Algorithmically Controlled Automated Decision-Making and Societal Acceptability: Does Algorithm Type Matter?

Camille Grange
HEC Montreal
camille.grange@hec.ca

Abstract

As technological capabilities expand, an increasing number of decision-making processes (e.g., rankings, selections, exclusions) are being delegated to computerized systems. In this paper, we examine the societal acceptability of a consequential decision-making system (university admission) to those subject to the decision (i.e., applicants). We analyze two key drivers: the nature of the decision-making agent (a human vs an algorithm) and the decision-making logic used by the agents (predetermined vs emerging). Consistent with uniqueness neglect theory, we propose that applicants will be more positive toward the use of human agents compared to computerized systems. Consistent with the theory of procedural justice, we further argue that applicants will find the use of a predetermined logic to be more acceptable than an emerging logic. We present the details and results of a factorial survey designed to test our theoretical model.

1. Introduction

The next-generation technological era will be marked by the need to cope with the prevalence of complex and contentious digital innovations. Among them are automation technologies, and more particularly algorithmically controlled, automated decision-making (ADM) systems. Such systems are increasingly being deployed to substitute human agents and automate a number of decisions in highly consequential domains (e.g., education, immigration, justice) and processes (e.g., the ranking, filtering, selecting of people) [1], [2].

Undoubtedly, the choices that individuals, organizations, and societies will make regarding the design and usage of these innovations will severely impact the kind of society we live in and will hand over to next generations. On that matter, there are no doubts that ADM can help make the world a better place. They are instrumental in making several organizational and personal processes be more consistent, efficient, and scalable. For example, they are often praised for their existing and potential contribution to pubic administration [3] and to healthcare (e.g., diagnosing diseases, selecting treatments, discovering remedies), including most recently in the context of supporting COVID-19 diagnoses based on X-ray images [4]. However, several concerns have surfaced regarding the socio-ethical risks accompanying their widespread deployment, such as privacy violations, unwarranted surveillance, uninformed control, unfair discrimination [5]–[7]. Considering this new reality, some efforts have been made to regulate ADM, such as article 22 of the General Data Protection Regulation (GDPR) and Article 29 Working Party's Guidelines on automated decision-making and profiling in Europe. However, these existing frameworks remain far from perfect at adequately addressing problems of opacity and discrimination related to machine learning processing and the explanations of automated decision-making [8].

The angle taken in this paper to study this phenomenon is that of societal acceptability, which involves a broader perspective on technology acceptance [9]. This perspective is not just focused on acceptability from the point of view of the direct user of a system. Instead, it considers other stakeholders, including individuals facing the consequences of putting the system to use. This stance is absent from the core information systems outlets, but it has been prevalent in other field such as energy and medicine [9].

Societal acceptability is a particularly relevant research gap in the context of the diffusion of automated decision-making systems. We have known for a while that the economic success of organizations depends on the ability of its people to make good decisions, but except a few studies published in information systems e.g., [10] as well as adjacent fields e.g., [11]–[13], the perceptions from those who are subject to these decisions (e.g., candidates, employees, customers, patients, constituents), which we refer to as decision recipients [14], have been much less of a focus of interest. Nonetheless, they are critical for the

HICSS

responsible diffusion of ADM systems and the reputation of the organizations deploying them.

In addition to this research gap, we bring attention to another limitation in the literature, that of the lack of distinction made between different types of algorithmically controlled automation. Several studies frame their study of automation around the concept of "artificial intelligence". For example, Longoni et al.'s study of patient resistance to artificial intelligence (AI) have defined AI as "any machine that uses any kind of algorithm or statistical model to perform perceptual, cognitive, and conversational functions typical of the human mind, such as visual and speech recognition, reasoning, and problem solving" [13] p. 630. The concepts of algorithms and AI can be fuzzy and mean different things to different communities [15], [16]. To be more specific, we distinguish two families of algorithmically controlled automation logic: (i) *rules-based* systems, in which the algorithmic logic is hardcoded by developers, leaving little flexibility once deployed, and (ii) *machine learning* systems, for which the logic "emerges" from a model that learns from being applied to a data set [17], [18].

Our research is thus led by the following question: *Do automation and automation logic matter to the societal acceptability of a decision-making process*? To explore this question, we theorize on the effects of two factors: 1) the agent in charge of the decision (a human vs. an algorithm), and 2) the logic underlying the decision-making process (following a predetermined vs. an emerging logic).

The rest of the paper presents the conceptual framing for this research, the hypotheses that we intend to test, the empirical design - a factorial experiment based on manipulated scenarios of a university admission process, and some key results and implications.

