7-6-2023

Unveiling AI Aversion: Understanding Antecedents and Task Complexity Effects

Md Jabir Rahman

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation
https://digitalcommons.memphis.edu/etd/3068

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.
UNVEILING AI AVersion: UNDERSTANDING ANTECEDENTS AND TASK COMPLEXITY EFFECTS

by

Md Jabir Rahman

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Major: Business Administration

The University of Memphis
August 2023
DEDICATION

To my wonderful wife Samira Rahman and my beautiful daughter Rufaida Amatur Rahman.
ACKNOWLEDGEMENTS

I express my deepest gratitude to my esteemed Ph.D. advisor, Dr. Huigang Liang, for his unwavering support and guidance throughout my Ph.D. journey. Dr. Liang has consistently demonstrated his dedication to my academic growth and has played a pivotal role in shaping my research. Working under his mentorship has been an invaluable learning experience for me. From him, I have acquired the ability to approach problems as opportunities and take systematic and logical steps to develop innovative solutions. I am profoundly grateful to him for his patience, especially when dealing with the errors in my drafts, continuous encouragement, and provision of invaluable suggestions and guidance. Dr. Liang went above and beyond what was expected, genuinely embodying the role of a supportive supervisor. He is undeniably one of the most outstanding scholars and mentors I have encountered in my academic career.

I also extend my heartfelt appreciation to all the members of my dissertation committee for their valuable comments and feedback, which have significantly enhanced the quality of my work. I am particularly thankful to Dr. Mark Gillenson for his unwavering support and encouragement, Dr. Joseph Zhang for his kind words and encouragement, generously devoting his time to my research, and Dr. Yinan Yu for her valuable time and support.

Furthermore, I am sincerely grateful to my colleagues and all the faculties and members of the Department of Business Information and Technology. Their support and collaboration have been instrumental in my academic growth. I especially thank Dr. Sandra Richardson, Dr. Srikar Velichety, and Ms. Christine Smith for their constant support and assistance.

I must also acknowledge the immeasurable support from my parents, whose unwavering support and prayers have been a constant source of strength. I am deeply thankful to my siblings,
friends, and well-wishers for your continuous encouragement and support, which have propelled me forward. To my beloved wife, I cannot express enough gratitude for your unwavering patience and support, without which, at times, it would have felt impossible to continue. And to my dearest daughter, your arrival brought joyous news and blessings to our lives. Finally, I express my gratitude to all my teachers, from elementary school through graduate school, whose knowledge and guidance have played an instrumental role in shaping my academic journey.
**PREFACE**

My dissertation comprises two essays that have garnered three conference acceptances. Additionally, I have adapted these essays to meet the submission requirements of top IS journals. Specifically, the first essay is intended for submission to the Journal of Management Information Systems, while the second essay aligns with the submission criteria of Management Information Systems Quarterly.
ABSTRACT

Artificial Intelligence (AI) has generated significant interest due to its potential to augment human intelligence. However, user attitudes towards AI are diverse, with some individuals embracing it enthusiastically while others harbor concerns and actively avoid its use. This two essays' dissertation explores the reasons behind user aversion to AI. In the first essay, I develop a concise research model to explain users' AI aversion based on the theory of effective use and the adaptive structuration theory. I then employ an online experiment to test my hypotheses empirically. The multigroup analysis by Structural Equation Modeling shows that users' perceptions of human dissimilarity, AI bias, and social influence strongly drive AI aversion. Moreover, I find a significant difference between the simple and the complex task groups. This study reveals why users avert using AI by systematically examining the factors related to technology, user, task, and environment, thus making a significant contribution to the emerging field of AI aversion research.

