HEC MONTRÉAL

Decision Recipients' Perspective on the Use and Acceptance of Automated Decision-Making Systems

By
Yinghui Gao

Management Sciences

(User Experience in a Business Context)

Thesis submitted in partial fulfilment of the

requirements for a Master of Science in Management

December 2022

© Yinghui Gao, 2022

Summary

This thesis focuses on the social aspects of automated decision-making systems. The European Commission (Bodea & Karanikolova, 2018) defines automated decision-making systems (ADMS) as decisions by technological means without human involvement, including various algorithms. Various technologies or algorithmic types, such as machine learning, expert system, and natural language processing, are employed to process decisions automatically or partially. These decisions involve various processes, from supporting human decision-makers to fully automated processes across contexts (Araujo et al., 2020).

Despite the socio-technical nature of ADMS and the risks (e.g., discrimination, loss of trust) they induce (Araujo et al., 2020; Kordzadeh & Ghasemaghaei, 2021; Marabelli et al., 2021), many of them are still implemented without a thorough analysis of their acceptability by decision recipients. This thesis addresses this overarching issue, taking a human-computer interaction (HCI) lens and examining decision recipients' responses to ADMS. The specific research objectives are threefold: to explore what is important to decision recipients in accepting ADMS, to investigate how decision recipients perceive algorithmic type, and to investigate how algorithmic type influences ADMS use.

To address the research objectives, coding and thematic analysis were performed on qualitative data from an online scenario-based experiment n=300 and interviews n=13. According to the results, participants discussed issues with human expert decision-making agents over algorithms. Themes brought up to explain their acceptability included whether algorithms could fulfill moral principles to render important judgements, the double standard between algorithms and human experts, and the favouritism towards outcome rather than the fairness of the process. In the interview, we discovered that participants perceive algorithmic types differently and mostly prefer machine learning algorithms over rule-based ones. When two algorithmic types were compared, participants distinguished themselves by their algorithmic characteristics' comprehensibility of decision-making process, adaptability to new information, robustness to protect itself from malicious behaviour, the quantity and quality of data input to make decisions, and bias in the decision-making process (historical bias and human expert bias). The participants' understanding of algorithmic characteristics influences their preference and behaviour with ADMS. When exposed to rule-based algorithms, participants will seek to game the system. However, participants will utilize the system as intended when exposed to machine learning. This thesis concludes that it is important to consider the decision-recipient's perspective in developing ADMS to prevent economic and social damage. From an individual's point of view, our results organize their perception towards ADMS, supporting informed decision-making.

Keywords: automated decision-making, algorithmic experience, perceived algorithmic fairness, transparency, algorithmic type, consequential decision, machine learning, rule-based algorithm, human bias.

HEC MONTRĒAL

Comité d'éthique de la recherche

CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains et qu'il satisfait aux exigences de notre politique en cette matière.

Projet # : 2023-4843

Titre du projet de recherche : Drivers Facilitating Acceptability in Automated Decision Making Systems (ADMs)

Chercheur principal :
Ying Hui Gao,
HEC Montréal

Directeur/codirecteurs :
Camille Grange
Professeur - HEC Montréal

Date d'approbation du projet : May 30, 2022

Date d'entrée en vigueur du certificat : May 30, 2022

Date d'échéance du certificat : May 01, 2023

Maurice Lemelin
Président
CER de HEC Montréal

Signé le 2022-05-31 à 16:29

Exporté le 2022-12-15 20:57 par Gao, Ying Hui — CODE DE VALIDATION NAGANO: hec-333c5f57-faba-44dd-ad04-ff28b40dfe69https://cer.hec.ca/verification/hec-333c5f57-faba-44dd-ad04-ff28b40dfe69

Table of Contents

List of tables

List of Figures

List of abbreviations

AI: Artificial intelligence

ADMS: Automated decision-making systems or algorithmic decision-making systems

AX: Algorithmic experience

FAccT: Fairness, accountability, and transparency

HCI: Human-computer interaction

ML: Machine learning (algorithmic type)

MLB ADMS: Machine learning based automated decision-making systems

RB: Rule-based (algorithmic type)

RB ADMS: Rule-based automated decision-making systems

Acknowledgements

I would like to start by thanking my partner for his support in my academic decision, which has been a long goal of mine. His moral support helped me get through the difficult moments.

In addition, I would like to thank Dr. Camille Grange, my thesis director, for single-handedly guiding and supervising the beginning and completion of my thesis. Dr. Grange assisted me in making tough life decisions, for which I am thankful.

Finally, I want to thank my family, specifically my mother, my sister, and my father, for their moral and financial support throughout my academic career.

Preface

My interest in automated decision-making systems (ADMS) stems from my frustration with applying and adopting an AI-based forecast and decision support system in my professional experience. This interest brought me to work with Dr. Grange and her current research program in the responsible adoption and diffusion of information technology. Amongst others, one of its applications is automated decision-making systems.

This study is motivated by an innovative growth of ADMS in various fields such as health, juridical, and business and an increase in a variety of opinions raised by society on the ethics of such implementations. These voices are essential because they have a role in adopting and implementing ADMS. This inevitably affects the system's success and the firms' return on investment.

Chapter 1: Introduction

1.1 Context

The advancement of algorithms is apparent in everything we do by reshaping traditional decision-making. Employing automated decision-making systems (ADMS) is becoming a preferred choice for businesses in fields ranging from health to recruitment and juridical for its efficiency. As ADMS begin to make a high level of consequential decisions, it alters people's thinking and behaviour by causing unexpected societal effects like prejudice and bias, which leads to a lack of trust and, accidentally, legal action and, eventually, termination of the ADMS. This is often strategically hidden from the decision recipient's experience to protect its intellectual property, prevent gaming the system, and provide a better experience (Burrell, 2016; Eslami et al., 2019). However, at other times, the decision-making process is involuntarily hidden from managers and engineers themselves for algorithms (e.g., machine learning and natural language processing) are labelled "opaque" for their black-box nature (Burrell, 2016; Marabelli et al., 2021).

ADMS are socio-technical artifacts rooted in several settings, connections, and societal functions (Araujo et al., 2020; Marabelli et al., 2021). It is critical to investigate the social aspects of ADMS to create a better product and to avoid unexpected consequences. As in the use of AI for a performance evaluation which lead to an unexplained job termination, causing employee frustration and legal disputes (Park et al., 2021), or the discontinuation of AI-based hiring due to unexpected gender discrimination ("Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women," 2018).

Social acceptability is concerned with the aspects that influence decision recipient's experience and the acceptance of its interactions (Koelle et al., 2019). Within the design cycle, constructing successful ADMS which are helpful to society necessitates paying attention to forecasting, influencing, and assessing alternatives for various elements of societal acceptability (Koelle et al., 2019). However, this conceptual framework receives little attention in the literature on responsible technology invention and spread (Grange, 2022; Tabourdeau & Grange, 2020).

One of the social aspects is within decision recipients. Even if algorithms are mathematically demonstrated to be fair, the decision's outcome or process may be seen as unjust since differing ideas of fairness exist among and between stakeholders (Delecraz et al., 2022; Lee et al., 2017). As a result, it is critical to take a human-centred approach to the repercussions. We have seen minimal work in the human-computer interaction (HCI) literature addressing algorithmic experience (AX), that is, the decision recipient's response to the usage of algorithmic systems such as ADMS (Alvarado & Waern, 2018).

In the information science (IS) literature, social consequences are categorized into perceived fairness, accountability, and transparency commonly abbreviated as FAT or FAccT. Despite the fact that many academic researchers have discussed these topics independently (e.g., (Diakopoulos, 2016; Kizilcec, 2016; Wang et al., 2020), Shin & Park (2019) suggests that these issues are interrelated. Specifically, Shin & Park's (2019) algorithmic acceptance model advocates that perceived fairness, accountability, and transparency are an antecedent of trust.

The current thesis seeks to identify factors influencing the use and acceptance of ADMS from the standpoint of the decision receiver. First, we emphasize factors critical to ADMS acceptance, and then analyze the role of algorithmic opaqueness in ADMS adoption.

1.2 Research Objectives

The overarching objective of this thesis is to explore the impact of individuals interacting with ADMS. Specifically, we seek to understand what is important to decision recipients in accepting ADMS, to investigate how decision recipients perceive algorithmic type, and to investigate how algorithmic type influences ADMS use.

Overall, we argue that it is important to include the decision recipient's perspective in the development of ADMS with high a degree of consequentiality to prevent or mitigate unintended consequences. Table 1-1 summarizes the specific research questions, motivation, and key findings.

Table 1-1 – Summary of Research Questions, Motivation and Key Insights

Main Objective	**Research Questions**	**Motivation**	**Key Insights**
Understanding the impact of individuals interacting with ADMS.	RQ1: What drives the acceptability of automated decision-making systems?	To follow up on Grange's (2022) online scenario-based research and identify elements influencing the algorithmic experience from the perspective of the decision recipient.	• An algorithm processing university application is unethical. • Double standards exist between algorithms and human experts. • Favouritism towards outcome rather than during the process.
	RQ2: What algorithmic characteristics do decision recipients perceive as important in ADMS?	To determine if the algorithmic type can aid in the communication of the decision-making process in ADMS.	• Comprehensibility of decision-making process. • Adaptability to new information. • Robustness to protect itself from malicious behaviour. • Quantity and quality of data input to decision-making. • Bias in the decision-making process (historical bias and expert bias).
	RQ3: What is the effect of algorithmic type in the use of ADMS?	To determine if algorithmic types facilitate unintended behaviour in ADMS.	• Algorithmic characteristics influence individuals' preference and usage pattern. • Decision recipients tend to game the system when exposed to RB ADM • Decision recipients use the system as intended when exposed to MLB ADM

1.3 Contributions

This thesis applies the concept of social acceptability and the FAccT framework to explain the social consequences of ADMS. The findings of this thesis raise awareness of the significance of decision recipients' perceptions of the usage and acceptance of ADMS, and incorporating decision recipients' perceptions into ADMS development can lead to a more socially responsible diffusion of ADMS. This also includes recommendations to improve communication with decision recipients. Lastly, the application of qualitative analytical methodologies complements existing quantitative research in the literature on the responsible diffusion of ADMS within the domain of IS and human-computer interaction (HCI).

1.4 Article 1

The first study investigates factors influencing ADMS usage by analyzing secondary data using coding and thematic analysis. The secondary data comes from Grange's (2022) study from an unexplored open-ended question from an online scenario-based experiment of n=300. This experiment investigates whether automation and automation logic relate to the societal acceptability of a decision-making process in the context of university admission. According to the findings, key factors influencing their acceptability in ADMS included the moral capability for ADMS to render important judgements, the double standard between algorithms and human experts, and the favouritism of outcome rather than the fairness of the process. Findings from qualitative data support existing quantitative studies in the body of knowledge in the field of HCI and ADMS. The concept of social acceptability in ADMS advocates for the advancement of the socially responsible implementation of ADMS.

1.5 Article 2

The second study uses semi-structured interviews n=13 to evaluate how algorithmic type and consequential decision impact the usage of ADMS. The interview presents the use of chatbot employment interviews via several scenarios describing the decision-making process employing two different algorithmic types (i.e., rule-based and machine learning based) in the context of high level of consequential judgments. The fairness, accountability, and transparency (FAccT) framework (Shin & Park, 2019) inspired the formulation of interview questions. The qualitative analysis employed the coding method and the thematic analysis in accordance with the grounded theory process. The findings show that participants' preferences and behaviour when using ADMS are influenced by their understanding of algorithmic characteristics. This finding has several implications, one of which is that people's preconceptions about how the types of algorithms function can be biased and have unintended effects. For a more responsible implementation ADMS, we propose sharing essential information about algorithmic characteristics to decision recipients to ensure a basic homogeneous understanding. In addition, we suggest providing clear instructions to decision recipients on how effectively employ ADMS for optimal results.

1.6 Thesis Structure

The rest of the thesis is organized as follows. In chapter two, we present the first article that investigates the factors that influencing ADMS acceptance in the context of university admission. We continue with the second article in chapter three to investigate how algorithmic type influences a decision recipient's perspective of using ADMS in the context of a chatbot supported job interview. Finally, in chapter four, we summarise the research questions, findings, limitations, contribution, future research opportunities from both articles, and my takeaway from the research experience.

Chapter 2: What Makes Decision Recipients Accept Automated Decision-Making Systems? - An Exploratory Study

ABSTRACT

Automated decision-making systems (ADMS) are increasingly employed to make more efficient, highly consequential health, hiring, and juridical decisions. ADMS can be valued for its innovative technology to solve problems, but its overall success and benefit depends on whether it is accepted by society. This paper explores different components of societal acceptability in automated decision-making systems. Using secondary data from Grange's (2022) investigation of the societal acceptability of algorithmic type in the automated university admission process, we analyzed participants' preferences of four decision-making processes: a rule-based human process, an experienced-based human process, a rule-based algorithm, and a machine learning based algorithm. The findings suggest that four factors contribute to the societal acceptability of ADMS: moral principles in algorithmic judgement, expectations of subjectivity in ADM, favourableness of the decision outcome, and human-ADM relationship. Consistent with Grange's (2022) initial research, the algorithm's capability to process an individual's uniqueness affects their acceptance of ADMS. However, little information is provided on the preference for different algorithmic types due to the lack of structure within the qualitative data. The article presents the details of the findings.

Keywords: ADMS, automated decision-making systems, social acceptance, responsible innovation, HCI, human-centered interaction

Research method: Qualitative analysis using coding method and thematic analysis

2.1 INTRODUCTION

Automated decision-making systems (ADMS) are defined as a set of instructions processing information in the form of input data using an algorithm to generate an output of some kind (Grange, 2022). These systems integrate algorithmic types such as rule-based systems (RB) and machine learning algorithms (ML), a subset of artificial intelligence (AI), to assist or replace human decisions (Benbya et al., 2021). AI has distinct capabilities from automation, such as learning from large historical data and solving complex problems creating a variety of decision-making opportunities (Ibrahim & Abdulazeez, 2021; van Ginneken, 2017). With AI-based algorithms introduced into ADMS, it is becoming an ideal solution to solve societal level problems in various fields as they are prized for their flexibility, speed, scalability, decision-making and personalization (Wilson & Daugherty, 2018).

Numerous companies in various industries are using ADMS to solve societal level problems. In the healthcare industry, ADMS are able of recognizing body organs from medical images, classifying lung diseases, and detection of lungs nodules, resulting in improved diagnosis and treatment outcomes (Qayyum et al., 2020). Products such as Google's Deep Mind, IBM's Watson, and Caption Health healthcare providers analyze and interpret results, leading to better decision-making and reduced healthcare costs by prioritizing health management over disease treatment, resulting in fewer hospitalizations (Bohr & Memarzadeh, 2020).