2. Theoretical background

2.1. Automated decision-making (ADM)

For better or for worse, our world is rife with automated processes, and it has been so for a few decades. However, the phenomenon and its consequences have started to generate serious concerns as of lately. The increasing availability of data and sophistication of algorithms (including the rebirth of machine learning / neural networks [19]) has enabled more uses and misuses of algorithmically controlled, automated decision-making (ADM, for short) [15]. The

scaling of such innovations has happened in a context where algorithms, which still bear little accountability [8], [20], are involved in automating news recommendations [21], [22], advertising [23], but also work [24], [25][26] and other highly sensitive processes such as social ranking, crime prediction, and bail, parole and criminal sentencing [27]–[29] .

Experts and researchers have offered different definitions for ADM. For example, the European Commission describes it as "decisions by technological means without human involvement"[1]. The Canadian government, in its Directive on automated decision-making, defines automated decision systems as "any technology that either assists or replaces the judgement of human decision-makers. These systems draw from fields like statistics, linguistics, and computer science, and use techniques such as rules-based systems, regression, predictive analytics, machine learning, deep learning, and neural nets."[2] For Newell and Marabi, algorithmic decision-making occurs when "data are collected through digitized devices carried by individuals such as smartphones and technologies with inbuilt sensors – and subsequently processed by algorithms, which are then used to make (data-driven) decisions" [26] pp. 4.

In other words, the several definitions of ADM indicate that there are several shades of grey in the world of automated decision-making – from algorithms *supporting* and *augmenting* human agents' decision-making to humans fully *delegating* the task to algorithms [11], [15], [30], [31]. This is in line with Rai et al.'s view that as digital technologies continue to penetrate our work and daily lives, we are observing the design and deployment of a variety of human-AI hybrids that range in human-AI interdependence from substitution (AI replaces humans) to augmentation (humans and AI augment one and another) and assemblage (AI and humans are dynamically brought together to function as an integrated unit) [32]. This contemporary take on ADM relates to the distinction made in the early literature on expert systems (ES) and decision-support systems (DSS). A DSS was considered to be *helping* people (e.g., managers, consumers) make decisions (often characterized as ill-structured problems) [33]. In turn, an ES was viewed as a computer program (often based on AI techniques) that *mimicked* experts at making complex, non-algorithmic decisions [34], [35].

In general, ADM tends to promote technological autonomy at the expense of human agency. As we explain next, ADM systems are often, but not always,

---

[1] Data protection working party. Directive 95/46/EC § Articles 29 and 30

[2] https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592#appA

supported by sophisticated machine-learning based artificial intelligence [36].

2.2. AI, algorithms, and algorithmic transparency

We purposefully distinguish between ADM and AI because the latter remains a fuzzy-defined term that encompasses ideas larger than ADM [15]. For example, according to the European Commission, AI refers to "systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g., voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g., advanced robots, autonomous cars, drones or Internet of Things applications)." (European Commission, p. 10).

ADM systems process information in the form of input data using an algorithm to generate an output of some kind. According to Gillespie [16], an algorithm is "a recipe composed in programmable steps," (p. 19) designed for the purpose of "organizing and acting on a body of data to quickly achieve a desired outcome." (p. 19). Simply put, it is a set of instructions.

Importantly, algorithms vary in the extent to which these instructions are made transparent. The notion and demand for algorithmic transparency is not new. A first example of legal regulation of algorithmically controlled, automated decision-making in the US dates back to 1984 when it became required that the criteria used in the algorithms be provided (to any person upon request) [15]. More recently, the European Union AI Strategy states that "in order to increase transparency and minimise the risk of bias or error, AI systems should be developed in a manner which allows humans to understand (the basis of) their actions" (pp. 10)[3]. The need for algorithmic transparency has been raised in various domains of application including the news media [37]. In that domain, the idea of algorithmic transparency is to "publicly show and reveal the workings of computational systems in order for users to discern the embedded values of specific algorithms" [38] pp. 253.

The issue of algorithmic transparency becomes particularly salient in the context of machine learning algorithms, which are known to be more prone to opacity [39]. Modern algorithms rely on technique such as machine learning, meaning that they are "trained" using a large, existing corpus of data, which allows the algorithm to classify and "generalize beyond the examples in the training set." [40] pp. 80. This differs from "manual programing" as Domingos put it, or rules-based programing, in which the instructions do not emerge from data but are predetermined by those developing the algorithms.

2.3. Psychology of automation

A stream of literature is fast developing about people's responses to the growing role of automation and algorithms.

One stream of thought contends that machines are supposed to and perceived to be more systematic and freer of ideological bias, leading to a preference toward delegating decisions to machines instead of putting human agents in charge [41]. Consistent with this idea, Logg et al. found that participants who faced "black box" algorithms in their experiments were still willing to rely on that advice despite its mysterious origins [12]. Two other studies also reported evidence of such leaning, commonly referred to as algorithmic appreciation [11], [42]. In Araujo et al. [11], the decisions taken automatically by an AI agent were often evaluated on par or even better than in the case of human experts for specific decisions.

