journal	Methodology
19	Eining et al. (1997)	Auditing: A J. of Practice & Theory	E
20	Er (1988)	Decision Support Systems	C
21	Fisher (2008)	Industrial and Organizational Psychology	C
22	Goodyear et al. (2016)	Frontiers in Human Neuroscience	E
23	Goodyear et al. (2017)	Social Neuroscience	E
24	Green and Hughes (1986)	J. of Management Information Systems	E
25	Hafenbrädl et al. (2016)	J. of Applied Research in Memory and Cognition	C
26	Highhouse (2008a)	Industrial and Organizational Psychology	C
27	Highhouse (2008b)	Industrial and Organizational Psychology	C
28	Huber (1992)	MIS Quarterly	C
29	Jarrahi (2018)	MIS Quarterly	C
30	Kahn and Baron (1995)	J. of Consumer Psychology	E
31	Klimoski and Jones (2008)	Industrial and Organizational Psychology	C
32	Kuncel (2008)	Industrial and Organizational Psychology	C
33	Lamberti and Wallace (1990)	MIS Quarterly	F
34	Landsbergen et al. (1997)	J. of Public Administration Research and Theory	E
35	Lim and O'Connor (1996)	Decision Support Systems	E
36	Lodato et al. (2011)	J. of Managerial Psychology	E
37	Mackay and Elam (1992)	Information Systems Review	E
38	Martin (2008)	Industrial and Organizational Psychology	C
39	Moldafsky and Kwon (1994)	Computers in Human Behavior	C
40	Montazemi (1991)	International J. of Man-Machine Studies	E
41	Mosier & Fischer (2010)	Review of Human Factors and Ergonomics	C
42	Muir (1987)	International J. of Man-Machine Studies	C
43	Mullins and Rogers (2008)	Industrial and Organizational Psychology	C
44	O'Brien (2008)	Industrial and Organizational Psychology	C
45	Önkal et al. (2009)	J. of Behavioral Decision Making	E
46	Pagano et al. (2016)	Wiley Interdisciplinary Reviews: Water	C
47	Patterson (2017)	Human Factors	C
48	Phillips and Gully (2008)	Industrial and Organizational Psychology	C
49	Prahl and Van Swol (2017)	J. of Forecasting	E
50	Rao et al. (1992)	MIS Quarterly	C
51	Robey (1992)	MIS Quarterly	C
52	Robey and Taggart (1982)	MIS Quarterly	C
53	Sage (1981)	IEEE Transactions on Systems, Man, and Cybernetics	C
54	Sanders and Courtney (1985)	MIS Quarterly	F
55	Scherer et al. (2015)	Health Psychology	E
56	Sieck and Arkes (2005)	J. of Behavioral Decision Making	E
57	Sutherland, Harteveld, and Young (2016)	ACM Transactions on Interactive Intelligent Systems	E
58	Swinney (1999)	Intelligent Systems in Accounting, Finance & Management	E
59	Thayer (2008)	Industrial and Organizational Psychology	C
60	Westin et al. (2016)	IEEE Transactions on Human-Machine Systems	C
61	Whitecotton (1996)	Behavioral Research in Accounting	E

[a]Methodology: E = experimental; C = conceptual/review of literature; F = ethnography/field study.

TABLE 2 Concept matrix

#	Reference	Variable(s) in focus		
		Human	Algorithm	Environment
1	Alavi & Henderson	X		X
2	Alexander et al. (2018)	X		X
3	Arkes et al. (1986)	X		X
4	Arkes et al. (2007)			X
5	Ashton et al. (1994)	X		
6	Benbasat and Taylor (1978)	X	X	
7	Benbasat & Taylor (1981)	X	X	X
8	Brown and Jones (1998)	X	X	X
9	Brown and Vari (1992)	X	X	X
10	Brown (2015)		X	X
11	Carey and Kacmar (2003)	X	X	X
12	Carrigan, et al. (2004)		X	
13	Choragwicka and Janta (2008)	X		X
14	Christin (2017)			X
15	Colarelli and Thompson (2008)	X		
16	Dietvorst et al. (2015)	X	X	
17	Dietvorst et al. (2016)	X	X	
18	Eastwood et al. (2012)			X
19	Eining et al. (1997)		X	
20	Er (1988)	X	X	X
21	Fisher (2008)	X		
22	Goodyear et al. (2016)	X		
23	Goodyear et al. (2017)	X		
24	Green and Hughes (1986)	X		
25	Hafenbrädl et al. (2016)	X	X	X
26	Highhouse (2008a)	X		
27	Highhouse (2008b)	X		X
28	Huber (1992)	X		
29	Jarrahi (2018)	X	X	
30	Kahn and Baron (1995)	X	X	X
31	Klimoski and Jones (2008)			X
32	Kuncel (2008)		X	X
33	Lamberti and Wallace (1990)	X	X	X
34	Landsbergen et al. (1997)	X	X	X
35	Lim and O'Connor (1996)	X	X	
36	Lodato et al. (2011)	X		
37	Mackay and Elam (1992)	X		
38	Martin (2008)	X		X
39	Moldafsky and Kwon (1994)	X		
40	Montazemi (1991)	X		X
41	Mosier & Fischer (2010)	X	X	X
42	Muir (1987)		X	X
43	Mullins and Rogers (2008)			X
44	O'Brien (2008)	X		
45	Önkal et al. (2009)	X		X