In the transportation industry, the use of self-driving systems can revolutionize the transportation industry and provide numerous benefits for society. Self-driving systems such as those from Tesla, Argo.ai, and Maymobility, can perform driving functions without human intervention, improving productivity, reducing traffic congestion, increasing efficiency, and minimizing environmental impact on the roads (Ryan, 2020).

While ADMS are increasingly being deployed to substitute tasks previously processed by humans through automation (Grange, 2022; Waytz et al., 2014), they are also beginning to make crucial judgments that directly affect the lives of those on the receiving end. Examples of this include, measuring the risk of recidivism (e.g., COMPAS) and giving mortgages (e.g., Rocket Mortgage). An ADMS can be prized for its innovative technology to solve problems, but its overall success and benefit depend on whether or not they are being accepted by society. Indeed, ADMS are socio-technical artifacts embedded in layers of contexts, relationships and societal roles (Araujo et al., 2020; Marabelli et al., 2021). Just like the utilization of AI in a performance evaluation leading to an unexplained job termination causing employee frustration and bringing about legal disputes (Park et al., 2021), the discontinuing of AI-based hiring due to unexpected discrimination against gender ("Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women," 2018) or the manipulation of videos of politicians using AI software such as Deepfake to say something they did not (Winder, 2019). Exploring the social side of ADMS is important in order to develop even better products and prevent unintended negative consequences.

Developing successful ADMS that are beneficial to society requires attention to predicting, influencing, and evaluating options for various aspects of social acceptability within the design cycle (Koelle et al., 2019). Social acceptability focuses on factors affecting decision recipient's experience and the acceptance of its interactions (Koelle et al., 2019). However, limited attention is given to this conceptual framework within the literature responsible for the innovation and diffusion of technology (Grange, 2022; Tabourdeau & Grange, 2020).

This research explores what drives the acceptability of ADMS by detecting factors affecting the algorithmic experience (AX) from the decision recipient's perspective. As factors in social acceptability differ given the context and perspective (Koelle et al., 2019), we specifically analyze the factors in the context of university admission from the perspective of university students. The data analysis and findings are built upon secondary data from a study that also adopted the lens of societal acceptability, notably the societal acceptability of algorithmic types and decision-making agents in ADMS (Grange, 2022). This study is further discussed in the literature review.

We investigated our research question by utilizing previously unexplored qualitative data from Grange's (2022) work. The coding method and thematic analysis were employed to uncover various factors affecting social acceptability in ADMS. The findings contribute to the advancement of developing a responsible innovation of ADMS within the domain of information science (IS) and human-computer interaction (HCI).

The remainder of the paper provides an overview of the relevant background literature followed by a description of the data collection and analysis method, the study findings, and its practical implications. We close the paper by discussing limitations, future research opportunities, and concluding remarks.

2.2 LITERATURE REVIEW

The objectives of the literature review are two-fold: provide background information on the origin of the data, which this study builds upon, and present the concepts related to the acceptability of ADMS.

2.2.1 Research Method

In the first round of research, we used the Web of Science database using keywords such as "algorithm (and algorithmic) decision making, and automated decision making" within the basket of eight in information systems journals published within the last ten years. Then, the resulting titles and abstracts yielded by the search were read thoroughly. An article was deemed relevant and kept when it addressed automated-decision making or algorithmic decision-making (ADM) and human-computer interaction (HCI) topics (ethics, behaviour, social, psychology, perception). These articles usually have at least one of the keywords from the search. Articles referring to applying and developing mathematical and statistical algorithms (technical and methodological aspects) were excluded. The results were first extracted using Zotero, a reference management system, by storing the references from each keyword in a designated folder. The articles retained allowed the identification of other related articles, which were found using the Google Scholar search tool. Most of the literature review was founded upon the CHI conference and the FAccT conference from the Association for Computing Machinery (ACM) publisher.

CHI: Conference on Human Factors in Computing Systems - Year: 2012-2022
FAccT: Fairness, Accountability, and Transparency - Year: 2012-2022

The keywords for the article search included: automated (and algorithmic) decision-making systems, machine learning, rule-based algorithms, perceived algorithmic bias, human bias, and perceived algorithmic fairness/accountability/transparency.

Below an area of research related to using and adopting ADMS are presented. Specifically, these topics will provide background on how the perception of algorithmic type and task importance affects perception towards ADMS.

2.2.2 Response Towards Algorithms

As algorithms become entangled in people’s daily life (e.g., Facebook, Netflix, and Uber), there is a growing response towards the use of ADMS. This response can be categorized into two streams of thoughts; those who trust the output of the algorithm in decision-making (algorithm appreciation) and those who do not, thus trusting the decision of a human expert, which is often called (algorithm aversion) (Logg et al., 2019). However, these responses are not a simple yes or no decision but rather a spectrum where the difference lies in the context and their pre-established perception of algorithms in the given tasks of a specialized system (Koelle et al., 2019; Logg et al., 2019).

Previous works show that responses toward algorithms are context-dependent. Some demonstrated that people relied on algorithms for decisions, such as guessing information from the picture, forecasting song popularity, and predicting online matchmaking (Logg et al., 2019). Consistently, other works found that algorithms were evaluated on par or even better than experts for recommending information in the field of media, health, and justice (Araujo et al., 2020). However, when it came to a high level of consequential decisions such as: university admission, management decisions in human resources (i.e., hiring and firing), or child maltreatment hotline screening, decision recipients tended to favour experts (De-Arteaga et al., 2020; Grange, 2022; Park et al., 2021). It is noticeable that the higher the degree of consequentiality to the individuals, the higher the level of algorithm aversion.

Algorithms have the capacity to analyze and lead quicker and more consistent decisions than humans, thriving particularly in a larger volume of data (Bodea & Karanikolova, 2018). However, decision recipients do not perceive ADMS the same as humans (Delecraz et al., 2022; Dietvorst, 2015; Grange, 2022). Each decision-making agent (algorithm vs. human) has their own attributes, which make them appropriate for different types of decision-making. For example, Dietvorts et al. 's (2015) found that participants believed the algorithm forecast was good at avoiding mistakes and weighting various attributes appropriately, whereas humans were better at improving with practice, learning from mistakes, and finding underappreciated candidates.

An individual’s response towards algorithms is not set in stone, though decision recipients generally appreciate algorithms; once they experience an erroneous or upsetting event, they quickly lose confidence in algorithms (Bucher, 2017; Dietvorst, 2015). An individual’s characteristics can also influence their perception of algorithms. For example, Wang et al. (2020) found that motivation and education level played a significant factor in algorithmic appreciation, where participants tended to favour the algorithm only if the outcome was favourable to themselves, and participants who were knowledgeable about algorithms tended to show algorithm appreciation. In addition, an expert in a field tends to rely less on advice from algorithms within the same field (Logg et al., 2019). An individual’s overconfidence and egocentrism also put more weight on their intuitive judgment over other advisors (Logg et al., 2019).

Overall, an individual’s response toward algorithms is not rigid, but layered and complex. It falls within a spectrum based on various deciding factors that are both context and non-context specific. With this, the remainder of the article focuses on the perspective of students in the context of university admission.

2.2.3 Grange’s (2022) Research

Recognizing the importance and the numerous benefits and risks associated with the use of ADMS for society, Grange (2022) presented two concepts that could help advance the literature on the responsible diffusion of these systems. First, she highlighted the concept of societal acceptability to support the diffusion of ADMS, which was lacking in the literature. Then, giving attention to the broad concept of artificial intelligence and automation, she distinguished it with different types of algorithms. Grange’s (2022) study involved linking the above-mentioned concepts. Specifically, whether automation and automation logic matter in the societal acceptability of a decision-making process by exploring two factors: the decision-making agent (human vs. algorithm) and the decision-making process (predetermined vs. emerging). Predetermined logic is defined as a decision-making process relying on defined rules and guidelines, whereas emerging logic is defined as a process relying on experience.

With the goal to advance the responsible diffusion process of ADMS, Grange conducted quantitative research consisting of exploring whether automation and automation logic matter to the societal acceptability of a decision-making process. Notably, in her research model, she conceptualized societal acceptability through the decision recipient’s attitude as it is considered an overall judgement and key component to acceptability. In addition, a decision-making agent (i.e., human vs. algorithm) is differentiated by the perceived ability to capture an individual's unique characteristic (i.e., uniqueness neglect) formulated into H1a and H1b. Then, relying on the concept of procedural justice, the decision-making process (predetermined vs. emerging) is differentiated by a lack of fairness in a decision-making process.

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' attitudes toward the decision-making process.
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' attitudes toward the decision-making process.

Grange's (2022) online scenario-based experiment, which consisted of a few sets of questionnaires and an open-ended question, was conducted using a sample of n=300 18-year-old or older undergraduate or recent graduate (> 2 years) students recruited through Qualtrics from North America. The online scenario-based experiment used a 2x2 factorial design, decision-making authority (human vs. algorithm), and logic-based decision-making (predetermined vs. emerging). Each participant was randomly assigned to one of four treatment groups (T1 n=78, T2 n=73, T3 n=74, T4 n=75). Four decision-making procedures were established: an expert using a predetermined set of criteria (human - predetermined), an expert using their professional experience to make a prediction (human - emerging), an algorithm using predetermined rules (algorithm - predetermined), and an algorithm relying on historical data to make a prediction (human - emerging) (algorithm - emerging).

In the context of university admission, participants were presented with different decision-making process scenarios, and then answered in Likert-tyle questions to capture the measured variables (uniqueness neglect, procedural justice, and attitude). Lastly, participants were asked to set a price for each decision-making process, and to provide a rationale for it in an in an open-ended format.

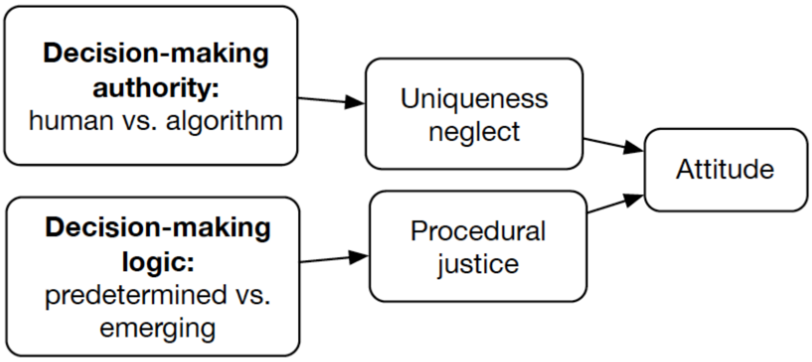

Figure 2-1 – Grange's (2022) Research Model

Table 2-1 – Grange's (2022) Statistical Findings

	Test	Results	Conclusion
Model Validity Check	Cronbach Alpha	Uniqueness Neglect: 0.80 Procedural Justice: 0.78	Acceptable consistency of items in measurements.

		Attitude: 0.93	
	One-way ANOVA test	MC1: p<0.001 MC2: p<0.001	Significant difference between treatment groups.
Hypothesis Testing	Two-way ANOVA test	H1a: p<0.001 H1b: p<0.001 H2a: p=0.18 H2b: p<0.001	H1a: Supported H1b: Supported H2a: Not Supported H2b: Supported

The results supported H1a, H1b, H2b, and rejected H2a showing that decision-making authority and decision-making logic are factors influencing societal acceptability of decision-making procedures. Specifically, a significantly higher level of uniqueness neglect was observed for algorithmic decisions compared to human decision-making (H1a). Additionally, uniqueness neglect (H1b) and procedural justice (H2b) were both significant predictors of attitude. Unexpectedly, not enough evidence supported the proposition that algorithmic type would affect attitude (H2a), which could be further explained in the open-ended question not yet explored, reflecting on participant's rationale on how they perceived the different algorithmic types.

As our research question builds upon the qualitative data from the above-mentioned online scenario-based experiment, its findings can provide further insights into why algorithmic type does not affect attitude (H2a).

2.3 METHOD

To explore our research question given the existing qualitative data, a coding method and thematic analysis for qualitative analysis were employed from Saldaña's *The coding manual for qualitative researchers* to group and identify themes (Saldaña, 2016). The analysis followed a cyclical process of four cycles, along with its respective analysis summarized in Table 2-2.

2.3.1 Data

This study used data originating from Grange's (2022) online scenario-based experiment on the societal acceptability of ADMS. We specifically used the unexplored open-ended question from this online experiment. In that context, our participants from the four treatment groups were presented with a standard fee of $70.00 for processing their university application. Each group was presented with a different decision-making process. These four decision-making processes included: an academic advisor relying on school guidelines, an academic advisor relying on their own experience, an algorithm relying on school guidelines and an algorithm relying on historical data to make this prediction. We then informed participants that the acceptance rate into the programs was the same for the four decision-making processes. With this knowledge, participants were asked to set a price ranging from $0-140.00 for each decision-making procedure. They were provided with an open form to explain their rationale for the price they chose. We assumed that the application cost participants set indicated the varying level of importance they placed on each decision-making process. In the open-ended question, the rationale for how close or far the application fee for each decision-making procedure was to the standard fee was investigated.

2.3.2 Coding Structure and Approach

Given the open-ended nature of the question, there was limited control over the structure of the answers containing information on one or several treatment groups. Thus, to organize and classify codes for the analysis, each code followed a structure which included the treatment group with the intended information.

For the data extraction, two coding methods were used: in vivo coding – which uses one to three words drawn from the participant's language as a code, and holistic coding – a preparatory approach that attempts to grasp general themes or issues from the data before the more detailed coding process (Saldaña, 2016). The raw data was then uploaded on Reframer by Optimal Sort, a qualitative analysis tool that allows for creating, visualizing, and managing codes.

2.3.3 Procedure

The procedure consisted of four cycles to refine the analysis. In the first round of coding, a sample of n=15 was randomly selected from each treatment group to understand the type of qualitative answers provided by the online scenario-based experiment.

In the second coding round, all n=300 data were used. Each data was coded either S1, S2, S3, or S4 based on the participant's respective treatment group from the online scenario-based experiment. On Reframer, the main themes of participants' responses were categorized: Human, Algorithm, Data, Rules, Cost, and Disregard (see Table 2-3). An extract of coding can be found in Appendix 1.