In contrast , earlier research argued that humans tend to irrationally dismiss the superior accuracy of statistical model over human intuition and to distrust algorithms [43]. The underlying idea of humans being preferred to computerized agents for advice or recommendation has been corroborated in several situations. For example, people tend to prefer friends over algorithmic recommender for preferential choice problems (e.g., finding a book or a movie) [44]. Patients are more reluctant to utilize healthcare provided by AI than a human provider, in part by fear of uniqueness neglect, that is, a concern that AI providers are less able than human providers to account for consumers' unique characteristics and circumstances [13]. People also tend to hate being managed by algorithms [45]. In this context of work and automated managerial decisions, human responses have been shown to be contingent on the task being automated. Lee observed that with mechanical tasks, algorithmic and human-made decisions were perceived as equally fair and trustworthy and evoked similar emotions. However, with human tasks (when human judgment is needed), algorithmic decisions were perceived as less fair and trustworthy and evoked more negative emotion than human decisions [46].

---

[3] https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe

Overall, existing findings indicate that there is no clear and definite pattern when it comes to assessing people's response to automation. Human beings tend to oscillate between algorithmic appreciation [12] and algorithm aversion [47] depending on contextual factors such as the expertise of the human agent and the task being automated.

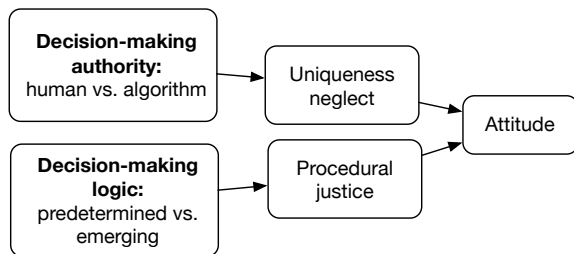## 3. Hypotheses development

Our research model is depicted in Figure 1.

Figure 1. Research Model

We conceptualize societal acceptability through decision recipients' **attitude** toward the decision-making process. We examine attitude because it refers to an overall evaluative judgment of a psychological object [48], and it has been considered a key facet of acceptability [49]. Our hypotheses rely on two theoretical mechanisms regarding the effects of automation (i.e., decision-making authority) and that of the underlying decision-making logic on acceptability.

The first mechanism, based on people's fear of uniqueness neglect, predicts that decision recipients will be less receptive to automated decision-making because of a concern that their unique characteristics, circumstances, and symptoms are more likely to be neglected by a machine than by a human (i.e., *uniqueness neglect*). Several theories in psychology (e.g., self-identity theory, social comparison theory [50][51]) stress the idea that individuals feel idiosyncratic, that is, unique and different from others. However, individuals realize that computerized systems tend to (over) standardize and dehumanize processes [52], [53], preventing the ability to consider and value their uniqueness [54]. Based on this, we propose the following hypotheses associated with the effects of decision-making authority:

H1a. Decision recipients experience higher feelings of human neglect when the decision-making authority is an algorithm.

H1b. The feeling of uniqueness neglect has a negative effect on decision recipients' attitude toward the decision-making process.

The second mechanism is that decision recipients will be less receptive (i.e., acceptability will be lower) when the decision-making logic generates inferences made on the basis of *emerging* rules because of a lack of perceived fairness in the evaluation procedure (*procedural justice*). This logic relies on the idea that predetermined, rules-based inferences are more effective at enabling algorithmic transparency than inferences that are 'adaptive', i.e., that emerge based on historical data, because the logic is explicitly coded in the algorithm. According to procedural justice, a decision process is perceived to be fair if it meets procedural justice criteria like consistency, lack of bias, and correctability [55], [56]. Such characteristics are more likely to be perceived and salient when individuals are subject to decisions that are the result of a standardized process whose evaluation criteria are clear and observable, which fits better with a rules-based decision-making logic. Therefore, we propose the following hypotheses:

H2a. Decision recipients perceive a lower degree of procedural justice when the decision-making logic is emerging.

H2b. Procedural justice has a positive effect on decision recipients' attitude toward the decision-making process.

4. Methodology

To test our hypotheses, we employed a factorial survey method, which applies an experimental approach to a scenario-based survey [57], [58]. The scenario part of this method enables the development of rich contextual situations while the experimental part enables the manipulations of these situations to explore dimensions of theoretical interest. The factorial survey method is frequently used to study human beliefs and judgments [59], and it is starting to gain prominence in information systems research e.g., [60], [61]. In this study, we have chosen a full-factorial design, where two factors are manipulated over two dimensions (*decision-making authority*: human vs. automation; *decision-making logic*: predetermined vs. emerging) yielding four treatments groups in total.