Next, while trust and distrust have been recognized as influential factors shaping users' attitudes towards IT artifacts, their intricate relationship with task characteristics and their impact on AI aversion remains largely unexplored. In my second essay, I conduct an online randomized controlled experiment on Amazon Mechanical Turk to bridge this critical research gap. My comprehensive analytic approach, including structural equation modeling (SEM), ANOVA, and PROCESS conditional analysis, allowed me to shed light on the intricate web of factors influencing users' AI aversion. I discovered that distrust and trust mediate between task complexity and AI aversion. Moreover, this study unveiled intriguing differences in these mediated relationships between subjective and objective task groups. Specifically, my findings demonstrate that, for objective tasks, task complexity can significantly increase aversion by
reducing trust and significantly decrease aversion by reducing distrust. In contrast, for subjective tasks, task complexity only significantly increases aversion by enhancing distrust. By considering various task characteristics and recognizing trust and distrust as vital mediators, my research not only pushes the boundaries of the human-AI literature but also significantly contributes to the field of AI aversion.
# TABLE OF CONTENTS

List of Tables ix
List of Figures x
Introduction 1
Chapter 1: why are people not using AI? an empirical study on AI aversion 7
   Abstract 7
   Introduction 7
   Literature Review 10
   Theoretical Development 12
      Theoretical Framework 12
      Research Model and Hypotheses 13
   Methodology 20
      Data Collection 20
         AI Tools and The Manipulation of Tasks 21
            Measurements 21
      Results 22
         Measurement Model 23
         Manipulation Check 23
         Structural Model 24
         Moderating Effect of Task Complexity 25
   Discussion 26
      Implications for Research 26
      Implications for Practice 28
      Limitations and Future Research 30
   Conclusion 31
   References 31
   Appendix 40
Chapter 2: what tasks lead to AI aversion: the moderated mediation of trust and distrust 43

Abstract 43
Introduction 43

Literature Review 45
AI vs. Algorithm 45
AI Aversion 46
Distrust and Trust 47
Task Complexity 48
Distrust and AI Aversion 48
Trust and AI Aversion 49
Task Complexity and Distrust 50
Moderating Role of Task Subjectivity 51

Methodology 54
Procedure 54
Experimental Design 54
Measurements 55
Analytical Approach 55

Results 55
Measurement Model 55
Manipulation Check 56
Structural Model 56
Moderating Effect of Task Subjectivity 57
Supplemental Analysis 60

Discussion 61
Implications for Research 61
Implications for Practice 63
Limitations and Future Research 63

Conclusion 64
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Statistics and correlations</td>
<td>23</td>
</tr>
<tr>
<td>1.2</td>
<td>Model testing results</td>
<td>25</td>
</tr>
<tr>
<td>1.A1</td>
<td>Demographics of Respondents</td>
<td>40</td>
</tr>
<tr>
<td>1.A2</td>
<td>Goodness-of-fit indices for the structural model</td>
<td>40</td>
</tr>
<tr>
<td>1.A3</td>
<td>Measures</td>
<td>41</td>
</tr>
<tr>
<td>1.A4</td>
<td>Literature Review (Algorithm Aversion)</td>
<td>41</td>
</tr>
<tr>
<td>2.1</td>
<td>AI grouping based on task type</td>
<td>54</td>
</tr>
<tr>
<td>2.2</td>
<td>Descriptive statistics and correlations</td>
<td>56</td>
</tr>
<tr>
<td>2.3</td>
<td>Multigroup analysis results</td>
<td>57</td>
</tr>
<tr>
<td>2.4</td>
<td>Conditional Indirect Effect of Task Complexity on AI Aversion for Two Task Groups</td>
<td>61</td>
</tr>
<tr>
<td>2.A1</td>
<td>Literature review for AI aversion</td>
<td>75</td>
</tr>
<tr>
<td>2.A2</td>
<td>Demographics of Respondents</td>
<td>78</td>
</tr>
<tr>
<td>2.A3</td>
<td>Goodness-of-fit indices for the structural model</td>
<td>78</td>
</tr>
<tr>
<td>2.A4</td>
<td>Measures</td>
<td>79</td>
</tr>
<tr>
<td>2.A5</td>
<td>Literature Review for Algorithm Aversion</td>
<td>79</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1.1 Research Model 13

1.2 Model Testing Results (Base model) 24

1.3 Model testing results for a simple task (numbers on top) and complex task (numbers at the bottom) conditions 25

2.1 Model Testing Result (Base model for main effects) 57

2.2 Model testing results for a subjective task (top) and an objective task (bottom) 58

2.3 Effect of task complexity on distrust for objective vs subjective tasks 58

2.4 Effect of task complexity on trust for objective vs. subjective tasks 59

2.5 Effect of task complexity on AI aversion for objective vs subjective tasks 60

2.6 Effect of Task complexity on AI aversion for objective vs subjective tasks when including trust and distrust as covariates 60
Introduction