Table 2-2 – Definition of Categories

Category	**Description**	**Count**
Human	Responses describing human characteristics when making decisions.	90
Algorithm	Responses describing algorithmic characteristics when making decisions.	51
Data	Responses based on emerging historical information to make decisions. This does not differentiate decisions made by algorithms or humans.	8
Rules	Responses based on predetermined rules and criteria to make decisions. This does not differentiate decisions made by algorithms or humans.	10
Cost	Responses that justify the cost they set.	69
Disregarded	Responses that do not provide enough information relevant to the question.	115

The third cycle consisted of merging codes with similar meanings and starting to map them. The in vivo coding method uses participant's language to name the codes, thus each code is unique, but many have the same meaning. Therefore, we combined these codes together to simplify the analysis. Then, by connecting the codes, we discovered emerging patterns.

The fourth cycle was an exploratory grouping. To distinguish the range of different opinions, we further divided the main categories into positive and negative. Codes referring to a supporting opinion (e.g., transparent, accurate, responsible) were grouped in the former sub-category and codes referring to an undermining opinion (e.g., error, bias, lacks nuance) will be grouped in the latter.

Table 2-3 – Summary of Coding Cycle

Cycles	Description of Cycles	Observations
1	Preliminary analysis with n=15	● Most answers were two sentences long ● Most answers were rejected due to a lack of clarity
2	Analysis with n=300 separated by four treatment groups. Answers were tagged by codes which were then grouped by themes.	Categories: ● Social Aspect ● Data Quality ● Predetermined Rules ● Admission Cost ● Disregarded
3	Merging redundant codes	
4	Exploratory grouping	Factors: ● Moral Principles in Algorithmic Judgement ● Expectation of Subjectivity in ADMS ● Favourableness of the Decision Outcome ● Human-ADMS Relationship

2.4 FINDINGS

Participants' responses to automated decision-making systems differed depending on the decision-making method and agent used. These differences were grouped by themes to support the research question. Findings from the interview were categorized into four themes: the moral capability of ADMS to render important judgements, the double standard between algorithms and human experts, the favouritism towards outcome rather than the fairness of the process, and the idea of human-ADMS relationship for improved decision-making.

Overall, we found that participants' rationale from the scenario-based online experiment was brief—generally two sentences—and to the point. Most answers consisted of the positive and negative points of view of the four treatment groups and the opinion and criticism towards the individual factors (human, algorithm, data, and rules). Comments that lacked clarity were removed from the analysis.

2.4.1 Moral Principle in Algorithmic Judgement

Many participants believed that giving an algorithm the authority to process university applications was unethical. Ethics and morality are principles that individuals believed were important. As ethics and morality played an important role in this scenario, many believed that a human should render judgement on an important decision like university admission acceptance. The reason being, that there are decisions only humans are worthy of making as they have the ability to see another individual as a person rather than as a number, and to make decisions based on circumstances by assessing worthiness. Examples of this include:

"I would trust an academic advisor because they can use their morals and knowledge."

"A real person making these decisions will always be worth more than computer algorithms."

2.4.2 Expectation of Subjectivity in ADMS

There is a different set of expectations towards decision-making agents. The topic of subjectivity in decision-making was mentioned in the online scenario-based experiment. On the one hand, an algorithm was viewed negatively due to its lack of subjectivity (i.e., historical bias in the data, unable to understand nuances or read emotions). At the same time, humans were viewed negatively for being subjective due to their potential biases (i.e., individual bias, potential for error, favouritism). In our study, given the double standard over subjectivity, we observed that when making decisions, subjectivity was desired in algorithms while less desired in humans. Examples of this includes:

"[…] I am willing to pay more when it is a real person as it takes their own personal time to complete. I am also willing to pay more for a review based on quality criteria and standards than a subjective review."

"Regardless, students are not given the chance to show why or how they fit into the program. If a human evaluated each individual application and went based on both standardized and quality of subjective criteria, then I'd maybe pay more."

"I feel that humans are not as good as technology for this situation because that leaves too much room for human error and especially bias or favouritism."

"*If it's an algorithmic process based on a good data set, that's reasonable as long as the broad underlying factors are well known. I can know whether or not to bother, in other words. I see no point in paying for a computer to inflexibly apply a subjective standard set at some point in the past. At least with people, you can "read the room" and know if you're truly a fit for a program."*

"I don't like having to pay the admin fees at all, but I'm familiar with all the flaws in an automated system. A one size fits all approach lacks nuance, and AI-based on historical data carries with it historical bias that I believe would be detrimental to be as an afab nonbinary person."

2.4.3 Favourableness of the Decision Outcome

For some participants, the acceptance of an ADMS depended more on the outcome that most benefited themselves rather than the decision-making process itself. Here, participants valued more the favourability of the outcome regardless of the decision-making process. For example:

"You should pay the most for getting a better chance of being accepted."

"You have a better chance with a person processing your information than you do with a computer doing it, in my opinion, because computers can't tell your actual worth."

2.4.4 Human-ADMS Relationship

Participants understood the benefits and weaknesses of each decision-making agent. When the decision-making agent was an algorithm, it had the benefits of reducing human bias while saving time and providing accurate and consistent results. However, the setbacks were the lack of ability to understand emotions, distinguish nuances, and use biased data from the past. In addition, some participants felt disrespected and morally unsettled to be judged by an algorithm.

When the decision-making agent was a human, they were seen as being flexible in understanding an individual's circumstances and were trusted to make complicated and complex judgements. On an individual level, it was also perceived as more respectful to cast judgment on others than be judged by algorithms. Still, human decision-makers are subject to individual bias, favouritism, and tend to make more errors.

It is important to acknowledge that algorithms complement human weaknesses and vice versa. Having a human review over the algorithm in decision-making is an ideal scenario.

"Truthfully, a combination of each of the scenarios might be ideal (computers could compare data, but advisors could make the final decision based on a number of factors)."

2.4.5 Summary of Findings

Table 2-4 summarize the findings from the analysis.

Table 2-4 – Summary of Factors Influencing Social Acceptability in ADMS

Factor Influencing Social Acceptability in ADMS	Description of Each Factor
Moral Principle in Algorithmic Judgement	Social acceptability towards ADMS is driven by the respect given to the decision recipients with the belief that some decisions can only be made by another human.
Expectation of Subjectivity in ADMS	Social acceptability towards ADMS is driven by whether the individual's expectation towards the algorithm is met, especially on the topic of subjectivity.
Favourableness of the Decision Outcome	Social acceptability towards ADMS is driven by an individual's motivation to select based on their chances of being admitted.
Human-ADMS Relationship	Social acceptability is driven by building a relationship between humans and algorithms to make an even decision.

2.5 DISCUSSION

There are a myriad of responses observed for the same topic where some participants found the algorithms to be biased. Others believed the opposite, some participants accepted algorithms if they were able to process an individual's uniqueness, and others responded the same if the outcome was reviewed by an expert. This is consistent with the notion that social acceptability is a spectrum rather than a binary decision (Koelle et al., 2019). An individual's response towards ADMS is responsive to the information brought to them by context.

Participants' response towards the use of ADMS was subjective and dependent on individual values and morals (Luccioni & Bengio, 2020). Just like the findings from the study, some were receptive to the use of ADMS for university admission, while others expressed it as conflicting with their values and morals. An advisor processing university admission was viewed as a sign of respect for the decision-recipient and to maintain their dignity. In addition, building an interpersonal relationship between the decision agent and the decision recipient is an important skill to develop as humans are social beings. ADMS systems have the potential to improve society, but researchers and practitioners must be able to properly leverage their power by taking into consideration the ethical principles that society is run by. We suggest the value-sensitive design approach (Friedman et al., 2008; Hoven, 2013) in the design cycle of ADMS for a more socially and ethical practice using algorithms that contribute to the morally responsible innovation and diffusion of ADMS.

As participants' expectation of decision-making is different, participants' motivation also plays a role in social acceptability. Participants expect ADMS to have subjective traits, while they expect the opposite when it comes to human decision-making. This contradictory finding is aligned with previous studies which found that people see algorithms differently from humans (Delecraz et al., 2022; Dietvorst, 2015; Grange, 2022) and that algorithms are judged more harshly than humans (Dietvorst, 2015). Participants' acceptance of the outcome of ADMS is different, and the ADMS are accepted if the outcome is in their favour. Similarly, studies mentioned human egocentrism (Logg et al., 2019) and self-interest (Wang et al., 2020), favouring the good of the self over the good of the collective within a similar context. Biases are more the norm than they are the exception (Delecraz et al., 2022), everyone has their own definition of truth and acceptability. It is inevitable for humans to be biased either consciously or unconsciously. However, recognizing that ADMS can provide objectivity in decision-making that humans cannot and vice versa, just like some participants suggested, the idea of a human-ADMS relationship or the concept of human-in-the-loop (De-Arteaga et al., 2020) seems to be the most acceptable and ethical option.

Findings from this study provided a limited distinction of acceptability between algorithmic types in the societal acceptability of ADMS from Grange's (2022) work. There were no clear comparisons among algorithm types, but a few explanations for this are possible. Between the decision-making agent and the algorithmic type, participants focused more on the differences between decision-making agents as it was the easier and more important factor to make the distinction in acceptability. Another possibility is that acceptance of algorithms goes beyond the understanding of the algorithm itself but rather is based on individual differences and morals (Park et al., 2021). Given the context of automating university admission, a novel context for many students, some people might find it harder to be accustomed to change. In addition, some students might not have enough knowledge of the algorithm's capabilities to make an informed decision on the acceptance of ADMS.

2.6 LIMITATIONS AND FUTURE RESEARCH

Just like all research, this study has limitations. Since secondary data was used, the quality of the questions and responses was limited. Ending the online scenario-based experiment with an open-ended question could have influenced how much attention was paid to the question and answer. Furthermore, considering the short length of the responses, it is conceivable that not all answers were correctly comprehended or interpreted. As a result, future work can be improved by conducting an interview and asking specific questions. Moreover, an extra set of eyes could help the data analysis process to strengthen the coding procedure and improve the quality of findings

In the IS research community, unintended consequences of algorithms can be categorized into perceived fairness, accountability, and transparency (or FAT). Notably, Shin & Park's (2019) algorithmic acceptance model proposes perceived fairness, accountability, and transparency as an antecedent of trust. It could be a useful framework for future research in the domain of ADMS.

2.7 CONCLUSION

The process of automated decision-making is done behind the interface and is often hidden from participants. However, decision recipients are aware of its existence and have concerns over its use. This study offers insights into what factors affect social acceptance of ADMS. These findings from the decision recipient perspective can provide insights to researchers and practitioners regarding factors to consider prior to or during automating a process that is socially acceptable. Furthermore, findings from qualitative data support existent quantitative studies in the body of knowledge in the field of HCI and ACM. The concept of social acceptability in ADMS advocates for the advancement of the socially responsible implementation of it.

Chapter 3: Decision Recipients' Perspective of Algorithmic Types Towards the Use of Automated Decision-Making Systems : A Qualitative Analysis

ABSTRACT

Algorithmic innovation is positively contributing to the economy, and businesses across different domains are increasingly integrating artificial intelligence-driven algorithms to make critical and consequential decisions. However, algorithms are not always transparent to the public, and its decision-making process is subject to bias. For these reasons, decision recipients perceive and behave differently preventing a successful implementation of automated decision-making systems. In a set of 13 semi-structured interviews (N=13), we asked participants how algorithmic type (rule-based and machine learning based) influences their use of chatbots in the hiring scenario. For interview, we used the fairness, accountability, and transparency (FAccT) framework. Following the grounded theory process, we then analyzed the qualitative data using rigorous coding and thematic analysis. Our findings show that participants' use of ADM was driven by their preconception of algorithmic types and their motivations. For a more responsible use of ADMS, we suggest focusing on transparent guidelines on how to better complete the tasks. The following presents the details of the findings

Keywords: Human-computer factors, HCI, algorithmic experience, AX, ADMS, automated decision-making, perceived algorithmic fairness, accountability, and transparency, FAccT

Research method: Semi-structured interviews

3.1 INTRODUCTION

The advancement of algorithms is apparent in everything we do as it is reshaping traditional decision-making. As artificial intelligence algorithms, notably machine learning, progresses, their performance draws closer to the level that of human capability (Luccioni & Bengio, 2020). For example, machine learning can recognize and display emotions by recognizing pattern from the given data (Benbya et al., 2021). Algorithmic innovation also positively contributes to the economy; businesses across different domains are increasingly integrating artificial intelligence-driven algorithms to make critical and consequential decisions. Some examples of these decisions include self-driving systems that automate driving such as Tesla and Comma.ai to improve safety, reduce traffic congestion, and increase efficiency on the road (Ryan, 2020). Automate customer service using in chatbot such as Netomi and Zendesk Answer Bot to provide 24/7 assistance, improve response times, and reduce the workload of customer service representatives (*A Guide to the Best AI Chatbot*, 2023). Improve health diagnosis with Google's Deep Mind and IBM's Waston Health to help healthcare professionals make more accurate diagnoses and improve outcomes (Bohr & Memarzadeh, 2020). Personalized marketing campaigns systems notably Phrasee, and Smartwriter.ai to increase sales (Payani, 2023).

In this study, we adopted the European Commission's (Bodea & Karanikolova, 2018) definition of automated decision-making (ADMS): decisions by technological means without human involvement, including a variety of types of algorithms. These decisions involve various processes, from supporting human decision-makers to fully automated processes across contexts (Araujo et al., 2020). We also recognize that ADMS are not neutral (Koenig, 2020) as developers and designers carry their own biases and values in the development cycle of ADMS (Alvarado & Waern, 2018; Kordzadeh & Ghasemaghaei, 2021) and that an unbiased algorithm does not exist (Delecraz et al., 2022).

With ADMS providing new business opportunities along the way, these implementations have been shown to develop consequences that are both positive and negative to the benefit of society (Luccioni & Bengio, 2020). Indeed, ADMS do not function in isolation; they are socio-technical artifacts embedded in layers of contexts, relationships and societal roles (Araujo et al., 2020; Kordzadeh & Ghasemaghaei, 2021; Marabelli et al., 2021). Behind every automated decision-making are algorithms that are given the authority role in the decision-making process.

The algorithmic decision-making process is often strategically hidden from the decision-recipient's experience to protect its intellectual property, prevent gaming the system and provide a better experience (Burrell, 2016; Eslami et al., 2019). However, at other times, the decision-making process is involuntarily hidden from the managers and engineers themselves for algorithms (e.g., machine learning and natural language processing) are labelled "opaque" for its black-box nature (Burrell, 2016; Marabelli et al., 2021). Many researchers have pointed out the negative side-effects of hiding algorithmic decision-making processes, either intentionally or unintentionally. Some issues include: hidden agenda/deception, discrimination, lack of human experience, and valuing different priorities such as fairness, accountability, and transparency (Eslami et al., 2019; Lee et al., 2017; Mehrabi et al., 2021; Pethig & Kroenung, 2022; Rutjes et al., 2019).