(Continues)

TABLE 2 (Continued)

#	Reference	Variable(s) in focus: Human	Algorithm	Environment
46	Pagano et al. (2016)	X	X	
47	Patterson (2017)	X	X	
48	Phillips & Gully (2008)	X		
49	Prahl and Van Swol (2017)	X		
50	Rao et al. (1992)	X	X	
51	Robey (1992)	X		X
52	Robey and Taggart (1982)	X	X	
53	Sage (1981)	X		X
54	Sanders and Courtney (1985)			X
55	Scherer et al. (2015)	X		
56	Sieck and Arkes (2005)	X		
57	Sutherland et al. (2016)			X
58	Swinney (1999)	X		X
59	Thayer (2008)	X		
60	Westin et al. (2016)		X	X
61	Whitecotton (1996)	X		

TABLE 3 Concept matrix

#	Reference	Judgement and decision making orientation[a]: Heuristics and biases	Fast and frugal	N/A
1	Alavi & Henderson			X
2	Alexander et al. (2018)			X
3	Arkes et al. (1986)	X		
4	Arkes et al. (2007)			X
5	Ashton et al. (1994)	X		
6	Benbasat and Taylor (1978)			X
7	Benbasat & Taylor (1981)	X		
8	Brown and Jones (1998)			X
9	Brown and Vari (1992)	X	X	
10	Brown (2015)			X
11	Carey and Kacmar (2003)			X
12	Carrigan et al. (2004)		X	
13	Choragwicka and Janta (2008)			X
14	Christin (2017)			X
15	Colarelli and Thompson (2008)	X		
16	Dietvorst et al. (2015)	X		
17	Dietvorst et al. (2016)	X		
18	Eastwood et al. (2012)	X		
19	Eining et al. (1997)	X		
20	Er (1988)			X
21	Fisher (2008)	X		
22	Goodyear et al. (2016)			X
23	Goodyear et al. (2017)			X

(Continues)

TABLE 3 (Continued)

#	Reference	Judgement and decision making orientation[a]		
		Heuristics and biases	Fast and frugal	N/A
24	Green and Hughes (1986)			X
25	Hafenbrädl et al. (2016)		X	
26	Highhouse (2008a)	X		
27	Highhouse (2008b)	X		
28	Huber (1992)			X
29	Jarrahi (2018)			X
30	Kahn and Baron (1995)	X		
31	Klimoski and Jones (2008)			X
32	Kuncel (2008)	X		
33	Lamberti and Wallace (1990)			X
34	Landsbergen et al. (1997)			X
35	Lim and O'Connor (1996)	X		
36	Lodato et al. (2011)	X		
37	Mackay and Elam (1992)			X
38	Martin (2008)			X
39	Moldafsky and Kwon (1994)	X		X
40	Montazemi (1991)		X	
41	Mosier & Fischer (2010)		X	
42	Muir (1987)	X		
43	Mullins and Rogers (2008)			X
44	O'Brien (2008)	X		
45	Önkal et al. (2009)			X
46	Pagano et al. (2016)	X		
47	Patterson (2017)		X	
48	Phillips & Gully (2008)	X		
49	Prahl and Van Swol (2017)			X
50	Rao et al. (1992)			X
51	Robey (1992)			X
52	Robey and Taggart (1982)		X	
53	Sage (1981)		X	
54	Sanders and Courtney (1985)			X
55	Scherer et al. (2015)	X		
56	Sieck and Arkes (2005)	X	X	
57	Sutherland et al. (2016)			X
58	Swinney (1999)	X		
59	Thayer (2008)			X
60	Westin et al. (2016)		X	
61	Whitecotton (1996)			X

[a]Judgement and decision making orientations were classified based on the article's descriptions of decision making processes (i.e. optimization or satisficing), the assumed statistical structure of the decision making environment (i.e. risk or uncertainty), and the referenced theories and authors (i.e. Kahneman/Tversky or Gigerenzer). Those articles with the "N/A" classification did not explicitly mention JDM theory, nor did they cite relevant authors.