4.1. Scenario design

We developed our scenarios in the context of university admission, showcasing a process in which the agent responsible for analyzing and selecting candidates and the process used to conduct this task varied. We selected university admission because education is a high-stakes domain in which automation is appealing for efficiency reasons but is also likely to generate ambivalent or contentious opinions [62], [63].

The scenarios were articulated as follow. The first part remained the same across conditions:

"Imagine that you are applying to a graduate program at a university. You are requested to fill an online form and to upload complementary documentation (including your grades as well as letters of intent and recommendation). This information will be used by the university to make admission offers. More specifically, the program admission website specifies that:"

The second part was treatment-specific:

[T1: algorithm x predetermined] "Candidates are ranked by a **computerized system** which relies on **quality criteria and standards specified in the publicly available school's admission**"

[T2: algorithm x emerging] "Candidates are ranked by a **computerized system,** which relies on **historical admission data** to predict the likeliness of whether a candidate will successfully complete the program"

[T3: human x predetermined] "Candidates are ranked by the **program's academic advisor** who relies on **quality criteria and standards specified in the publicly available school's admission**"

[T4: human x emerging] "All files are processed by a **faculty member** who relies on their **experiential knowledge** to predict the likeliness of whether a potential applicant will successfully complete the program."

4.2. Study sampling and procedure

The data comes from a sample of 18 years old or older North American adults recruited via the services of a well-established online market research company (Qualtrics). Given the study context, recruited participants needed to have been exposed to a university admission process in the past. Thus, only those who were either currently enrolled as undergraduate students of a four-year degree or had recently graduated from such program (less than 2 years ago) were eligible. We assembled a sample of N=300 complete responses, split just about evenly between the four treatment groups (T1: N=78; T2: N=73; T3: N=74; T4: N=75).

Study participants were informed that our research intended to improve our understanding of student preferences regarding university admission. Participants first answered demographics questions (age, gender, level of education). On the survey's following page, they were randomly assigned to one of the four treatments presenting the associated scenario. We urged participants to read the scenario presented to them very carefully. Moving to the following pages, participants answered questions intended to capture the research model's variables. We also included realism and

attention check questions [64], and incorporated two manipulation check questions (7-pt Likert scale): "in the scenario that was exposed to you…:"

(i) "*A human assessed the quality of the candidates' applications*" (participants in treatments 3 and 4 would need to report high scores on this question for us to consider the manipulation effective while those in treatments 1 and 2 would need to report low scores),

(ii) "*The analysis relied only on criteria and standards specified in the school's admission policy*" (participants in treatments 1 and 3 would need to report high scores on this question for us to consider the manipulation effective while those in treatments 2 and 4 would need to report low scores).

4.3. Measurements

Most constructs were measured using multi-item 7-point Likert-type scales ranging from strongly disagree to strongly agree. Measurement items (Table 1) were adapted from scales validated in earlier studies.

Table 1. Measurement scales

Uniqueness neglect: a concern that one's unique characteristics, circumstances, and symptoms will be neglected (Longoni et al. 2019 [13])
To which extent would you be worried that the faculty member responsible for processing your application
1. Would not recognize the uniqueness of your application 2. Would not consider your unique circumstances 3. Would not tailor the criteria of admission to your unique case
Procedural justice: the perceived fairness of the decision-making process (Enosken 2015 [56]; Otting and Maier 2018 [65])
To what extent do you think the <u>evaluation process</u> described in the scenario
1. Leads to admission decisions that are free of bias 2. Relies on accurate information from the candidates 3. Enables consistent evaluations of candidates' application 4. Leads to decisions that can be explained to candidates 5. Enables a transparent assessment of candidates 6. Is amenable to correction if mistakes are made
Attitude: the overall degree of favourability toward the decision-making process (Ajzen 2001 [48])
In terms of its value to candidates and society in general, I would say that the admission process described in this scenario is (7-pt scale):
1. Bad . . . . Good

2. Foolish . . . . Wise
3. Irresponsible . . . . Responsible
4. Ineffective . . . . . . Effective
5. Harmful . . . . . . Beneficial
6. Immoral . . . . . . Moral
7. Detrimental . . . . Desirable

5. Preliminary insights

We first ran some preliminary validity tests. The Cronbach alphas values were all within an acceptable range: they were 0.80, 0.78, and 0.93 for the uniqueness neglect, procedural justice, and attitude constructs, respectively. The one-way ANOVAs ran on our two manipulation check questions (Table 2) confirmed the expected significant differences between treatment groups: [$F(3, 294) = 58.41$, $p< 0.001$] for MCI and [$F(3, 294) = 12.2$, $p < 0.001$] for MC2. Post-hoc comparisons using the Tukey HSD test indicated that the mean score of MC1 for treatments 1 and 2 was significantly different than for treatments 3 and 4, and that the score of MC2 for treatments 1 and 3 was significantly different than for treatments 2 and 4 (Table 2). Those results suggest that our manipulations (scenarios) functioned as expected.