When the root cause of ADMS' negative side effects is not properly identified and prevented, they can lead to great economic loss and social distress. For example, the discontinuation of AI-based hiring from the unexpected discrimination against gender ("Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women," 2018) or the utilization of AI in performance evaluation for unexplained job termination causing employee frustration resulting in legal disputes (Park et al., 2021). Researchers, managers, designers, and engineers of the ADMS share the responsibility of mitigating the risks for society (Lima et al., 2021; Luccioni & Bengio, 2020).

One of the underlying causes lies within the decision recipient's mind. Even if algorithms are proven fair mathematically, the outcome or process of the decision might still be perceived as unfair as different concepts of fairness exist within and across stakeholders (Delecraz et al., 2022; Lee et al., 2017). Thus, it is important to undertake a human-centered approach to the consequences. Within the human-computer interaction (HCI) literature, there is limited work investigating algorithmic experience (AX), that is, the decision recipient's response to the use of algorithmic systems such as ADMS (Alvarado & Waern, 2018). Recent works studied several contextual factors contributing to the algorithmic experience, such as decision outcome (Wang et al., 2020), decision-making process (i.e., algorithmic type, explanation type) (Grange, 2022; Wang et al., 2020; Kizilcec, 2016), individual differences (e.g., education and demographic) (Wang et al., 2020), and type of tasks (Araujo et al., 2018; Logg et al., 2019) in contexts such as health, media, justice, and hiring (Araujo et al., 2020; Delecraz et al., 2022). These studies employed quantitative measures to discover factors contributing to the algorithmic experience, but further work employing qualitative approaches is needed to understand the root cause of their perception by focusing on the individual's beliefs and thoughts.

Our study addresses this gap by using qualitative data to focus on the beliefs and thoughts of the individuals. Given the rise of AI and its capabilities of making a high level of consequential decisions and leaving a great social impact (Luccioni & Bengio, 2020), we focus our research on the study of how algorithmic types impact a decision recipient's perception. Here, we define algorithmic type as a characteristic of algorithms based on the opacity level of the decision-making process that is either static and transparent (e.g., rule-based) or dynamic and a "black box" (e.g., machine learning) (Marabelli et al., 2021). Thus, we investigate how algorithmic types affect the decision recipient's perspective towards the use of automated decision-making systems with the following research questions:

RQ1: What algorithmic characteristics do decision recipients perceive as important in ADMS?
RQ2: What is the effect of algorithmic type in the use of ADMS?

Using semi-structured interviews, we explored the research questions' algorithmic type (rule-based and machine learning based) in the field of HCI and IS. We used the Algorithmic Acceptance Model (Shin & Park, 2019) as a framework to explore factors influencing the use of ADMS. We selected the job hiring process for being a socially relevant context and a chatbot job interview run by either rule-based or machine learning based algorithms for a highly desired vs. mundane job. Our work extends the literature on people's experience with algorithms (algorithmic experiences) by discussing the algorithmic characteristics and its associated behaviour to help researchers, managers, designers, and engineers gain new insights useful to the design, development, and management of ADMS to prevent economic and social distress.

Our results suggest that algorithmic type influences the decision recipient's perception and behaviour towards ADMS. Everyone has their presumption of algorithms; some are accurate, and others are biased or incomplete. An individual's presumptions of algorithmic characteristics will consequently influence their behaviour, from the type of algorithm they prefer to the way they will use it. Consequently, we recommend offering information on how to effectively employ ADMS for the best results.

The remainder of the paper provides an overview of the relevant background literature followed by a description of the data collection and analysis method, the study findings, and the research and practical implications. We close the paper by discussing limitations and future research opportunities.

3.2 BACKGROUND LITERATURE

The objectives of the literature review are to present the concepts related to the research question related to the algorithmic types and the consequential decisions made by ADMS.

3.2.1 Research Method

In the first round of research, we used the Web of Science database using keywords such as "algorithm (and algorithmic) decision making and automated decision making" within the basket of eight in information systems journals published within the last ten years. Then, the resulting titles and abstracts yielded by the search were read thoroughly. An article was deemed relevant and kept when it addressed automated-decision making or algorithmic decision-making (ADMS) and human-computer interaction (HCI) topics (ethics, behaviour, social, psychology, perception). These articles usually have at least one of the keywords from the search. Articles referring to applying and developing mathematical and statistical algorithms (technical and methodological aspects) were excluded. The results were first extracted using Zotero, a reference management system, by storing the references from each keyword in a designated folder. The articles retained allowed the identification of other related articles, which were found using the Google Scholar search tool. Most of the literature review is founded upon the CHI conference and the FAccT conference from the Association for Computing Machinery (ACM) publisher.

CHI: Conference on Human Factors in Computing Systems - Year: 2012-2022
FAccT: Fairness, Accountability, and Transparency - Year: 2012-2022

The keywords for the article search included the following: automated (and algorithmic) decision-making systems, machine learning, rule-based algorithms, perceived algorithmic bias, human bias, and perceived algorithmic fairness/accountability/transparency.

Below we present a synthesis of the existing research related to using and adopting automated decision-making systems. This synthesis will provide background on how the perception of algorithmic type and task importance affects perception towards automated decision-making systems.

3.2.2 Rule-Based vs. Machine Learning Based Algorithms

Not all automated decision-making systems are created equal; they are bound by programs, algorithms, and functions carefully chosen by engineers and directed by managers (Lima et al., 2021; Luccioni & Bengio, 2020). Rule-based and machine learning are algorithmic types that fall under the artificial intelligence umbrella. Each algorithmic type has its advantages and weaknesses for the given problem's complexity and thought machine learning is becoming the algorithm of choice for its performance (van Ginneken, 2017).

Rule-based (RB) algorithms use numerous if-then statements to express a step-by-step method to an output (Liao, 2005; van Ginneken, 2017). Because human knowledge is used to define rules, they are known as expert systems or GOFAI (Good Old-Fashioned Artificial Intelligence) (van Ginneken, 2017). Rule-based algorithms are frequently employed due to their ease of use, adaptability, and ability to categorize extreme values well (Uzuner et al., 2009). However, they have difficulty categorizing nuances, and their ability to understand a complicated set of judgments is limited by experts' knowledge (Benbya et al., 2021; Uzuner et al., 2009; van Ginneken, 2017). As a result, in rule-based algorithms, if-then statements represent expert knowledge that is extracted and then written and implemented in programmes.

Machine learning (ML) consists of sets of algorithms that allow a computer to learn, in other words, let the computer find patterns in the data (Ayodele, 2010), solve problems, recognize and display emotions, and create an outcome in diverse domains (Benbya et al., 2021). On the one hand, programs extract knowledge from a large amount of data during the training period. On the other hand, machine learning has features that automatically learn from the data during the training process without being directly programmed by experts (Ibrahim & Abdulazeez, 2021; van Ginneken, 2017). Machine learning-based systems are often prized for their performance and capabilities to process complex unstructured problems and a large amount of information (van Ginneken, 2017). However, research shows that ML tend to misinterpret words (Uzuner et al., 2009), their decision-making process is not transparent, also called “black box”, making it difficult to understand the logic behind the output (Burrell, 2016), and can classify categories that are unethical (e.g., discrimination based on gender) (Mehrabi et al., 2021).

3.2.3 Algorithmic Type and Algorithmic Transparency

The consequences of the algorithmic opaqueness in ADMS have opened the door for research on algorithmic transparency (Eslami et al., 2019; Kizilcec, 2016). While a lack of information leads to a lack of trust in the ADMS, it turns out that the opposite is true when too much information is provided (Kizilcec, 2016). Furthermore, the algorithm becomes vulnerable to leaking the trade secrets of businesses or allowing malicious users to game the system (Eslami et al., 2019). The amount and type of information shared with decision recipients must be adjusted to shape a positive attitude when decision recipients interact with ADMS. Fine-tuning necessitates determining the optimal level of transparency. However, in order to understand what information to adjust and how, it is necessary to first understand the decision recipient's perception of the algorithms and how they are shaped by it (Eslami et al., 2019).

Leveraging the increased use of artificial intelligence in ADMS, we explore the roles of algorithmic type from the perspective of decision recipients. Although quantitative studies have shown that algorithmic type (i.e., rule-based and machine learning) were not supporting factors influencing perceived fairness and trust (Grange, 2022; Wang et al., 2020), a qualitative analysis is needed to further understand and explain these findings.

3.2.4 Human Perceptions for Highly Consequential Decisions

AI-based ADMS are becoming a business opportunity hub that is deployed into the world at a fast speed (Luccioni & Bengio, 2020). From social media to health, to justice, to logistics to military etc., today, many decisions, if not all, are either a product or a by-product of algorithms. These decisions are shifting the locus of action, choice, control, and power to algorithms away from humans (Benbya et al., 2021). As algorithms take on more complex decisions, algorithmic-driven decisions impact people's daily in a way that is beginning to cross the fine line of morality.

As pointed out by many researchers, there is a need to understand human perception and a human-centered approach to the development of responsible ADMS (Alvarado & Waern, 2018; Delecraz et al., 2022; Marabelli et al., 2021). Human perception of the use of automated decision-making systems is layered, nuanced, subject to bias, and carries different expectations (Delecraz et al., 2022; Jussupow et al., 2020; Wang et al., 2020). For example, in the context of humanitarian food allocation logistics, each stakeholder has their own definition of fairness. While some consider efficient allocation as fair, other believe that equitable or equal allocation is needed (Lee et al., 2017). Additionally, in the context of passing an online qualification, a participant who received a "fail" evaluated the algorithm less fair than a participant who received a "pass" (Wang et al., 2020). Thus, an individual's expectations and motivation are also important factors influencing the use of ADMS (Delecraz et al., 2022; Lee et al., 2017; Wang et al., 2020).

Individuals process information differently, and their judgement is inevitably subject to bias. However, bias does not result in an error; instead, working around factors influencing the perception of ADMS during the development of ADMS is key to prevent and mitigate negative consequences (Alvarado & Waern, 2018; Marabelli et al., 2021).

3.2.5 The Fairness, Accountability and Transparency (FAccT) Framework

Ethical values such as fairness, accountability, and transparency are frequently brought up in the study of people's interaction with algorithms (Shin & Park, 2019). Shin & Park's (2019) proposed the Algorithm Acceptance Model, which sets perceived fairness, accountability, and transparency (FAccT) as the antecedent to trust and use. According to their exploratory interpretive analysis from the interview and online scenario-based experiment, their findings are two-fold. First, it shows that the FAccT model is acceptable and significantly affects decision recipient satisfaction; thus, decision recipient's perception of an algorithm plays a role in their satisfaction. In addition, it found that trust plays a moderating role in the effects of FAccT on satisfaction.

3.2.5.1 FAccT Definition

People interpret, perceive, and process FAccT differently because its issues are complex and abstract concepts (Shin & Park, 2019). Consequently, there is no agreed definition of FAccT. As a starting point for our research, we employ Shin and Park's (2019) definition. Fairness in algorithms is the principle that decisions made by an algorithm should not create discriminatory or unjust consequences. The concept of transparency involves the details of the service reasoning, and of other types of data management, including sensible data. Accountability in algorithms is an essential concept for designers and managers who are responsible for the consequences or impacts an algorithmic system has on stakeholders and society. Each of the terms are interrelated and overlapping which can be further divided into subcategories, see Table 3-1.

Table 3-1 – FAccT Definition

Fairness	**Definition**
Indiscrimination	The perception that the outcome is not derived from favoritism and does not discriminate against people.
Accuracy	The perception that the source of data throughout an algorithm and its data sources should be identified, logged, and benchmarked.
Impartiality	The perception that the system follows due process of impartiality with no prejudice.

Accountability	**Definition**
Responsibility	The perception that the system requires a person in charge who should be accountable for its adverse individual or societal effects in a timely fashion.
Auditability	The perception that algorithms should be designed to enable third parties to examine and review the behaviour of an algorithm.
Equity	The perception that algorithms should be free from bias; the system must be liable for the results.

Transparency	**Definition**
Understandability	The perception that the evaluation and the criteria of algorithms used should be publicly released and understandable to people—alternatively, the extent to which information is easily comprehended.
Explainability	The perception that any outputs produced by an algorithmic system should be explainable to those affected or generate an explanation without external assistance.
Observability	The perception that algorithms should let people know how well the internal states of algorithms can be understood from knowledge of their external outputs.

3.2.6 Summary of the Literature Review

A rule-based algorithm, also called the expert system, imitate human judgments using various if-then statements that are formulated within a system to express a step-by-step method of a decision-making process (Liao, 2005; van Ginneken, 2017). This type of algorithm is constrained by experts' knowledge, so its capacity for categorizing large amounts of data and making complex decisions is limited (Benbya et al., 2021; Uzuner et al., 2009; van Ginneken, 2017). With technological advancements in artificial intelligence, new algorithm types (e.g., machine learning and natural language processing) have developed and been improved upon to process larger and more complex decisions, decreasing human involvement in the decision-making process (Ibrahim & Abdulazeez, 2021; van Ginneken, 2017. However, the new generation of algorithmic types is not without weaknesses. Their decision-making process is often hidden from the experts; it is also called "black box" (Mehrabi et al., 2021) and without human input, the decision-making process is often challenging to understand or accept by the decision recipients, which sparks a call for transparency (Eslami et al., 2019; Kizilcec, 2016).

Transparency in ADMS can raise awareness and encourage quality decision-making. However, it comes with a cost: too little transparency leads to a lack of trust in the ADMS; too much transparency discloses intellectual property for business and burdens the decision-recipient's experience (Eslami et al., 2019). Thus, the optimal level of transparency is somewhere in between and should be tailored to its context. Understanding decision recipients' perceptions of algorithms and how they are shaped is important in the development of ADMS.

Individuals process information differently, and their judgment is subject to bias. When developing ADMS, bias is a crucial factor to take into account (Delecraz et al., 2022; Lee et al., 2017; Wang et al., 2020), but it can be mitigated by understanding factors that affect how ADMS are perceived. Therefore, there is a need to focus on individuals' perceptions of ADMS (Alvarado & Waern, 2018; Marabelli et al., 2021). To find factors influencing the perception of ADMS, we employed Shin & Park's (2019) Algorithm Acceptance Model, which sets FAccT as an antecedent to the use of ADMS. This framework involves ethical values, namely fairness, accountability, and transparency (FAccT) in the study of people's interactions with algorithms.