While an algorithm is a finite set of rules to solve a specific problem (Knuth, 1997), Artificial Intelligence (AI) encompasses much more than that. The application of AI has gained tremendous attention due to its ability to assist in decision-making processes (Prahl & Van Swol, 2017). Artificial Intelligence (AI) has garnered widespread attention in recent years due to its potential to revolutionize various industries and applications (Topol, 2019; Adir et al., 2020; Johnson et al., 2021). The widespread popularity and trial of AI, like chatGPT, demonstrate the potential of AI applications to improve human lives and society (Biswa, 2023; Sallam, 2023; Surameery & Shakor, 2023). Accenture estimates that within a mere 18 months, insurers could save up to $7 billion by employing AI to automate core administrative functions (Hanover, 2021). Bolton et al. (2018) even predict that the annual growth in Gross Value Added (GVA) resulting from AI will nearly double by 2035, highlighting the vast possibilities and potential of this technology. AI is defined as an information processing system that can interpret external data correctly, learn from it, and use those learnings to achieve specific goals or complete tasks through adaptation to its environment (Haenlein & Kaplan, 2019).

AI's capabilities are numerous and varied, ranging from automating repetitive tasks, reducing human error, improving user experiences, and providing swift service delivery and improving existing tools to analyzing vast amounts of data and informing better decision-making (Haibe-Kains et al., 2020; Bolton et al., 2018; Buchanan, 2019; Goodell et al., 2021). Furthermore, AI holds the potential to cut costs and enhance productivity across industries. For instance, streaming giant Netflix saves nearly a billion dollars annually by leveraging an automated recommendation system (Gomez-Uribe & Hunt, 2015). Various organizations utilize AI for a multitude of purposes, ranging from offering personalized shopping recommendations
(Amazon) to enabling autonomous vehicles (Tesla), healthcare advancements (IBM Watson),
chatbot interactions (EVA), gaming enhancements (AlphaGo), speech recognition (Siri),
customer service support (Watson Assistant), computer vision, recommendation systems, and
even automated stock trading (IBM Education, 2020). It finds extensive applications in the
financial service industry, such as fraud detection, banking robo-advisors, and algorithmic
trading (Buchanan, 2019), and has given rise to the sub-domain of Fintech (Qi & Xiao, 2018;
Cao et al., 2021). The healthcare industry also benefits from AI applications, including AI-led
drug discovery, clinical trials, and patient care (Chan et al., 2019; Woo, 2019; Neill, 2013). AI's
potential applications extend to customer service (e.g., chatbots), education (e.g., tracking
student progress and making recommendations), transportation (e.g., route optimization),
manufacturing (e.g., identifying bottlenecks and predicting maintenance), retail (e.g., analyzing
customer data and making personalized product recommendations), marketing (e.g., targeted
campaigns), and other areas (Chen et al., 2020; Huang & Rust, 2018; Abduljabbar et al., 2019;
Chien et al., 2020).

Despite the significant potential of AI to benefit society and improve human lives, there
exists a prevailing aversion towards it among individuals (Fildes & Goodwin, 2007; Dietvorst et
al., 2020; Prahl & Swol, 2017). AI aversion, as defined in the extant literature (Dietvorst et al.,
2020; Jussupow et al., 2020), refers to the negative affective reactions, cognitions, and behaviors
exhibited by individuals towards AI. This aversion is widespread and can lead to substantial
losses. Particularly in situations where algorithms provide suggestions or recommendations or
when the underlying task involves uncertainty, this aversion tends to manifest (Dietvorst et al.,
2020). For example, surveys and studies have shown that professional forecasters often disregard
or give minimal weight to forecasting algorithms (Fields & Goodwin, 2007). Similarly, the
acceptance of driverless cars or robots as caregivers among most Americans remains low (Smith & Anderson, 2017). Consumers also exhibit a preference for human interaction over chatbots, with 86% expressing a preference for humans (Press, 2019). Notably, IBM's artificial intelligence system, Watson, touted as a multi-billion-dollar tool for bringing about industrial revolutions and profit, faced failure (Kolker et al., 2016; Lohr, 2021). Despite IBM's significant investment in Watson's development, the tool was eventually sold for a fraction of its cost (O'Leary, 2022). Physicians' perception of Watson's ineffectiveness was one key concern contributing to its failure (O'Leary, 2022). Factors such as providing known suggestions, inconsistency, and unsuitable recommendations led to negative evaluations and perceptions among physicians, especially concerning Watson for Oncology, which was trained on synthetic cases rather than actual patient data, potentially incorporating physicians' personal biases into recommendations (Ross & Swetlitz, 2018).