Table 2. Manipulation check results

Treat. groups	MC1. *A human assessed the quality of the candidates' applications**	MC2. *The analysis relied only on criteria and standards specified in the school's admission policy***
T1	2.63	**5.65**
T2	2.89	4.71
T3	**5.23**	**5.64**
T4	**5.56**	4.29

Bold: expected higher values.

Next, we ran some preliminary hypothesis tests. Two two-way ANOVAs were conducted to compare the main effects of decision-making authority and decision-making logic on uniqueness neglect (H1a) and procedural justice (H2a). In the first test (**H1a**), the main effect of decision-making authority yielded an F ratio of $F(1,296) = 11.63$, $p<0.001$), indicating a significant difference between the *algorithm* (M=5.20, SD=1.26) and *human* authority (M=4.70, SD=1.29) conditions. The results indicate a positive effect, in line with what we had hypothesized: respondents in the algorithmic decision-making condition reported higher level of uniqueness neglect than those in the human decision-making authority condition. Thus, H1a was supported. In addition, note that decision-making logic and the interaction of decision-making logic and decision-making authority had no significant effects on uniqueness neglect.

In the second test (**H2a**), the main effect of decision-making logic on procedural justice yielded an F ratio of $F(1,296) = 1.791$, $p=0.18$), indicating a non-significant difference between a *predetermined* (M=4.20, SD=1.15) and *emerging* logic (M=4.02, SD=1.19). Thus, H2a was rejected. Note that decision-making authority and the interaction of decision-making logic and decision-making authority had no significant effects on procedural justice either.

We tested H1b and H2b using simple linear regressions. The results indicate that uniqueness neglect is a significant predictor of attitude, $F(1,298) = 40.9$, $p < 0.001$ with an R-square of 11.77%. Attitude decreases by 0.37 for each point gained on the uniqueness neglect scale. This result is in line with what we had hypothesized in **H1b** (a negative effect of uniqueness neglect on attitude). Running a similar analysis with procedural justice as the independent variable yielded an F ratio of $F(1,298) = 173.9$, $p < 0.001$ with an R-square of 36.64%. The results indicate that attitude increased by 0.71 for each point on the procedural justice scale, which is in line with what we hypothesized (a positive effect of procedural justice on attitude, **H2b**).

In **summary**, we found strong significant effects of both uniqueness neglect and procedural justice on the societal acceptability of decision-making procedures, conceptualized through decision recipients' attitude toward the decision-making procedure. We also observed that automation mattered because algorithmic authority generated significantly higher feelings of uniqueness neglect compared to when a human was in charge. Finally, we did not find any evidence that algorithm type matters to acceptability, a finding that goes against what we had expected based on the body of work highlighting the value that people (including decision recipients) place in fair and transparent decision rules and processes.

6. Limitations and next steps

This study has limitations that need to be acknowledged. First, in retrospect, we have realized that the sampling approach has excluded people who went through the process of applying to university but ended up being unsuccessful (that is, their applications were denied). These people might have had a systematically different experience than that of those who were successful in the process and who constituted the totality of our sample. Second, it has been a challenge operationalizing the emerging logic treatments, both for a human and an algorithmic authority, and this may contribute to explain the lack of effect on procedural justice that we observed. We first intended to mention

the term "machine learning algorithm" in the scenarios but we reasoned that it makes more sense to expose potential decision recipients to key characteristic of an emerging (algorithmic) logic instead of just referring to machine learning (which may mean many different things to different people). This led us to focus operationalization of emerging logic on the idea of an algorithmic logic relying on historical data as a predictor of admission success, but we are aware that it is incomplete and that other instantiations are possible.

Aside from method, it is possible that our research model put too strong of an emphasis on the expected value of algorithmic transparency (thus procedural justice). Recent research has noted that complete transparency may not be fully feasible or desirable and suggests that accountability may be a stronger driving force of acceptability e.g., [66], [67], an avenue that future work would need to explore further.

7. Conclusion

Our study explores the social acceptability of algorithmically controlled automated decision making with a unique emphasis on the effect of different evaluation procedures (e.g., algorithms based on a predetermined logic vs algorithms based on an emerging or 'adaptive' logic. Given the increased use of opaque adaptive algorithms (often referred to as machine learning) to automate decisions in highly consequential domains, it is critical to account for the perceptions of those subject to the decisions delegated to algorithms. The societal acceptability lens adopted in this paper is in line with this position and contributes to the advancement of a more responsible diffusion process of ADM.