In this study, we were interested in finding out what differences' individuals perceive between the old and the new algorithmic types, namely rule-based algorithms, and machine learning. We attempted to determine whether revealing the decision-making process was a good type of information to disclose to encourage transparency. Finally, we constructed the interview questions using the FAccT framework to learn a person's impression of the algorithmic type.

3.3 RESEARCH METHOD

We addressed our research questions using a semi-structured interview to identify and understand participants' responses toward the acceptability of ADMS. We created scenarios in the context of chatbot interviews for job recruitment. The questions showcased the importance of the decision supported by the ADMS (high vs. low degree of consequentiality) and the type of algorithm used (rule-based vs. machine learning based). The current interview was approved by our institution's research ethics committee, CER de HEC, with the number 2023-4843 (see Appendix 7 and 8). In addition, all participants signed the consent form as required.

Job interview contexts are relatable to the population at large. Thus, participants were recruited using our personal networks. The interview consisted of four scenarios in the context of a job interview using two types of algorithms (rule-based and machine learning based) to evaluate the subjects for two kinds of jobs (high and low degree of consequentiality). Then, they were informed how their interview would be assessed, rule-based followed by machine learning based.

The FAccT framework inspired the follow-up questions to help reveal their values and factors influencing their use of ADMS. Participants were first asked about their impressions and understanding regarding rule-based algorithms for low-impact jobs. We used that first part as a common starting ground and a baseline before getting into comparison exercises involving discussions of machine learning algorithms and decision-making situations involving high-impact jobs. The interview guideline is provided in Appendix 4.

The interviews were first recorded and transcribed using Zoom, a telecommunication software, and lasted between 30 minutes to one hour. The transcripts were then revised and imported into NVIVO, a qualitative analysis software for coding, and our preliminary analysis of these transcripts (discussed next) yielded 350 codes. Lastly, the codes were exported to MIRO, a whiteboard for visual representation and finalized in a data structure on Microsoft Excel.

3.3.1 Procedure

Recruitment To qualify for our study, candidates had to have already applied for a job online. This criterion is added to ensure they understand the online job application process. Two potential participants were not eligible for the study as they had only applied to jobs in person or by referral.

Pre-Interview In the pre-interview, participants were first asked to explain in their own words what an algorithm was to gauge their general understanding of that concept. If their definition was far off-topic, the definition was provided to them and ensured they understood the concept before continuing. In the pool of thirteen participants, all participants' answers were aligned with the definition and grasped the concept of the algorithm using keywords such as patterns, logic, automation, formula, calculation, code, conditions, and programming. If an individual is uncertain about an algorithm, it may lead to mistrusting it. For the case of the interview, we conceptualized an algorithm as "a set of rules that englobes a mathematical formula," inspired by Logg et al.'s (2019) definition of "a series of mathematical calculations."

Chatbot Interview Scenario Participants were asked to identify their dream job and the least preferred job to represent degrees of consequentiality, respectively. Participants were placed in a scenario where they were offered a chatbot interview for their least favorite job that used a rule-based algorithm to evaluate the interview. They were then asked questions following the FAccT framework

Repeat We repeated the chatbot interview scenario using a machine learning based algorithms.

3.3.2 Participants

Thirteen participants based in Quebec, Canada were recruited in September 2022. All of whom were acquired through personal and professional networks. The participants were between the ages of 21-32 years old, came from different occupational background (engineer, entrepreneur, consultant, student, analyst, and researcher) and all lived in Montreal, Quebec, Canada. Participant's answers became repetitive as of the tenth participant, however, we stopped the interview after the thirteenth participant to ensure there are information for backup. Reaching saturation meant that no further information (properties, dimensions, conditions, actions, or consequences) seemed to emerge during the interview. Table 3-2 below shows the participants' demographic profile.

Table 3-2 – Participant's Demographic

Demographic profile	
Age	Population
20-25	3/13
26-30	9/13
30-35	1/13
Gender	

Male	5/13
Female	8/13
Highest Level of Education	
High School Diploma	2/13
Bachelor's Degree	6/13
Master's Degree	5/13

3.3.3 Data Analysis Approach

To analyze the content and summarize findings from the semi-structured interview, a coding method from (Saldaña, 2016) and the application of grounded theory from Gerlach and Cenfetelli (2020) and Gioia et al. (2013) were followed. While coding enables rigorous analysis and interpretation of the data by arranging information in a systematic order based on shared characteristics or themes (Saldaña, 2016), a grounded theory approach helps theorize, organize, and present data (Gioia et al., 2013).

3.3.4 Data Structure

We used two coding methods: in vivo and holistic coding, that were employed to represent the first-order data structure analysis in the grounded theory's development (Gerlach & Cenfetelli, 2020). In vivo coding used one to three words drawn from the participant's language as a code and holistic coding, a preparatory approach that attempts to grasp basic themes or issues (Saldaña, 2016). In addition, capturing the different layers of information within a question was also used as a simultaneous coding method that applies several codes to each qualitative answer (Saldaña, 2016). In the first-order analysis, these codes captured participants' beliefs, thoughts, and behaviour. Appendix 5 depicts an example of codes.

In the second order analysis, we employed a thematic analysis to find patterns and group codes into themes. An affinity diagram concept organized and visualized the relationships amongst the codes within the themes and the research question. Appendix 6 groups the emergence of themes by the interpretation of the codes.

Tables 3-3 and 3-4 summarize the result of the first order and second order analysis. The complete data structure can be found in Appendix 6.

Table 3-3 – Algorithmic (RB ADMS and MLB ADMS) Characteristics

	RB ADMS	**MLB ADMS**
Understandability: the ability to make sense of the decision-making process	Participants perceive the process involved in RB ADMS as transparent because there is visibility and reasoning behind the rules.	Participants perceive the process of MLB ADMS as less transparent. They are aware of the general concept of the "black box."
Adaptability: the capability of the decision-making process to manually adjust to new information or evaluation criteria	Participants believe that RB ADMS are more adaptable than MLB ADMS for its ability to change criteria.	Participants believe that MLB ADMS are less adaptable than RB ADMS because managers have limited ability to change criteria.

Robustness: the resilience of the decision-making process to detect untruthful answers.	Participants mentioned that it is easy to predict and write what the RB ADMS want to hear, which makes it vulnerable to untruthful answers.	Participants mentioned that it is more difficult than RB ADMS to determine what the MLB ADMS want to hear, which protects itself from untruthful answers.
Input Data: the quantity and quality of information used to make a decision.	Participants feel that keywords as the only data input is not enough for quality decision-making.	Participants feel that MLB ADMS can capture more data (e.g., words, sentences, tones) to make more holistic decisions.
Bias: prejudice or discrimination during the decision-making process	Participants feel like RB ADMS cannot understand the nuances behind the meaning of words provided as an input by the decision recipients. They also are not able to judge of the characteristics that are important in hiring. Participants have a bias in believing that the experts in the industry are not capable of configuring the rules better than MLB ADMS.	Participants feel like MLB ADMS can have a historical bias because they know that data used to train models can be a vehicle for prejudices or other types of biases.

Notes:

RB ADMS: Rule-based automated decision-making systems

MLB ADMS: Machine learning based automated decision-making systems

Table 3-4 – Behavior towards RB ADMS, MLB ADMS

	RB ADMS	**MLB ADMS**
Preference: the factors that influence a participant's preference towards a decision-making process.	Participants prefer RB ADMS for their Understandability of the decision-making process and their lack of Robustness to manipulate answers to their advantage.	Participants prefer MLB ADMS for their ability to process more Input Data for a better-quality decision and for Robustness to detect untruthful answers.
Use Pattern: how the participants will interact with the system knowing which type of algorithm types used.	Given the participant's understanding of RB ADMS properties, they were less compelled to use the chatbot as intended. That means attempting to game the system for their advantage.	Given the participant's understanding of MLB ADMS properties, they were compelled to use the chatbot as intended. That means providing more candid answers.

Notes:

RB ADMS: Rule-based automated decision-making systems

MLB ADMS: Machine learning based automated decision-making systems

3.4 FINDINGS

Several patterns emerged from the semi-structured interview. When we revealed the algorithmic types in the decision-making process, participants started comparing their characteristics from their knowledge. These capabilities are defined as algorithmic characteristics which englobe participants' presumed concepts of how an algorithm works by defining what is important to them and what is worth mentioning. Interestingly, these presumptions are sometimes consciously or unconsciously biased, which will then affect how they interact with the ADMS. Consequently, these presumptions guide their behaviour toward using the ADMS.

In what follows, we elaborate on the themes and concepts that affect the use of automated decision-making systems, along with their related observations. We cited findings from the literature review and new studies to further support the data and offered propositions.

3.4.1 Algorithmic Characteristics

We defined four concepts that were brought up by participants related to the algorithmic type: the participant's comprehensiveness of algorithmic characteristics, adaptability to new information, robustness of the ADMS, and quantity and quality of data input.

3.4.1.1 **Comprehensiveness of Algorithmic Characteristic**

According to our observations, for each algorithmic type, participants tried to make sense of the decision-making process based on the degree of transparency. Many participants shared the opinion that RB ADMS (rule-based automated decision-making systems) are easier to understand because the rules of the decision-making system are visible and have supporting reasoning. Meanwhile, it is the opposite for MLB ADMS (machine learning based automated decision-making systems), where the "black box" makes it harder to understand because the rule to decision-making systems is less visible and lacks supporting reasoning behind the rules. The following table illustrate these findings:

Table 3-5 - Example of codes for Comprehensiveness of Algorithmic Characteristic

Participants Quotes	**Characteristics**
P5*: I just feel like I would be more comfortable being reviewed by an algorithm that's using synonyms* [RB ADMS]*, so I have some sort of structure to be reviewed to make sure, like it's the scenarios, maybe are appropriate, the whole system is properly set up.*	RB ADMS, comprehensiveness, evaluation criteria are quantifiable, easier to understand the decision-making process
P4*: From the machine learning [MLB ADMS], I feel like it's exclusive, like not everybody gets picked, like there's a certain similarity with all these people because I see algorithms as like things on social media, right? That it's very it's hard to understand […] A little bit more nervous because unlike the other scenario, this one is really like random. It's like I don't have any say on if I'm going to get picked or not for real, even though I might say the right thing.* *I think I feel confident and comfortable with the predetermined [RB ADMS], I feel like I have enough pieces of knowledge to give sort of the right answers and good words that might you know correlate with what they're looking for.*	MLB ADMS, comprehensiveness, harder to understand the decision-making process, rules are random, evaluation criteria are not quantifiable RB ADMS, comprehensiveness, rules are stable

3.4.1.2 **Adaptability to New Information**

We also observed the importance of adaptability, which we define as an algorithm's ability to adjust manually to new information or evaluation rules. The following table illustrate this finding:

Table 3-6 - Example of codes for Adaptability to New Information

Participants Quotes	**Characteristics**
P9: *At a certain point, you need to change the criteria of the scores because it needs to be updated from time to time and validated, so it's fair versus machine learning.*	RB ADMS, adaptability, managers must constantly update variables to meet the job requirement
P1: *It's more biased with machine learning than with predetermine rules because there are more variables. But at the same time, I don't know, because you don't have control over it, and maybe it's more interesting, and you can trust it more, but at the same time, it's also the opposite because you don't have a controller it can start having biases, so it's a question.*	MLB ADMS, adaptability, managers have no control over bias

From the above observations, we see that RB ADMS are perceived to be more adaptable because managers can manually modify its rules to adapt and update the criteria used in the decision task, while MLB ADMS are perceived to be less adaptable because managers have limited control over the rules.

It is worth mentioning that although it is important to have human control over decision criteria, some participants saw benefits in MLB ADMS being able to automatically adapt to decision rules instead of adjusting criteria for every task (which is viewed as a burden).

3.4.1.3 **Robustness of the ADMS**

Robustness is defined as the resilience of the decision-making process to detect untruthful answers. During the interviews, participants expressed the ease of predicting and writing answers when the algorithmic type is RB ADMS (vs MLB ADMS). The capability of the algorithm to protect itself from "fake" answers or gaming the system can be perceived positively or negatively by an individual. Gaming the system means to exploit the properties of the system for a desired outcome rather than using it as intended (Baker et al., 2008). The two perspectives can be seen as being centered on the good of the self and being centered on the good of the collective. The following quotes illustrate these findings:

Table 3-7 - Example of codes for Robustness of the ADMS

Participants Quotes	Characteristics
P10: *I prefer predetermined [RB ADMS]. I guess it's it seems easier to predict the answers to know what answers would give a better chance.*	RB ADMS, robustness, easier to predict answers or keywords, easier to game the system
P4: *Machine learning would be better because you could teach him to recognize when someone is being extremely fake. Whereas if it's a set of rules, it can easily bypass it by being extremely fake and just shooting keywords and getting your way through, whereas machine learning might be much harder to trick.*	MLB ADMS, robustness, recognizes uncandid answers, harder to predict answers, harder to game the system RB ADMS, robustness, easier to predict answers or keywords, easier to game the system
P5: *I feel like predetermined [RB ADMS] has more favour for myself. I just feel like I can manipulate the words that I want to use. I can manipulate the context of the situations based on experiences to fit like like a puzzle versus the unknown and the closed group.*	RB ADMS, robustness, easier to predict answers or keywords, easier to game the system

This observation is aligned with Delecraz et al., 2022; Lee et al., 2017 and Wang et al., 2020 stating that an individual's expectations and motivations are factors influencing the use of ADMS. Just like our findings, participants motivated by having control over the chatbot preferred a lack of robustness to manipulate answers (or gaming the system) to their advantage. On the other hand, participants who wished to conform with the terms of use preferred MLB ADMS for its robustness to detect untruthful answers.