Given that consumers' acceptance plays a crucial role in the success of any IT artifact, it is vital to study AI aversion to foster a reasonable understanding of AI and derive maximum benefits from its technological advancements. Interestingly, the reasons behind AI aversion are multifaceted, as suggested by extant literature (table A1). For instance, while some studies propose that people prefer AI over humans when it outperforms humans (Longoni et al., 2019; Pezzo & Beckstead, 2020), others show that users are less likely to adopt a model with better performance if it is unreliable and does not produce consistent results (Dietvorst et al., 2020). But this is not a ubiquitous finding. For example, Logg et al. (2019) finds that people indeed prefer algorithmic over human judgement. Furthermore, although aversion towards AI can be influenced by the context of the task (Castelo et al., 2019) or the environment (Burton et al., 2020), existing literature on AI aversion has primarily focused on factors related to AI and the
user. Therefore, it is crucial to explore how user aversion towards AI is shaped in different contexts and to develop a comprehensive research model that can holistically explain this phenomenon. (Stein et al., 2019; Xie et al., 2022)

In my two-paper dissertation, I aim to address this concern. In the first paper, I seek to answer the following research question: Why do people exhibit aversion towards artificial intelligence, and how does the context in which AI is applied influence this aversion? To answer this question, we develop a research model based on the theory of effective use (TEU) and adaptive structuration theory (AST), employing a randomized online experiment to test our model's hypotheses. Our analysis reveals that AI aversion is influenced by perceived algorithmic bias (PAB) and perceived social influence (PSI), while perceived human dissimilarity (PHD) mitigates this aversion. Furthermore, these relationships are contingent upon the complexity of the task (TC). This study contributes significantly to understanding AI aversion behavior through several key findings. Firstly, by considering four crucial facets of AI (AI characteristics, task characteristics, environmental characteristics, and user characteristics), we offer a more comprehensive understanding of the factors contributing to AI aversion. Contrary to traditional beliefs, we challenge the notion that users prefer human-like AI and instead demonstrate that human dissimilarity is desirable, particularly for complex tasks. Additionally, by conducting task-dependent inter-group comparisons, we demonstrate that the impact of various antecedents on AI aversion varies depending on the task's complexity. These findings provide valuable insights into the reasons behind user aversion towards AI and shed light on how it varies within different contextual settings. Although our study explores the effects of specific antecedents, the theoretical implications extend beyond these findings, paving the way for future research to
investigate other constructs within the four facets we considered, which could significantly influence AI aversion.