TABLE 4 Results summary

Theme	References		Cause of algorithm aversion	Solution for algorithm aversion
Expectations and expertise	Alexander et al. (2018)	Mackay and Elam (1992)	False expectations: A human decision maker's proclivity to utilize an algorithmic aid is influenced by that decision maker's past experiences and expectations for how the algorithm should perform; algorithms and humans are held to different standards.	Algorithmic literacy: Human decision makers need training with algorithmic aids in addition to training in their professional decision domain. Core statistics concepts like error and uncertainty cannot be overlooked.
	Arkes et al. (1986)	Martin (2008)		
	Ashton et al. (1994)	Montazemi (1991)		
	Carey and Kacmar (2003)	Muir (1987)		
	Christin (2017)	Önkal et al. (2009)		
	Dietvorst et al. (2015, 2016)	Pagano et al. (2016)		
	Goodyear et al. (2016, 2017)	Prahl and Van Swol (2017)		
	Green and Hughes (1986)	Robey (1992)		
	Highhouse (2008a, 2008b)	Sutherland et al. (2016)		
	Kahn & Baron (1995)	Swinney (1999)		
	Lamberti and Wallace (1990)	Thayer (2008)		
	Landsbergen et al. (1997)	Whitecotton (1996)		
	Lodato et al. (2011)			
Decision autonomy	Ashton et al. (1994)		Lack of decision control: Human decision makers need an opportunity to interact with an algorithmic aid and provide input in order to calibrate their confidence in the aid.	Human-in-the-loop decision making: Algorithmic aids should be designed in a way that allows human decision makers to semisupervise their processes; increased transparency and participation.
	Brown and Jones (1998)			
	Carrigan et al. (2004)			
	Colarelli and Thompson (2008)			
	Dietvorst et al. (2015, 2016)			
	Eining et al. (1997)			
	Highhouse (2008a)			
	Landsbergen et al. (1997)			
	Lim and O'Connor (1996)			
	Muir (1987)			
	Scherer et al. (2015)			
	Whitecotton (1996)			

(Continues)

TABLE 4 (Continued)

Theme	References		Cause of algorithm aversion	Solution for algorithm aversion
Incentivization	Alavi and Henderson (1981)	Highhouse (2008b, 2008a)	Lack of incentivization: Human decision makers require incentivization to put in extra effort into consulting algorithmic insights. Economic and social incentives need to be framed in a way that suit the decision making environment's conditions.	Behavioral design: Context-specific nudges and boosts to remedy human decision makers' motivational deficiencies; aim to systematically change behaviors and social norms.
	Alexander et al. (2018)	Klimoski and Jones (2008)		
	Arkes et al. (1986)	Kuncel (2008)		
	Arkes et al. (2007)	Lodato et al. (2011)		
	Brown and Jones (1998)	Önkal et al. (2009)		
	Brown (2015)	Sanders and Courtney (1985)		
	Choragwicka and Janta (2008)	Sutherland et al. (2016)		
	Christin (2017)	Swinney (1999)		
	Eastwood et al. (2012)			
	Fisher (2008)			
	Hafenbrädl et al. (2016)			
Cognitive compatibility	Alavi and Henderson (1981)	Lamberti and Wallace (1990)	Combatting intuition: Humans vary in their decision strategies. Inflexible aids built on formal decision rules attempt to override natural processes in favor of normative decision making, which creates cognitive conflict between human and algorithm.	Engaging intuition: Modeling algorithms on intuitive decision processes to permit cognitive compatibility in problem solving; human and algorithmic decision processes must overlap, not run parallel to one another.
	Arkes et al. (1986)	Landsbergen et al. (1997)		
	Benbasat & Taylor (1978, 1981)	Lim and O'Connor (1996)		
	Brown & Jones (1998)	Moldafsky and Kwon (1994)		
	Brown (2015)	Mullins and Rogers (2008)		
	Carey and Kacmar (2003)	Rao et al. (1992)		
	Carrigan et al. (2004)	Robey (1992)		
	Christin (2017)	Robey and Taggart (1982)		
	Dietvorst et al. (2015)	Sage (1981)		
	Eastwood et al. (2012)	Sieck and Arkes (2005)		
	Eining et al. (1997)	Swinney (1999)		
	Er (1988)	Patterson (2017)		
	Green and Hughes (1986)	Thayer (2008)		
	Huber (1992)	Whitecotton (1996)		
	Jarrahi (2018)			

(Continues)

TABLE 4 (Continued)

Theme	References		Cause of algorithm aversion	Solution for algorithm aversion
Divergent rationalities	Benbasat and Taylor (1981)	Sage (1981)	Conflicting views of rationality: Humans and algorithms extract different information and statistical structures from identical environments; they differ in their definition of a good decision.	Aiding ecological rationality: Designing algorithmic decision aids for ecological rationality so algorithms are able to incorporate the perspective of a human user under real-world uncertainty.
	Brown and Vari (1992)	Sanders and Courtney (1985)		
	Brown (2015)	Sieck and Arkes (2005)		
	Er (1988)	Westin et al. (2016)		
	Fisher (2008)			
	Hafenbrädl et al. (2016)			
	Jarrahi (2018)			
	Kahn and Baron (1995)			
	Mosier & Fischer (2010)			
	O'Brien (2008)			
	Pagano et al. (2016)			
	Patterson (2017)			