3.4.1.4 **Quantity and Quality of Data Input**

Data input is defined as the quantity and quality of information used to make a decision. Participants believed using different sources of information was important to the hiring process. In RB ADMS, decision-making is limited to keywords, whereas MLB ADMS uses sources of information that go beyond keywords to make a decision, such as sentences and tone. Because MLB ADMS are perceived to have the ability to capture more information, in terms of data input, it is believed to make a more holistic decision compared to RB ADMS. The following table illustrate these findings:

Table 3-8 - Example of codes for Quantity and Quality of Data Input

Participants Quotes	Characteristics
P1: *I feel like machine learning [MLB ADMS] would be more accurate because using rules is very strict and you can like really miss some […] It's looking at the bigger picture than just like, for example, screening data. With all the previously hired*	MLB ADMS, quality of input, looks at the bigger picture, determines personality fit in the team and the job

employees and you can really know which kind of employees perform that company, and you have more variables to look at, then, just like the personality match.	RB ADMS, quality of input, less room to understand nuance in words
***P3**: It [MLB ADMS] would be better actually because you could teach him to recognize when someone is being extremely fake. Whereas if it's a set of rules, it can easily bypass it by being extremely fake and just shooting keywords and and getting your way through, whereas the machine learning might be much harder to trick*	MLB ADMS, quality of input, analyze the time and emotion in sentences, RB ADMS, quality of input, unable to analyze the tone and emotion in sentences
***P5:** I think it's [RB ADMS] a good way to interview candidates but not for the final interview process because it removes the ability to select to see fit with like the algorithm that can't really necessarily predict if it's like a personality fits do the company or the team and also removes the ability for the candidate to decide to be like the company or the position because they're speaking to other box.*	RB ADMS, quality of input, unable to determine personality fit in the team and job
***P12:** I would prefer being judged by an algorithm run by machine learning more than predetermined. Because there is more place for interpretation with machine learning to make a decision because it's based on a lot of background, for example, as I said, maybe it's difficult to explain an idea, but if you say three times to an algorithm that uses machine learning, it will be interpreted differently.*	MLB ADMS, quality of input, looks at the bigger picture,

Participants' conception of RB ADMS and MLB ADMS conforms with the literature. MLB ADMS can find patterns and learn from the data (Ayodele, 2010); these patterns can include a wide variety of data input, including emotions (Benbya et al., 2021). With RB ADMS, the amount of patterns is limited by the expert's capability (Liao, 2005; van Ginneken, 2017).

3.4.1.5 **Individual Bias**

Last, participants were concerned about the potential bias from the algorithms leading to inaccurate decision outcomes. We defined bias as algorithmic prejudice or discrimination during the decision-making process. The following table illustrates this finding:

Table 3-9 - Example of codes for Individual Bias in RB ADMS

Participants Quotes	**Characteristics**
***P5:** Predetermined [RB ADMS] removes the ability to select to see fit with like the algorithm that can't necessarily predict if it's like personality fits the company or the team and also removes the ability for the candidate to decide to like the company or the position because they're speaking to another bot.*	RB ADMS, individual bias, cannot grasp soft skills, harder for RB ADMS to measure fit (team, personality)
***P9:** For more culture fit or personality fit things that are related to soft skills that requires more of a human touch I don't think you can program it.*	RB ADMS cannot grasp soft skills
***P12:** The chatbot [RB ADMS] would just take my words as a series of characters more than the unification of it so maybe, for example as as right now, because my English is not the best, sometimes I will say things just to explain myself but because you are human you understand it, but the chat but wouldn't understand that so yeah, the idea is.*	RB ADMS cannot understand the meaning behind sentences

RB ADMS are seen as incapable of understanding nuances and other characteristics important to hire participants. However, RB ADMS or expert systems are limited by the rules experts set. Thus, contrary to the observations, RB ADMS are capable of understanding nuances and processing complex decisions if experts apply their knowledge rightfully (Benbya et al., 2021; Uzuner et al., 2009; van Ginneken, 2017). Therefore, participants have an unconscious bias toward experts configuring the rules and developing the RB ADMS algorithm.

Table 3-10 - Example of codes for Individual Bias in MLB ADMS

Participants Quotes	Characteristics
P8: *I think that it's a bit trickier with that with this one because it based itself on previously hired people. Then you're almost certain to have a homogeneous group of staff because people that have given great answers are the ones hired, and then those answers are seen as the status quo for the machine learning, plus if the results cannot be verified by the HR person at the end, then it makes the entire decision process based on AI.*	MLB ADMS, bias, historical bias in data leading to bias in a decision (stereotype in hired employees)

On the other hand, MLB ADMS are perceived to have a historical bias because the data used to train models can be a vehicle for prejudices or other types of biases. Consistent with the literature review, historical bias could lead to unconscious stereotypes in hired employees, creating a homogeneous pool of candidates that is considered unethical (e.g., discrimination based on gender) (Mehrabi et al., 2021).

Based on the above observations, we propose the following:

Proposition 1: The most salient algorithm characteristics assessed by decision recipients are characterized by comprehensiveness, adaptability, robustness, data input and bias.

3.4.2 Behaviour In Response to Algorithmic Type

Behaviour related to the algorithmic type is two-fold: algorithmic appreciation and adaptation strategy.

3.4.2.1 Preference for Algorithmic Type

Several factors influence participants' preference towards a type of decision-making process. Our observation is aligned with Delecraz et al., 2022, Lee et al., 2017 and Wang et al., 2020 stating that an individual's expectations and motivations influence the use of AaDMS. Just like our findings, participants motivated to game the system will prefer a lack of robustness to manipulate answers to their advantage. On the other hand, participants who wish to use the system as intended will prefer MLB ADMS for its robustness to detect untruthful answers. The following quotes illustrate these findings:

Table 3-11 - Example of codes for Preference for Algorithmic Type

Participants Quotes	Characteristics
P5: *I feel like predetermined [RB ADMS] has more favour for myself. I just feel like I can manipulate the words that I want to use. I can manipulate the context of the situations based on experiences to fit like like a puzzle versus the unknown and the closed group. […] I think I feel confident and comfortable with the predetermined [RB ADMS], I feel like I have enough pieces of knowledge to give sort of the right answers and good words that might you know correlate with what they're looking for.*	RB ADMS, preference, lack of robustness, understandability
P4: *Machine learning would be better because you could teach him to recognize when someone is being extremely fake. Whereas if it's a set of rules, it can easily bypass it by being extremely fake and just shooting keywords and getting your way through, whereas machine learning might be much harder to trick.*	MLB ADMS, preference, robustness to detect untruthful answer
P12: *I would prefer being judged by an algorithm run by machine learning more than predetermined. Because there is more place for interpretation with machine learning to make decisions because it's based on a lot of background, for example, as I said, maybe it's difficult to explain an idea, but if*	MLB ADMS, preference, input data for a better-quality decision

you say three times to an algorithm that uses machine learning, it will be interpreted differently.	

3.4.2.2 ADMS Usage Pattern

A participant's preconception of an algorithmic characteristic determines how the participants will interact with the ADMS. Providing the type of algorithms in a decision-making process is like giving out the rules of a game. It gives individuals a sense of control over how to perform better while providing them with an opportunity to adapt their answers.

For RB ADMS, when participants were told that the chatbot used predetermined keywords to rank their answers, all participants would use more keywords or write what the ideal solution might be expected. The following quote illustrates these findings:

Table 3-12 - Example of codes for RB ADMS Usage Pattern

Participants Quotes	**Characteristics**
P1: *A predetermined algorithm, I think I tried to analyze what they want to hear because maybe I would not get that way, and so I'd be more kind of logical about it because I know that if I say what I mean, so it's not going to fit right.*	*RB ADMS, game the system, write what they want to hear*
P 6: *It takes the pressure off one, I guess. If he was more words than you, there's more chance to get it right than the good keywords, so you'll have more chance to have the job.*	*RB ADMS, game the system, higher chance to get the job*

For ML ADMS, when participants were told that the chatbot used interview data from previously hired and ideal candidates to rank their answers, it motivates them to take the interview more seriously, put in more effort and answer with authenticity. The following quoted illustrate these findings:

Table 3-13 - Example of codes for Usage Pattern in MLB ADMS

Participants Quotes	**Characteristics**
P5: *To get the chatbot to select me over someone else, it's more unpredictable, so I guess I rather be myself, and so that will be my answer was just tried to be as much as I can, to be myself and hope that state ideal candidate.*	MLB ADMS, use as intended, be myself
P13: *I would be more at ease to just answer truthfully and completely wouldn't be thinking about what the algorithm might want to hear [...] Possibly check against the persons who were hired for the job since it would be easier for me to just be on my own skin wouldn't be too stressed*	MLB ADMS, use as intended, answer truthfully

These observations are in line with the coping model of user adaptation for IT artifacts (Beaudry & Pinsonneault, 2005; Stein et al., 2015). According to the IT coping model, the support-resistance and conformity of the terms of use lead to different use patterns such as personalization, gaming the system, being a good citizen, exercising discretion, and opting out (Stein et al., 2015). For example, individuals who resist conformity to the terms of use will either strive to game the system or opt out, whereas individuals who support conformity to the terms of use will often be good citizens by following the rules (Stein et al., 2015). Similarly, in the context of hiring, participants who prefer a RB ADMS for its lack of robustness often aim to game the system, whereas participants who prefer MLB ADMS for its robustness often aim to "answer truthfully" by being a good citizen.

Based on the above, we propose the following:

Proposition 2: Decision recipients are more likely to attempt gaming the system when interacting with a RB ADMS than when interacting with a MLB ADMS.

3.4.3 Summary of Findings

Figure 3-1 illustrates the emerging themes and concepts that emerged from the data analysis.

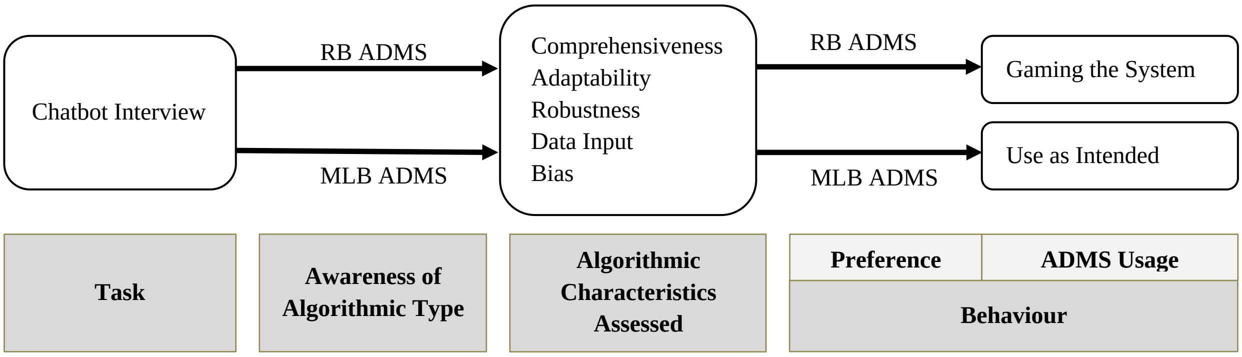

Figure 3-1 – Responses towards Algorithmic Types

3.5

3.6 DISCUSSION AND CONTRIBUTION

The discussion below connects the findings from the study with the research questions and the FAccT framework.

Our research shows that sharing the algorithmic type of the decision-making process tests participants' understanding of how specific algorithms work and their characteristics. This highlights the importance of an individual's perception of ADMS. Their general understanding of algorithm type provides a foundation for how they perceive it and how they interact with ADMS. Five algorithmic characteristics were brought up by participants, which we considered as important for the context of hiring using a chatbot: understandability, adaptability, robustness, data input, and bias. A participant's knowledge of algorithmic type is based on the following: how easy it is to understand the decision-making process (transparent or "black box"), its flexibility to adjust to new or unique information, its ability to withstand manipulation from decision recipients, the amount of information it takes to make decisions, and any prejudice the algorithm or the participant have in the decision-making.

Algorithmic characteristics serve as an a priori for their preference and use pattern. Indeed, an individual's knowledge and motivation generate different opinions and behaviours. For instance, if an algorithm is easy to understand, the algorithmic type can enable participants to game the system. Whereas, if an algorithm is hard to understand, the algorithmic type can promote the terms of use. For example, in rule-based, participants will use as many keywords as possible, whereas in machine learning, participants will do more background research on hired employees. Depending on the type of algorithm and the terms of use, the algorithmic type can be a gateway to help decision recipients make informed decisions while maintaining the integrity of the terms of use.

There is a difficult line to cross between transparency and protecting intellectual property, preventing gaming the system, and maintaining a quality experience (Burrell, 2016; Eslami et al., 2019). As seen from the findings, providing the algorithmic type provides only a part of the decision-making process; the decision recipients fill in the missing information based on their assumptions. It can positively or negatively impact the use of ADMS because an individual's knowledge or assumptions of algorithms can be biased. To continue the journey to fine-tune the type of information for a positive attitude towards using ADMS, given the individual's bias and motivation towards the algorithm, we suggest a different approach; rather than aiming for a transparent algorithmic decision-making process, we suggest shifting the focus away from the algorithm and put more emphasis on the possible use patterns. We propose a transparent use of decision-making systems by providing and offering information on how to effectively employ ADMS for the best results. For example, prior to using ADMS, managers can share important points on what constitute a good quality answer to decision recipients, such as drafting complete sentences or even suggesting tools for corrections for informed decision-making.

Finally, in the chatbot hiring context, we used the FAccT framework (Shin & Park, 2019) to develop the interview questionnaire. Because we emphasized the issues of fairness, accountability, and transparency during the interview, participants were able to better express themselves and explain their points of view on subjects they would not have considered otherwise. FAccT provided a useful foundation for delving into areas that the general public is unaware of but should consider for this study. Overall, using the FAccT framework for ADMS acceptability is useful to the framing of the interview questions and was a useful model for finding topics that lead to the acceptability of ADMS. Individuals' perceptions of ADMS are based on their knowledge of the subject, which can be biased at times. This demonstrates the importance of researchers, managers, and engineers providing adequate information to decision recipients for them to make informed decisions. As part of positively promoting the use of ADMS, we propose providing essential information about algorithmic characteristics to decision recipients to ensure a basic homogeneous understanding. By focusing on the individual's perspective to foster responsible use of innovative technologies for decision-making, this position contributes to the algorithmic experience and algorithmic transparency within the IS and HCI community.

3.6.1 LIMITATIONS AND FUTURE RESEARCH

The limitations of this study should be acknowledged. First, we employed chatbots in the context of the hiring process to generalize and emphasize high level of consequential automated decision-making systems. Because the scenarios were hypothetical, new responses may be generated by live interaction with ADMS. As a result, future researchers might put the same study to the test using an existing system. Furthermore, the interview is confined to a small sample size of participants drawn from a network of people who have previously applied for jobs online. As a result, the findings are constrained to the perspective of one group, which may not be indicative of society as a whole. Bringing in the perspectives of other stakeholders, such as regular citizens, business/HR managers, and software developers, will provide a more holistic perspective on the use of ADMS. Finally, to improve the accuracy of data interpretation, the data analysis segment could have been strengthened with another researcher's perspective.