In the second article, I investigate the effect of task characteristics (i.e. task complexity & task subjectivity) on AI aversion and the mediating role of trust and distrust in this relationship. Task complexity is a critical factor that influences users' perception and attitudes toward an IT artifact and can affect their acceptance and use of AI systems. While trust and distrust can play a significant role in mediating this relationship, the subjective nature of a task can further alter it (Castelo et al., 2019). Therefore, investigating the effect of task complexity and subjectivity on AI aversion can provide valuable insights for designing more user-friendly and trustworthy AI systems. Hence, in this study, I aim to address the research gap by examining the effect of task complexity on AI aversion and the mediating role of trust and distrust in this relationship. I propose that users' perception of task complexity positively influences AI aversion, and that this relationship is mediated by their trust and distrust of AI systems. Furthermore, I suggest that these relationships will vary based on the subjectivity of the task. To achieve these objectives, I conduct an online experiment using Amazon Mechanical Turk (MTurk) to collect users' perceptions of task complexity, trust, distrust, and AI aversion. I manipulate the complexity and subjectivity of tasks by randomly assigning participants to one of four task groups and providing them with different AI tools that vary in complexity and subjectivity. Finally, I analyze the data using structural equation modeling in AMOS and examine the moderating effect of task subjectivity using multigroup analysis. I further run additional analysis (i.e., ANOVA) for robustness checks. The findings indicate that the effect of task complexity is significantly mediated by both distrust and trust and that these relationships differ both in strength and direction based on the subjectivity of a task. I expect the findings to contribute to the fields of
Human-AI interaction in general and AI aversion in particular. Because by enhancing our understanding of the factors related to task complexity that contribute to AI aversion, this study provides insights for designing more user-friendly and trustworthy AI systems. Moreover, it empirically validates the importance of considering users' trust and distrust as independent constructs in the study of AI aversion. More importantly, it offers a framework for future research on AI aversion and paves the way for investigating the complex interplay among task complexity, trust, distrust, and AI aversion.
CHAPTER 1: WHY ARE PEOPLE NOT USING AI? AN EMPIRICAL STUDY ON AI AVERSION

Abstract

Artificial Intelligence (AI) has received significant interest for its potential to augment human intelligence. Interestingly, extant literature indicates that users have mixed opinions regarding AI. While some users zealously embrace AI, others express deep concerns and try to avoid it. Intrigued by this phenomenon, in this study, we aim to understand why users engage in AI aversion. We develop a research model to explain users' AI aversion based on the theory of effective use and the adaptive structuration theory. We then employ an online experiment to test our hypotheses empirically. Our multigroup analysis by Structural Equation Modeling shows that users’ perceptions of human dissimilarity, AI bias, and social influence are strong drivers of AI aversion. Moreover, we find a significant difference between the simple and the complex task groups. This study reveals why users avert using AI by systematically examining the factors related to technology, user, task, and environment, thus making a significant contribution to the emerging field of AI aversion research.

Keywords: Artificial Intelligence; Human Intelligence; AI Aversion; The Theory of Effective Use; Adaptive Structuration Theory; Task Complexity

Introduction

While an algorithm is a finite set of rules to solve a specific problem (Knuth, 1997), Artificial Intelligence (AI) is much more than that. Based on the extant literature (Haenlein & Kaplan, 2019), AI is "an information processing system that can interpret external data correctly, learns from it, and uses those learnings to achieve specific goals or complete tasks through adaptation to its environment." Because AI can help people make better decisions (Prahl & Van Swol, 2017), AI application has gained tremendous attention. Organizations utilize it for many purposes, such as offering shopping recommendations (Amazon), autonomous vehicles (Tesla), healthcare (IBM Watson); chatbot (EVA); gaming (AlphaGo), speech recognition (Siri); customer service (Watson assistant); computer vision, recommendation and automated stock trading (IBM Education, 2020). AI is highly desirable because it can automate tasks, reduce human error, improve end users' experience, or provide quick service delivery (Haibe-Kains et al., 2020; Bolton et al., 2018). Moreover, it has the potential to reduce costs and improve
productivity. For example, Netflix, the streaming giant, saves nearly a billion dollars per year by using an automated recommendation system (Gomez-Uribe & Hunt, 2015). A calculation by Accenture estimated that within just 18 months, insurers could save up to $7 billion by using AI in automating core administrative functions (Hanover, 2021). Another study by Bolton et al. (2018) shows that the annual GVA growth because of AI will almost double by 2035. Thus, the possibility and potential of AI are enormous.