8. References

[1]   P. Molnar and L. Gill, "Bots at the gate: A human rights analysis of automated decision-making in Canada's immigration and refugee system," 2018.

[2]   C. O'Neil, *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown Books, 2016.

[3]   B. W. Wirtz, J. C. Weyerer, and C. Geyer, "Artificial Intelligence and the Public Sector— Applications and Challenges," *International Journal of Public Administration*, vol. 42, no. 7, pp. 596–615, May 2019, doi: 10.1080/01900692.2018.1498103.

[4]   G. Marques, D. Agarwal, and I. de la Torre Díez, "Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network," *Applied Soft Computing*, vol. 96, p. 106691, Nov. 2020, doi: 10.1016/j.asoc.2020.106691.

[5]   A. G. Ferguson, *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: NYU Press, 2017.

[6]   D. Lyon, *Surveillance as social sorting: Privacy, risk,and digital discrimination*. Routledge, 2003.

[7]   K. Macnish, "Unblinking eyes: the ethics of automating surveillance," *Ethics and information technology*, vol. 14, no. 2, pp. 151–167, 2012.

[8]   C. Castets-Renard, "Accountability of Algorithms in the GDPR and Beyond: A European Legal Framework on Automated Decision-Making," *Fordham Intellectual Property, Media and Entertainment Law Journal*, vol. 30, no. 1, p. 91, Jan. 2019.

[9]   G. Tabourdeau and C. Grange, "From User Acceptance to Social Acceptance," presented at the International Conference on Information Systems (ICIS), Hyderabad, India, Dec. 2020. [Online]. Available: https://aisel.aisnet.org/sighci2020/10

[10]   J. Ochmann and S. Laumer, "Fairness as a determinant of AI Adoption in Recruiting: An interview-based study," presented at the ICIS, Munich, 2019. Accessed: Jan. 29, 2020. [Online]. Available: https://www.researchgate.net/publication/338646040_Fairness_as_a_determinant_of_AI_Adoption_in_Recruiting_An_interview-based_study

[11]   T. Araujo, N. Helberger, S. Kruikemeier, and C. H. de Vreese, "In AI we trust? Perceptions about automated decision-making by artificial intelligence," *AI & Soc*, vol. 35, no. 3, pp. 611–623, Sep. 2020, doi: 10.1007/s00146-019-00931-w.

[12]   J. M. Logg, J. A. Minson, and D. A. Moore, "Algorithm appreciation: People prefer algorithmic to human judgment," *Organizational Behavior and Human Decision Processes*, vol. 151, pp. 90–103, 2019.

[13]   C. Longoni, A. Bonezzi, and C. K. Morewedge, "Resistance to Medical Artificial Intelligence," *J Consum Res*, vol. 46, no. 4, pp. 629–650, Dec. 2019, doi: 10.1093/jcr/ucz013.

[14]   L. Heuer, S. Penrod, and A. Kattan, "The role of societal benefits and fairness concerns among decision makers and decision recipients," *Law and Human Behavior*, vol. 31, no. 6, pp. 573–610, 2007.

[15]   M. Spielkamp, "Automating society - Taking stock of automated decision-making in the EU." AlgorithmWatch, Jan. 2019. [Online]. Available:

https://algorithmwatch.org/de/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf

[16] T. Gillespie, *Digital Keywords - A Vocabulary of Information Society and Cultyre*. Princeton: Princeton University Press, 2016. Accessed: Jun. 14, 2021. [Online]. Available: algorithmic tran

[17] K. Kane, A. Young, A. Majchrzak, and S. Ransbotham, "Avoiding an Oppressive Future of Machine Learning: A Design Theory for Emancipatory Assistants," *Management Information Systems Quarterly*, vol. 45, no. 1, pp. 371–396, Mar. 2021.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[19] O. Temam, "The rebirth of neural networks," 2010.

[20] K. Martin, "Ethical Implications and Accountability of Algorithms," *J Bus Ethics*, Jun. 2018, doi: 10.1007/s10551-018-3921-3.

[21] B. Bodó, "Selling News to Audiences – A Qualitative Inquiry into the Emerging Logics of Algorithmic News Personalization in European Quality News Media," *Digital Journalism*, vol. 7, no. 8, pp. 1054–1075, Sep. 2019, doi: 10.1080/21670811.2019.1624185.

[22] N. Thurman and S. Schifferes, "The Future of Personalization at News Websites," *Journalism Studies*, vol. 13, no. 5–6, pp. 775–790, Oct. 2012, doi: 10.1080/1461670X.2012.664341.