3.7 CONCLUSION

To understand the implications of algorithmic opacity in a human-centered approach, this study explores the algorithmic characteristics that influence decision recipients' behaviour towards the use of ADMS using a semi-structured interview. We found that algorithmic type can influence how a decision recipient completes a task. Decision recipients use their preconceptions of algorithms to evaluate algorithmic types, and with that information, they adjust their ADMS usage pattern. However, we found that an individual's preconception of algorithms can be biased, which has unintended implications. We recommend concentrating on transparent usage guidelines for a more responsible implementation of transparency.

Chapter 4: Conclusion

4.1 Summary of Research Objectives and Main Results

The main goal of this thesis is to determine what factors influence the social acceptability of ADMS from the perspective of the decision receiver. The research questions were:

RQ 1-1 What factors drive the acceptability of automated decision-making systems by detecting factors affecting the algorithmic experience from the decision recipient's perspective?

RQ 2-1 What algorithmic characteristics do decision recipients perceive as important in ADMS?

RQ 2-2 What is the effect of algorithmic type in the use of ADMS?

Overall, we observed that moral principles, expectations, and self-interest are factors driving the decision recipient's acceptability of ADMS and that revealing algorithmic type revealed important algorithmic characteristics such as understandability, adaptability, robustness, and data input. However, that information is limited to the individual's knowledge. Furthermore, an individual's preference and how they use the ADMS are influenced by their knowledge of algorithmic characteristics.

4.2 Contributions

The study's findings emphasize the importance of an individual's perception and role in the success of ADMS. An individual's perception of ADMS is based on their knowledge, which can be biased and lead to negative consequences for the economy and society. However, as we discovered in the studies, everyone has their own perception of what is acceptable and what algorithms can and cannot do. This bias is both conscious and unconscious at times. This data is significant because it emphasizes the importance of algorithmic transparency and education as part of the problem and solution.

For managers intending to implement ADMS in which the outcome directly affects decision recipients, we recommend including their opinion into consideration during the development process of ADMS to ensure the values are aligned. We believe that disclosing algorithmic type might be important information to promote algorithmic transparency only when the algorithm involves machine learning. We also recommend that managers give meaningful guidance on how decision recipients should do their tasks in order to encourage acceptable behaviour.

FAccT provided a useful foundation for delving into areas that the public is unaware of but should consider for this study. Overall, using the FAccT framework for ADMS acceptability is useful to the framing of the interview questions and was a useful model for finding topics that lead to the acceptability of ADMS. Additionally, the FAccT framework for ADMS acceptability helped frame the interview questions and was a useful model for finding topics that lead to the acceptability of ADMS.

Finally, most HCI and ADMS literature studies in information science utilizes quantitative research methods to test hypotheses. People's attitudes toward ADMS shift over time as they interact with it more. As a result, understanding people's perspectives is equally vital in order to uncover new facts. As a result, we expect that this thesis will add to current studies and explain their findings.

4.3 Limitations and Future Research

The limitations of this study should be acknowledged. First, we employed chatbots in the context of the hiring process and automated university admission to generalize a high level of consequentiality in automated decision-making systems. Because the scenarios were hypothetical, new responses may be generated by live interaction with an ADMS system. As a result, future researchers might put the same study to the test using an existing system. Additionally, it is possible that not all answers were properly understood or well interpreted; thus, the data analysis process could be improved with an additional pair of eyes.

Furthermore, because secondary data was employed, the quality of the questions and answers are limited. Setting an open-ended question at the end of the online scenario-based experiment could determine how much attention the question and answer receive. More quality control is thus required for future online scenario-based experiments with open-ended questions.

Moreover, the interview is confined to a small sample size of participants drawn from a network of people who have previously applied for jobs online. As a result, the conclusions are limited to the viewpoint of one group, which may not be representative of society as a whole. Including the opinions of additional stakeholders, such as ordinary citizens, business managers, and software engineers, will provide a more comprehensive view of ADMS's utilization. Finally, the data analysis segment may have been improved using another researcher's perspective to improve the accuracy of data interpretation.

4.4 Final Thoughts

Overall, this thesis helped me better understand the complex relationship that individuals have with ADMS. I witnessed participants expressing concerns first hand and was also reminded of the need to stay open-minded and unbiased in order to properly understand their underlying values. Qualitative analysis is a key approach for digging deeper into problems and discovering the unknown. However, due to the unstructured nature of the information, organizing it has been more difficult than anticipated.

In fact, the qualitative analysis was, the biggest challenge I faced in completing this thesis, but it is this challenge that taught me not to give up trying. It often takes a few tries before you can come up with a good concept, and with each attempt, you get better. Although the process was difficult, it was also incredibly rewarding, and for that I am thankful.

Bibliography

A guide to the best AI chatbot. (2023). Zendesk. https://www.zendesk.com/service/messaging/chatbot/

Alvarado, O., & Waern, A. (2018). Towards algorithmic experience: Initial efforts for social media contexts. *Proceedings of the 2018 Chi Conference on Human Factors in Computing Systems*, 1–12.

Amazon scraps secret AI recruiting tool that showed bias against women. (2018, October 10). *Reuters*. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Araujo, T., De Vreese, C., Helberger, N., Kruikemeier, S., van Weert, J., Bol, N., Oberski, D., Pechenizkiy, M., Schaap, G., & Taylor, L. (2018). Automated decision-making fairness in an AI-driven world: Public perceptions, hopes and concerns. *Digital Communication Methods Lab*.

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY*, *35*(3), 611–623. https://doi.org/10.1007/s00146-019-00931-w

Ayodele, T. O. (2010). Types of machine learning algorithms. *New Advances in Machine Learning*, *3*, 19–48.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why students engage in “gaming the system” behavior in interactive learning environments. *Journal of Interactive Learning Research*, *19*(2), 185–224.

Beaudry, A., & Pinsonneault, A. (2005). Understanding user responses to information technology: A coping model of user adaptation. *MIS Quarterly*, 493–524.

Benbya, H., Pachidi, S., & Jarvenpaa, S. L. (2021). Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. *Journal of the Association for Information Systems*, *22*(2), 281–303. https://doi.org/10.17705/1jais.00662

Bodea, G., & Karanikolova, K. (2018). *Automated decision-making on the basis of personal data that has been transferred from the EU to companies certified under the EU-U.S. Privacy Shield*. European Commission. https://ec.europa.eu/info/sites/default/files/independent_study_on_automated_decision-making.pdf

Bohr, A., & Memarzadeh, K. (2020). Chapter 2—The rise of artificial intelligence in healthcare applications. In A. Bohr & K. Memarzadeh (Eds.), *Artificial Intelligence in Healthcare* (pp. 25–60). Academic Press. https://doi.org/10.1016/B978-0-12-818438-7.00002-2

Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 2053951715622512.

De-Arteaga, M., Fogliato, R., & Chouldechova, A. (2020). A case for humans-in-the-loop: Decisions in the presence of erroneous algorithmic scores. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12.

Delecraz, S., Eltarr, L., Becuwe, M., Bouxin, H., Boutin, N., & Oullier, O. (2022). Responsible Artificial Intelligence in Human Resources Technology: An innovative inclusive and fair by design matching algorithm for job recruitment purposes. *Journal of Responsible Technology*, *11*, 100041. https://doi.org/10.1016/j.jrt.2022.100041

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, *59*(2), 56–62.

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

Eslami, M., Vaccaro, K., Lee, M. K., Elazari Bar On, A., Gilbert, E., & Karahalios, K. (2019). User Attitudes towards Algorithmic Opacity and Transparency in Online Reviewing

Platforms. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3290605.3300724

Friedman, B., Kahn, P. H., & Borning, A. (2008). Value Sensitive Design and Information Systems. In *The Handbook of Information and Computer Ethics* (pp. 69–101). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470281819.ch4

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, *16*(1), 15–31.

Government of Canada. (2019, February 5). *Directive on Automated Decision-Making*. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592#appA

Grange, C. (2022). Algorithmically Controlled Automated Decision-Making and Societal Acceptability: Does Algorithm Type Matter? *HICSS*, 1–10.

Hoven, J. van den. (2013). Value Sensitive Design and Responsible Innovation. In *Responsible Innovation* (pp. 75–83). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch4

Ibrahim, I., & Abdulazeez, A. (2021). The role of machine learning algorithms for diagnosing diseases. *Journal of Applied Science and Technology Trends*, *2*(01), 10–19.

Jussupow, E., Benbasat, I., & Heinzl, A. (2020, June 28). WHY ARE WE AVERSE TOWARDS ALGORITHMS? A COMPREHENSIVE LITERATURE REVIEW ON ALGORITHM AVERSION. *Proceedings of the 28th European Conference on Information Systems (ECIS)*.

Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2390–2395.

Koelle, M., Olsson, T., Mitchell, R., Williamson, J., & Boll, S. (2019). What is (un)acceptable? Thoughts on social acceptability in HCI research. *Interactions*, *26*(3), 36–40. https://doi.org/10.1145/3319073

Koenig, A. (2020). The Algorithms Know Me and I Know Them: Using Student Journals to Uncover Algorithmic Literacy Awareness. *Computers and Composition*, *58*, 102611. https://doi.org/10.1016/j.compcom.2020.102611

Kordzadeh, N., & Ghasemaghaei, M. (2021). Algorithmic bias: Review, synthesis, and future research directions. *European Journal of Information Systems*. https://doi.org/10.1080/0960085X.2021.1927212

Lee, M. K., Kim, J. T., & Lizarondo, L. (2017). A Human-Centered Approach to Algorithmic Services: Considerations for Fair and Motivating Smart Community Service Management that Allocates Donations to Non-Profit Organizations. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3365–3376. https://doi.org/10.1145/3025453.3025884

Liao, S.-H. (2005). Expert system methodologies and applications—A decade review from 1995 to 2004. *Expert Systems with Applications*, *28*(1), 93–103.

Lima, G., Grgić-Hlača, N., & Cha, M. (2021). Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3411764.3445260

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.

Luccioni, A., & Bengio, Y. (2020). On the morality of artificial intelligence [Commentary]. *IEEE Technology and Society Magazine*, *39*(1), 16–25.

Marabelli, M., Newell, S., & Handunge, V. (2021). The lifecycle of algorithmic decision-making systems: Organizational choices and ethical challenges. *Journal of Strategic Information Systems*, *30*(3), 101683. https://doi.org/10.1016/j.jsis.2021.101683

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, *54*(6), 1–35.

Park, H., Ahn, D., Hosanagar, K., & Lee, J. (2021). Human-AI Interaction in Human Resource Management: Understanding Why Employees Resist Algorithmic Evaluation at Workplaces and How to Mitigate Burdens. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3411764.3445304

Payani, A. (2023). *Council Post: Embracing The Future: How AI Is Revolutionizing Marketing And Sales*. Forbes. https://www.forbes.com/sites/forbesbusinesscouncil/2023/03/08/embracing-the-future-how-ai-is-revolutionizing-marketing-and-sales/

Pethig, F., & Kroenung, J. (2022). Biased humans,(un) biased algorithms? *Journal of Business Ethics*, 1–16.

Qayyum, A., Qadir, J., Bilal, M., & Al-Fuqaha, A. (2020). Secure and robust machine learning for healthcare: A survey. *IEEE Reviews in Biomedical Engineering*, *14*, 156–180.

Rutjes, H., Willemsen, M. C., & IJsselsteijn, W. A. (2019). Beyond Behavior: The Coach's Perspective on Technology in Health Coaching. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3290605.3300900

Ryan, M. (2020). The Future of Transportation: Ethical, Legal, Social and Economic Impacts of Self-driving Vehicles in the Year 2025. *Science and Engineering Ethics*, *26*(3), 1185–1208. https://doi.org/10.1007/s11948-019-00130-2

Saldaña, J. (2016). *The coding manual for qualitative researchers* (2nd ed.). SAGE.

Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284.

Stein, M.-K., Newell, S., Wagner, E., & Galliers, R. (2015). Coping with Information Technology: Mixed Emotions, Vacillation, and Nonconforming Use Patterns. *MIS Quarterly*, *39*, 367–392. https://doi.org/10.25300/MISQ/2015/39.2.05

Tabourdeau, G., & Grange, C. (2020). *From User Acceptance to Social Acceptance*.

Uzuner, Ö., Zhang, X., & Sibanda, T. (2009). Machine learning and rule-based approaches to assertion classification. *Journal of the American Medical Informatics Association*, *16*(1), 109–115.

van Ginneken, B. (2017). Fifty years of computer analysis in chest imaging: Rule-based, machine learning, deep learning. *Radiological Physics and Technology*, *10*(1), 23–32.

Wang, R., Harper, F. M., & Zhu, H. (2020). Factors Influencing Perceived Fairness in Algorithmic Decision-Making: Algorithm Outcomes, Development Procedures, and Individual Differences. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376813

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, *52*, 113–117.

Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review*, *96*(4), 114–123.