Surprisingly, despite AI's potential to benefit society and improve human lives, individuals often show aversion towards it (Fildes & Goodwin, 2007; Dietvorst et al., 2020; Prahl & Swol, 2017). Based on extant literature (Dietvorst et al., 2020; Jussupow et al., 2020), we define AI aversion as "the human assessment of an AI that manifests in a negative affective reaction with concomitant cognitions and behaviors towards it." AI aversion is widespread and could lead to severe losses. When the algorithm provides suggestions or recommendations or the underlying task involves uncertainty, this tendency prevails (Dietvorst et al., 2020). For example, a survey shows that many professional forecasters either did not use the forecasting algorithm or weighed it lightly (Fields and Goodwin, 2007). Similarly, surveys find that most Americans are reluctant to accept driverless cars or robots as caregivers (Smith & Anderson, 2017). Press (2019) shows that 86% of consumers prefer humans over chatbots. Watson, a multi-billion-dollar IBM artificial intelligence supposed to bring industrial revolutions and profit for businesses, failed (Kolker et al., 2016; Lohr, 2021). IBM spent a couple of billions to develop Watson but sold it for a billion later (O’Leary, 2022). Among other reasons (e.g., insufficient data, complexity, and massiness in the data) behind Watson’s failure, one key concern was physicians’ perception of that tool’s being ineffective (O’Leary, 2022). Making known suggestions, inconsistency, and unsuitable recommendations, are some of the factors behind physicians' negative evaluations and perceptions of IBM Watson. Moreover, because Watson for Oncology was trained on synthetic cases compiled by doctors and IBM engineers (rather than actual patient data), it has the potential to accrue physicians’ personal biases toward a particular recommendation (Ross & Swetlitz, 2018). Because consumers' acceptance is essential for the success of any IT artifact, it is crucial to study AI aversion so that individuals, organizations, and society can develop a reasonable understanding of AI and benefit from the advancement of AI technologies.
A thorough review of extant literature indicate several issues that demand for a fresh study on AI aversion. First, despite AI falling under the broader definition of algorithm, AI is much more powerful, and distinct than generic algorithm (more discussion on this later). Hence, findings from algorithm aversion literature cannot be directly applied to understand AI aversion without further investigation. Second, while most studies on algorithm aversion offer online surveys to identify participants’ preferences between a human and an algorithm (e.g., Kießling 2021) or investigate various antecedents of algorithm aversion behavior (Mahmud et al. 2022), to the best of our knowledge none offers a concise research model that can holistically explain the phenomenon of AI aversion. Third, current findings from algorithm aversion literature suggest contradictory results. For example, in response to Longoni et al. (2019), Pezzo and Beckstead (2020) claimed that people prefer AI over humans if it performs better than humans. On the contrary, Dietvorst et al. (2020) showed that people are less likely to use a model with better performance if it is unreliable and does not produce the same result every time. Therefore, it is imperative to independently investigate users’ AI aversion behavior while considering properties and dimensions that are more applicable to the definition and development of AI. Finally, despite studies indicating that algorithm aversion can be conditional on the task context (Castelo et al., 2019), we find only three articles on algorithm aversion (none on AI aversion) that considered task characteristics in their investigations. Moreover, while Castelo et al. (2019) suggests that people do not prefer algorithm for subjective tasks, results from other studies (e.g., Sohn et al. 2020) offer contradictory findings. Hence, it is also important to check how contextual factors like task characteristics can alter our primary relationships. Considering all these, in this paper, we aim on developing a concise research model that can comprehensively explain users’ AI aversion behavior while keeping the following research questions in mind: *Why do people demonstrate an aversion towards artificial intelligence, and how does the context in which AI is applied affect this aversion?*

To answer the aforementioned questions, we first develop the skeleton of our research model based on the theory of effective use (TEU) and adaptive structuration theory (AST). Next, we identify important antecedents of AI aversion grounded on extant literature and their relevancy to the fundamental definition and development of AI. Finally, we employ a randomized experiment to test our model. Our analysis shows that AI aversion is increased by perceived algorithmic bias (PAB) and perceived social influence (PSI) and conditionally
decreased by perceived human dissimilarity (PHD). Moreover, these relationships are contingent on task complexity (TC).

This study makes significant contributions to the understanding of AI aversion behavior. Specifically, we have made the following key contributions. First, by combining four key facets of AI (AI characteristics, task characteristics, environmental characteristics, and user characteristics), we have created a more holistic understanding of the antecedents of AI aversion. Second, we have challenged the traditional belief that users prefer a human-like AI and instead found that human dissimilarity is desirable for complex tasks. Third, through a task-dependent inter-group comparison, we have demonstrated that the effects of various antecedents on AI aversion vary based on the complexity of the task. These findings provide valuable insights into why users demonstrate AI aversion and how it varies depending on the context. While we have studied the effects of several specific antecedents, the theoretical implications of our research extend beyond these findings. Our study suggests that other constructs from the four facets we have considered could also significantly influence AI aversion and serves as a starting point for future research to investigate these unknown effects.