[23] S. C. Boerman, S. Kruikemeier, and F. J. Z. Borgesius, "Online Behavioral Advertising: A Literature Review and Research Agenda," *Journal of Advertising*, vol. 46, no. 3, pp. 363–376, Jul. 2017, doi: 10.1080/00913367.2017.1339368.

[24] N. Jabagi, A.-M. Croteau, and L. Audebrand, "Perceived Organizational Support in the Face of Algorithmic Management: A Conceptual Model," Jan. 2020. doi: 10.24251/HICSS.2020.489.

[25] M. Möhlmann, L. Zalmanson, O. Henfridsson, and R. Gregory, "Algorithmic Management of Work on Online Labour Platforms: when matching Meets Control," *MIS Quarterly*, Jan. 2021.

[26] A. Rosenblat, *Uberland: How algorithms are rewriting the rules of work*. Univ of California Press, 2018.

[27] D. Ensign, S. A. Friedler, S. Neville, C. Scheidegger, and S. Venkatasubramanian, "Runaway Feedback Loops in Predictive Policing," *arXiv:1706.09847 [cs, stat]*, Jun. 2017, Accessed: Mar. 03, 2019. [Online]. Available: http://arxiv.org/abs/1706.09847

[28] D. Kehl, P. Guo, and S. Kessler, "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing," *Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School.*, 2017, Accessed: Feb. 21, 2019. [Online]. Available: https://dash.harvard.edu/handle/1/33746041

[29] S. Newell and M. Marabelli, "Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification,'" *The Journal of Strategic Information Systems*, vol. 24, no. 1, pp. 3–14, Mar. 2015, doi: 10.1016/j.jsis.2015.02.001.

[30] S. Basu, A. Garimella, W. Han, and A. Dennis, "Human Decision Making in AI Augmented Systems: Evidence from the Initial Coin Offering Market," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021, p. 176.

[31] S. Raisch and S. Krakowski, "Artificial intelligence and management: The automation–augmentation paradox," *Academy of Management Review*, vol. 46, no. 1, pp. 192–210, 2021.

[32] A. Rai, P. Constantinides, and S. Sarker, "Next-Generation Digital Platforms: Toward Human–AI Hybrids," *MIS Quarterly*, vol. 43, no. 1, pp. iii–ix, Mar. 2019.

[33] B. Hosack, D. Hall, D. Paradice, and J. F. Courtney, "A look toward the future: decision support systems research is alive and well," *Journal of the Association for Information Systems*, vol. 13, no. 5, p. 3, 2012.

[34] F. Nelson Ford, "Decision support systems and expert systems: A comparison," *Information & Management*, vol. 8, no. 1, pp. 21–26, Jan. 1985, doi: 10.1016/0378-7206(85)90066-7.

[35] J. J. Sviokla, "An Examination of the Impact of Expert Systems on the Firm: The Case of XCON," *MIS Quarterly*, vol. 14, no. 2, pp. 127–140, Jun. 1990, doi: 10.2307/248770.

[36] A. Agrawal, J. Gans, and A. Goldfarb, *Prediction Machines: The simple economics of artificial intelligence*. Harvard Business Press, 2018.

[37] N. Diakopoulos and M. Koliska, "Algorithmic Transparency in the News Media," *Digital Journalism*, vol. 5, no. 7, pp. 809–828, Aug. 2017, doi: 10.1080/21670811.2016.1208053.

[38] M. Koliska and N. Diakopoulos, "Disclose, Decode, and Demistify - An empirical Gide to

Algorithmic Transparency," in *The Routledge Handbook of Developments in Digital Journalism Studies*, S. Eldridge and B. Franklin, Eds. Routledge, 2018, p. 564.

[39]   J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data & Society*, vol. 3, no. 1, p. 2053951715622512, Jun. 2016, doi: 10.1177/2053951715622512.

[40]   P. Domingos, "A few useful things to know about machine learning," *Commun. ACM*, vol. 55, no. 10, pp. 78–87, Oct. 2012, doi: 10.1145/2347736.2347755.

[41]   S. S. Sundar, "The MAIN model: A heuristic approach to understanding technology effects on credibility," in *Digital media, youth, and credibilit*, M. Metzger and A. Flanagin, Eds. The MIT Press, 2008, pp. 73–100.

[42]   N. Thurman, J. Moeller, N. Helberger, and D. Trilling, "My Friends, Editors, Algorithms, and I," *Digital Journalism*, vol. 7, no. 4, pp. 447–469, Apr. 2019, doi: 10.1080/21670811.2018.1493936.

[43]   P. E. Meehl, "Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.," 1954.

[44]   R. Sinha, and R. Sinha, and K. Swearingen, "Comparing Recommendations Made by Online Systems and Friends," 2001.