Winder, D. (2019). *Forget Fake News, Deepfake Videos Are Really All About Non-Consensual Porn*. Forbes. https://www.forbes.com/sites/daveywinder/2019/10/08/forget-2020-election-fake-news-deepfake-videos-are-all-about-the-porn/

Appendix 1: Coding Sample (article 1)

Table 3-14 – Example of Codes

Qualitative Answers	**Codes**
I would pay more for the advice of an accredited person than some computer coldly assessing me based on some list instead of the actual person that I am, which another person can see.	S3, algo_cold, human_experience, human_see
I wouldn't want a computer. I would pay the same for a person than for a computer. I would wish the admission person to look at my application just based on criteria, not the history of other students, because that does not demonstrate anything.	S1, data_inaccurate, human_rules
The computerized system seems less likely to have a bias	S4, algo_less_biased
I'm low income	S4, no_explanation
Because the application process should be free for college	S4, cost_free

Appendix 2: Coding by Themes (article 1)

Table 3-14 – Raw Codes by Themes

	Decision-Making Agent	
Themes	**ADMS**	**HDM**
Moral Principle in Algorithmic Judgement	• No Moral • No Rational • Decision Insulting • Don’t See Human Beings • Do Not Understand Emotions • Do Not Understand Uniqueness • Cold	• Worth • Moral Responsible • Unworthy Judging Application • Worth Human Time • Morals • Acknowledge • Rational • Human Touch • See
Expectation of Subjectivity in ADMS	Positive • Consistent • Unbiased • Accurate • Convenient • Fast • Safe • Easy • Value for money Negative • Lacks Subjective Standard • Sort uniqueness • Inaccurate • Lacks Nuance • Lacks Human Factors • No individual bias • Do not lie • Inflexible • No Feedback • No Lived Experience • No Special Circumstances • Historical Bias • Vulnerable To Mistakes • Detrimental To Gender • Misinterpret	Positive • Emotions • Empathy • Care • Human Element • Human Uniqueness • Skills/Knowledge • Circumstances • Customizable • Help • Flexible • Read The Room • Understand • Experience • Holistic With Great Intent Negative • Bias • Error • Favouritism
Favourability of the Decision Outcome	• Self Interest • Higher Success	
Human-ADMS relationship	• Human Check Algo	
Attitude towards ADMS	• Not Sold • Fair/Unfair • If consider uniqueness • Distrust • Reliable/Unreliable	• Fair Decision • Trustworthy • Trust • Credible

Appendix 3: Coding by Decision-Making Agent (article 1)

Table 3-15 – Unique Codes Supporting Human Authority

Supporting Human Authority			
Human Decision-Making		**Human Experience**	**Human Rules**
• Fair Decision • Emotions • Skills • Acknowledge • Customizable • Trustworthy • Morals • Knowledge • Review • Care • Rational • Read The Room • Check Algo • Human Element • Human Uniqueness	• Self Interest • Higher Success • Touch • See • Circumstances • Time • Help • Worth • Trust • Understand • Holistic • With Great Intent • Empathy • Contact • Over Algo Decision	• Experience • Uniqueness • Flexible • Holistic • Rules • Emotions • Decision • Credible	• Moral Responsible

Table 3-16 – Unique Codes Undermining Human Authority

Undermining Human Authority	
Human Decision-Making	**Human Experience**
• Bias • Error	• Bias • Error • Favouritism • Unworthy Judging Application

Table 3-17 – Unique Codes Supporting Algorithm Authority

Supporting Algorithm Authority			
Algorithm Decision-Making		**Machine Learning Algorithm**	**Rules-Based Algorithm**
• Do not lie • Convenient • Fast • Safe • Reliable • Accurate • Sort uniqueness	• Consistent • Easy • No individual bias • Transparency • Value for money • Unbiased	• If consider uniqueness	• Fair • Unbiased • Fast

Table 3-18 – Unique Codes Undermining Algorithm Authority

Undermining Algorithm Authority			
Algorithm Decision-Making		**Machine Learning Algorithm**	**Rule-Based Algorithm**
• No Moral • Lacks Human Factors • No Rational • Lacks Nuance • Decision Insulting • Unfair • Do Not Understand Uniqueness • Don’t See Human Beings Biased • Cold • Do Not Understand Emotions • Misinterpret	• Not Sold • Inaccurate • Distrust • No Feedback • No Lived Experience • No Special Circumstances • Detrimental To Gender • Unreliable • Vulnerable To Mistakes	• Bias • Historical Bias	• Inaccurate • Inflexible • Lacks Subjective Standard

Appendix 4: Interview Procedure (article 2)

Demographic
Anyone who applied to jobs.

Verbatim
Welcome <participant name>, and thank you for accepting our invitation. My name is Ying, and we will spend about 45 minutes together discussing your experience with the software. This interview is developed for my master's thesis at HEC Montreal.

You have been selected to participate in this study because we are interested in your opinion and feedback on your experience. In other words, you are not being evaluated, and there are no right or wrong answers. What is important is that you share your opinion as much as possible. Also, be very comfortable if there are things you don't understand or are unclear about. You will see that I am following a script, that is, to ensure I provide consistent information to all participants.
We'll start our session today with a series of pre-interview questionnaires followed by questions for you to share your impressions, preferences, and experience. Before we begin, we'll review a few things.
You'll notice that the discussion is being recorded. Your comments will remain anonymous - used only by the research team. You have already signed the consent form; however, we also need your verbal consent. First, do you consent to participate in this study? Do you agree to the recording of video, sound, and screen?

Do you have any questions at this point?

Great. We will begin the study.

Personal Information
Age:
Gender:
Ethnicity:
Year of graduation:
School:
Program:
Explain in your own words what is an algorithm: (a set of rules that englobes mathematical formulas)

Let's talk about jobs. Imagine you have a list of jobs ranked from best to worst. Can you tell me what job is at the top of the list? What job is at the bottom of the list?

(Here, the top list refers to a dream job, and the bottom is a job that you have as a part-time, a temporary job or a job that you won't enjoy as much)

Let's consider this scenario. Say you are looking for a job and found an opening for (your worst job), decided to apply anyway. You passed the first screening process. To proceed to the next step, you must complete a chat and video-based interview process via a chatbot. How it works is that a bot will ask you two sets of interview questions to assess your traits and communication skills and score you according to your fit for the role. (e.g., What is your name and contact information, what skill or experience do you have that makes you a great candidate for this role?). The best candidate will then advance to a non-ai video interview for the HR team to review, and the best candidate will receive an offer.

The chatbot assesses your traits and skills based on predetermined rules that filter and assigns your answer to a fit score. Therefore, based on predefined rules or criteria (e.g., skills in communications or teamwork), the algorithm will go through your interview answer, look for keywords and provide a fit score. The predetermined algorithm involves purely mathematical formulas which allows hiring personnel to clearly understand how a decision is derived.

Transparency

1. How would you feel about applying for (worst job) with a chatbot that uses predetermined algorithms?
2. In your opinion, why does this Company use chatbot interviews?
3. Based on your understanding, how are candidates being evaluated? (Expecting criteria and type of algorithm used to make decision)
4. How would you feel if the chatbot uses machine learning instead of predetermined rules to rank your traits and skills? Unlike predetermined algorithms, machine learning does not use predefined rules, instead it uses data from previously hired or ideal candidates to learn and predict where your traits stand. The exact decision process is invisible and impossible to understand or explain.
5. How would you do the chat interview differently if it was for (your dream job)? Does the type of algorithm matter in your approach?

Fairness (a decision with no discrimination, favouritism, and bias. Candidates are treated equally)

1. Let's talk about fairness. In your opinion, do you think using chatbot interview data is relevant for ranking candidates? Why or why not?
2. Will this process favour a particular type of candidate over others?
3. Do you think this process of ranking candidates is fair? Why or why not?
4. In terms of fairness, how would you feel differently if the chatbot uses machine learning prediction (opaque explanation) rather than rules (clear explanation)?
    1. If both algorithms were fair…
5. How would you do the interview differently for (your dream job)?

Accountability (someone takes responsibility if problems arise from the algorithms)

1. Let's talk about accountability. What will you do if you disagree with this interview process?
2. How would you feel about using the automatic chatbot interview if it was examined and reviewed by a third independent group (government, non-profit, ethics, associations)?
3. Regarding accountability, how would you feel differently if the chatbot uses machine learning prediction rather than a prediction?
4. How would you do the interview differently for (your dream job)?

Appendix 5: Coding Sample (article 2)

Table 3-19 – Example of Codes

Qualitative Answers for Q1	**Codes**
P1: I probably find it interesting because I've never applied for a job in that, like, I've never had this kind of experience with a chatbot, so I find it at least challenging. Even though I don't necessarily want a job I try to like do my best and and really put in the effort to see if I can pass this screening process.	Interesting, new experience, challenging, do my best, test my skill
P3: I think I feel confident and comfortable because I feel like I have enough pieces of knowledge to give sort of the right answers and good words that might you know correlate with what they're looking for.	Confident and comfortable, have the skills, adapt to algorithm
P10: I believe, like for that type of position [gaRB ADMSage man] it's probably unfair maybe because it's not something [chatbot] that I can actually show my skill. Let's say like a job that requires me to like lift heavyweights that I cannot actually show it with words, so actually have to like manually or like physically prove it, so I think in that situation, maybe if I'm actually a good fit for the job, maybe I'll feel that it's unfair and the way because it doesn't really actually evaluate the light, you can actually it cannot capture like how of a good fit I'll be for the job.	Unfair, words do not measure physical strength, writing skills unimportant in physical jobs

Appendix 6: Data Structure (article 2)

Table 3-20 – Data Structure

<u>Algorithmic type</u>: Rule based and machine learning	RB ADMS	MLB ADMS
<u>Comprehensiveness</u>: the ability to make sense of the decision-making process	Participants perceive the process involved in RB ADMS as transparent because there is visibility and reasoning behind the rules. · Easier to understand the decision-making process · The evaluation criteria are quantifiable (less bias) · The rules are stable (same rules for each person or job)	Participants perceive the process of MLB ADMS as less transparent. They are aware of the general concept of the "black box." · Harder to understand the decision-making process (black box) · The evaluation criteria are not quantifiable (possibility of using irrelevant information in decision-making) · The rules are random (changes for each job)
<u>Adaptability</u>: the capability of the decision-making process to adjust to new information or evaluation criteria	Participants believe that RB ADMS are more adaptable than MLB ADMS for its ability to change criteria: · Managers must constantly update variables to meet the job requirement	Participants believe that MLB ADMS are less adaptable than RB ADMS because managers have limited ability to change criteria: · Managers have no control over bias in MLB ADMS
<u>Robustness</u>: the resilience of the decision-making process to detect untruthful answers.	Participants mentioned that it is easy to predict and write what the RB ADMS want to hear, which makes it vulnerable to untruthful answers: · Easier to predict answers or keywords · Easier to game the system	Participant mentioned that it is more difficult than RB Participants mentioned that it is more difficult than RB ADMS to determine what the MLB ADMS want to hear, which protects itself from untruthful answers: · Recognizes uncandid answers · Harder to predict answers · Harder to game the system
<u>Input Data</u>: the quantity and quality of information used to make a decision.	Participants feel that keywords as the only data input is not enough for quality decision-making: · Unable to analyze the tone and emotion in sentences · Unable to determine personality fit in the team and job · There is less room for the grey zone to understand nuance in words	Participants feel that MLB ADMS can capture more data (e.g., words, sentences, tones) to make more holistic decisions: · Looks at the bigger picture, such as analyzing the text instead of just using words · Analyzes the tone and emotion in sentences · Determines personality fit in the team and the job
<u>Bias</u>: prejudice or discrimination during the decision-making process	Participants feel like RB ADMS cannot understand the nuances behind the meaning of words provided as an input by the decision recipients. They also are not able to judge of the characteristics that are important in hiring: · RB ADMS cannot grasp soft skills · Harder for RB ADMS to measure fit (team, personality) · RB ADMS cannot understand the meaning behind sentences Participants have a bias in believing that the experts in the industry are not capable of configuring the rules better than MLB ADMS.	Participants feel like MLB ADMS can have a historical bias because they know that data used to train models can be a vehicle for prejudices or other types of biases: · Historical bias in data leading to bias in a decision (stereotype in hired employees)

Notes:

RB ADMS: Rule-based automated decision-making systems

MLB ADMS: Machine learning based automated decision-making systems

Appendix 7: Certificate of Ethics Approval

HEC MONTRĒAL

Comité d'éthique de la recherche

May 30, 2022

To the attention of:
Ying Hui Gao
HEC Montréal

Re: Ethics approval of your research project

Project No.: 2023-4843

Title of research project: Drivers Facilitating Acceptability in Automated Decision Making Systems (ADMs)

Funding source : CRSH : 32 153 300 26 R2677

Bonjour,

Your research project has been evaluated in accordance with ethical conduct for research involving human subjects by the Research Ethics Board (REB) of HEC Montréal.

A Certificate of Ethics Approval attesting that your research complies with HEC Montréal's *Policy on Ethical Conduct for Research Involving Humans* has been issued, effective May 30, 2022. This certificate is **valid until May 01, 2023.**

In the current context of the COVID-19 pandemic, you must ensure that you comply with the directives issued by the Government of Quebec, the Government of Canada and those of HEC Montréal in effect during the state of health emergency.

Please note that you are nonetheless required to renew your ethics approval before your certificate expires using Form *F7 – Annual Renewal*. You will receive an automatic reminder by email a few weeks before your certificate expires.

When your project is completed, you must complete Form *F9 – Termination of Project*. (or *F9a – Termination of Student Project if certification is under the supervisor's name).* **All students must complete an F9 form to obtain the "Attestation d'approbation complétée" that is required to submit their thesis/master's thesis/supervised project.**

If any major changes are made to your project before the certificate expires, you must complete Form *F8 – Project Modification.* .

Under the *Policy on Ethical Conduct for Research Involving Humans*, researchers are responsible for ensuring that their research projects maintain ethics approval for the entire duration of the research work, and for informing the REB of its completion. In addition, any significant changes to the project must be submitted to the REB for approval before they are implemented.

You may now begin the data collection for which you obtained this certificate.

We wish you every success in your research work.

REB of HEC Montréal

HEC MONTRĒAL

Comité d'éthique de la recherche

CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains et qu'il satisfait aux exigences de notre politique en cette matière.

Projet # : 2023-4843

Titre du projet de recherche : Drivers Facilitating Acceptability in Automated Decision Making Systems (ADMs)

Chercheur principal :
Ying Hui Gao,
HEC Montréal

Directeur/codirecteurs :
Camille Grange
Professeur - HEC Montréal

Date d'approbation du projet : May 30, 2022

Date d'entrée en vigueur du certificat : May 30, 2022

Date d'échéance du certificat : May 01, 2023

Maurice Lemelin
Président
CER de HEC Montréal

Signé le 2022-05-31 à 16:29

Exporté le 2022-12-15 20:57 par Gao, Ying Hui — CODE DE VALIDATION NAGANO: hec-333c5f57-faba-44dd-ad04-ff28b40dfe69https://cer.hec.ca/verification/hec-333c5f57-faba-44dd-ad04-ff28b40dfe69

Appendix 8: Attestation of Completed Ethics Approval Form

HEC MONTRĒAL

Comité d'éthique de la recherche

ATTESTATION D'APPROBATION ÉTHIQUE COMPLÉTÉE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet des approbations en matière d'éthique de la recherche avec des êtres humains nécessaires selon les exigences de HEC Montréal.

La période de validité du certificat d'approbation éthique émis pour ce projet est maintenant terminée. Si vous devez reprendre contact avec les participants ou reprendre une collecte de données pour ce projet, la certification éthique doit être réactivée préalablement. Vous devez alors prendre contact avec le secrétariat du CER de HEC Montréal.

Projet # : 2023-4843 **-** Drivers of ADMS Preference

Titre du projet de recherche : Decision Recipients' Perspective of Algorithmic Types Towards the Use of Automated Decision-Making Systems : A Qualitative Analysis

Chercheur principal :
Ying Hui Gao

Directeur/codirecteurs :
Camille Grange

Date d'approbation initiale du projet : May 30, 2022

Date de fermeture de l'approbation éthique : December 15, 2022

Maurice Lemelin
Président
CER de HEC Montréal

Signé le 2022-12-15 à 21:20

NAGANO www.nagano.ca Attestation of completed ethics approval - End of project
Comité d'éthique de la recherche - HEC Montréal 1 / 1