**Literature Review**

An algorithm is a finite set of rules to solve a specific problem and is completed through a series of operations (Knuth, 1997). Similarly, Bhardwaj and Verma (2017) defined it as a sequence of computational steps that transforms inputs into outputs. For our study, we accept this definition and define an algorithm as a set of logical commands that take some inputs (either user provided or automatically generated) and produce some outputs (can be in many forms). The availability of large volumes of data, low data storage costs, and high computing power made it possible for companies to implement data-driven algorithmic approaches to market predictions, product suggestions, and prospective consumer identifications. Predictive algorithms use different computational models to develop suggestions, usually without human influence (Lee, 2018). The algorithm has improved from its basic computational form to a state that it can now learn from past experiences and improve future outputs (Castelo et al., 2019).

As a particular type of algorithm, AI has gained extensive attention recently, especially with significant advantages in deep learning (LeCun, Bengio, and Hinton, 2015; Silver et al., 2016). However, no widely accepted definition (e.g., Kirsh, 1991; Allen, 1998; Bhatnagar et al.,
2018; Monett & Lewis, 2018) is available for AI (Wang, 2019). Wang (1995, p.17) defined intelligence as "The capacity of an information-processing system to adapt to its environment while operating with insufficient knowledge and resources." This definition excludes the conventional computer systems or algorithms that merely depend on the set inputs to produce output based on set models. Similarly, Kaplan and Hae (Kaplan & Hae, 2019, p.15) defined AI as "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation." Based on these two definitions, we define AI as an information processing system that can learn from its interpretation of external data and apply the learned knowledge to achieve specific goals or tasks through adaptation to its environment.

Users' aversion towards algorithms has been studied for a long time. Aversion is "a negative affective reaction with concomitant behaviors and cognitions" (Meier, 1985, p.171). Jussupow et al. (2020) define algorithm aversion as a "biased assessment of an algorithm which manifests in negative behaviors and attitudes towards the algorithm compared to a human agent" (Jussupow et al., 2020, p.4). Drawing from the literature, we define AI aversion as the human assessment of AI that manifests in an adverse affective reaction with concomitant cognitions and behaviors towards AI. Several studies investigated the phenomenon of algorithm aversion. For example, Meehl (1954) presented a theoretical analysis and reviewed the evidence of the phenomenon known as "algorithm aversion," in which individuals prefer predictions made by human experts over those made by statistical algorithms. Dietvorst et al. (2015, p.114) find that although evidence-based algorithms more accurately predict the future than human forecasters, people often choose the former when deciding between a human forecaster or a statistical algorithm. Dietvorst et al. (2018) further show that people often refuse to use algorithms after learning they are not perfect.

Literature on algorithm aversion predominantly focused on users' aversion in the decision domain while comparing a human agent and an algorithmic recommendation under different scenarios (e.g., Meehl, 1954; Prahl & Swol, 2017; Pezzo & Beckstead, 2020; Dietvorst et al., 2015). While doing so, some studies focused on attributes like accuracy, speed, effectiveness, capabilities, and performance specific to the underlying mathematical model (Berger et al., 2020; Dietvorst et al., 2020; Baets & Harvey, 2020). Others focused on attributes like the inability to assess emotional issues (Niszczota & Kaszás, 2020; Castelo et al., 2019), lacking a human-like
mind (Bigman & Gray, 2018), disability to learn (Berger et al., 2020) or inability to provide explanations (Yeomans et al., 2019) that are relatively unique to human. A few studies concentrated on environmental properties, i.e., factors connected to how the algorithm functions under different contexts. For example, lack of incentivization (Burton et al., 2020; German & Merkle, 2020), human-in-the-loop (Burton et al., 2020; Jago, 2019; Jussupow et al., 2020; Logg et al., 2019; Wolf, 2014); ability to modify (Dietvorst et al., 2018) belong to this category. Some studied other factors, e.g., false expectations (Burton et al., 2020) and unfamiliarity (Logg et al., 2019). For a complete list, see Table A4 in the Appendix. Despite the valuable insights gained from research on algorithm aversion, there is still a need for an integrative approach that connects the various dimensions of artificial intelligence (AI) applications.