[45]   M. Möhlmann and O. Henfridsson, "What people hate about being managed by algorithms, according to a study of Uber drivers," *Harvard Business Review*, vol. 30, 2019.

[46]   M. K. Lee, "Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management," *Big Data & Society*, vol. 5, no. 1, p. 2053951718756684, Jan. 2018, doi: 10.1177/2053951718756684.

[47]   B. J. Dietvorst, J. P. Simmons, and C. Massey, "Algorithm aversion: People erroneously avoid algorithms after seeing them err.," *Journal of Experimental Psychology: General*, vol. 144, no. 1, p. 114, 2015.

[48]   I. Ajzen, "Nature and Operation of Attitudes," *Annual Review of Psychology*, vol. 52, no. 1, p. 27, 2001.

[49]   M. Sekhon, M. Cartwright, and J. J. Francis, "Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework," *BMC Health Serv Res*, vol. 17, no. 1, p. 88, 26 2017, doi: 10.1186/s12913-017-2031-8.

[50]   A. P. Buunk and F. X. Gibbons, "Social comparison: The end of a theory and the emergence of a field," *Organizational Behavior and Human Decision Processes*, vol. 102, no. 1, pp. 3–21, 2007, doi: 10.1016/j.obhdp.2006.09.007.

[51]   M. B. Brewer, "Optimal distinctiveness, social identity, and the self," *Handbook of self and identity*, pp. 480–491, 2003.

[52]   H. Nissenbaum and D. Walker, "Will computers dehumanize education? A grounded approach to values at risk," *Technology in Society*, vol. 20, no. 3, pp. 237–273, Aug. 1998, doi: 10.1016/S0160-791X(98)00011-6.

[53]   N. Haslam, "Dehumanization: An integrative review," *Personality and social psychology review*, vol. 10, no. 3, pp. 252–264, 2006.

[54]   C. R. Snyder and H. L. Fromkin, *Uniqueness: The human pursuit of difference*. Springer Science & Business Media, 2012.

[55]   J. A. Colquitt, D. E. Conlon, M. J. Wesson, C. O. Porter, and K. Y. Ng, "Justice at the millennium: a meta-analytic review of 25 years of organizational justice research.," *Journal of applied psychology*, vol. 86, no. 3, p. 425, 2001.

[56]   E. Enoksen, "Examining the dimensionality of Colquitt's Organizational Justice Scale in a public health sector context," *Psychological reports*, vol. 116, no. 3, pp. 723–737, 2015.

[57]   L. Wallander, "25 years of factorial surveys in sociology: A review," *Social Science Research*, vol. 38, no. 3, pp. 505–520, Sep. 2009, doi: 10.1016/j.ssresearch.2009.03.004.

[58]   K. Auspurg and T. Hinz, *Factorial survey experiments*, vol. 175. SAGE, 2015.

[59]   G. Jasso, "Factorial Survey Methods for Studying Beliefs and Judgments," *Sociological Methods & Research*, vol. 34, no. 3, pp. 334–423, Feb. 2006, doi: 10.1177/0049124105283121.

[60]   S. Al-Natour, H. Cavusoglu, I. Benbasat, and U. Aleem, "An Empirical Investigation of the Antecedents and Consequences of Privacy Uncertainty in the Context of Mobile Apps," *Information Systems Research*, vol. 31, no. 4, pp. 1037–1063, Oct. 2020, doi: 10.1287/isre.2020.0931.

[61]   A. Vance, P. B. Lowry, and D. L. Eggett, "Increasing accountability through the user interface design artifacts: A new approach to addressing the problem of access-policy violations," *Mis Quarterly*, vol. 39, no. 2, pp. 345–366, 2015.

[62]   B. Williamson, "The Automatic University A review of datafication and automation in higher education," Universtity and College Union Scotland, Jun. 2020.

[63] D. Sweeney, "Automation And The Challenge For Colleges And Universities," *Forbes*, 2021. Accessed: Jun. 15, 2021. [Online]. Available: https://www.forbes.com/sites/forbesbusinesscouncil/2020/01/07/automation-and-the-challenge-for-colleges-and-universities/

[64] H. Shamon and C. Berning, "Attention Check Items and Instructions in Online Surveys: Boon or Bane for Data Quality?," *Survey Research Methods, Forthcoming*, 2019.

[65] S. K. Ötting and G. W. Maier, "The importance of procedural justice in human–machine interactions: Intelligent systems as new decision agents in organizations," *Computers in Human Behavior*, vol. 89, pp. 27–39, 2018.

[66] M. Ananny and K. Crawford, "Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability:," *New Media & Society*, Dec. 2016, doi: 10.1177/1461444816676645.

[67] J. A. Kroll, "Outlining Traceability: A Principle for Operationalizing Accountability in Computing Systems," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 2021, pp. 758–771.