While some studies have treated AI and algorithms interchangeably, they are different because AI systems can learn from their inputs and adapt their recommendations over time, while algorithms always provide the same output for a given set of inputs. The learning capability of AI makes it difficult to rely on findings from algorithm literature to understand how users will react to AI technologies. Therefore, a fresh, focused examination of how users respond to AI suggestions and recommendations is necessary to understand better and address any potential concerns about the use of AI.

Theoretical Development

Theoretical Framework

We utilize the theory of effective use (Burton-Jones & Grange, 2013) and the adaptive structuration theory (DeSanctis & Poole, 1994) to inform the development of our theoretical framework. Like Liang et al. (2015), rather than strictly following the propositions of these theories, we rely on their organizing logic and rationales to justify the critical conceptual dimensions that might influence AI aversion and choose relevant constructs from these dimensions. We apply TEU and AST to identify construct dimensions that are responsible for users’ AI aversion. Based on TEU, we propose that the use of any AI involves a user, the AI itself, and a task to be completed with the support of the AI (Burton-Jones & Straub, 2006). Although TEU posits that AI use is socially constructed, TEU does not directly consider environmental factors (Liang et al., 2015). To complement this, we adopt AST, which suggests that "the major sources of structure for groups as they interact with an advanced information
technology are the technology itself, the tasks, and the organizational environment" (DeSanctis & Poole, 1994, p.128). AI is an advanced, intelligent IT artifact in the form of software or a combination of software and hardware. Thus, integrating TEU and AST, we contend that users' AI aversion is shaped by factors in four categories: task characteristics, technology (AI) characteristics, user characteristics, and environmental characteristics (Figure 1). Because it is impossible to consider all possible factors associated with a single study, we select one or two representative constructs with strong theoretical and practical relevance in each category. Based on a comprehensive literature review, we selected PAB and PHD as AI characteristics, PSI as the environment characteristic, PLA as user characteristic, and TC as task characteristic. The Measurement Development section under the Methodology section describes the detailed factor selection process.

**Research Model and Hypotheses**

Figure 1 shows our research model. We propose that *perceived human dissimilarity* negatively affects AI aversion, and *perceived AI bias, perceived lack of autonomy, and perceived social influence* positively impact AI aversion. We further posit that these relationships vary based on *task complexity*. To rule out alternative explanations, we control for age, education, gender, time spent using the tool, and prior AI experience (e.g., Dietvorst et al., 2020). We discuss the theoretical rationale behind each hypothesis as follows.

---

**Figure 1.1. Research Model**

![Research Model Diagram](image-url)
Perceived human dissimilarity is the degree to which one perceives an AI to be different from a human in the approaches or strategies it takes to complete a task (Hobman et al., 2004; Jackson et al., 1992; Jackson et al., 1995). In practice, AI is highly data-driven and makes decisions based on its learning from data (Gunning & Aha, 2019). Existing literature on explainable AI suggests that the data-driven approach followed by an AI often acts like a black box that is difficult for users to explain and comprehend (Adadi & Berrada, 2018). AI can discover hidden patterns and biases in data that are entirely unknown to humans. A critical difference between humans and AI is that AI relies on algorithms and mathematical models, while human intelligence comes from experiences, emotions, and senses (Zhou, 2021). AI systems often arrive at conclusions or recommendations through intricate calculations or by identifying underlying patterns that are not easily explainable in human terms. Despite AI superseding humans in different tasks, the dissimilarity in its approaches and the inability to interpret and explain its decisions due to a lack of transparency can create a sense of unfamiliarity to users leading to a lack of trust in the system (Zerilli et al., 2022).

Another difference between AI and human intelligence is that AI is often designed to perform specific tasks and trained based on specific criteria and constraints. While AI sticks to data, humans often holistically consider the rationality or the principle of logic with emotion and experience and make optimal decisions (Shafir & LeBoeuf, 2002; Korteling et al., 2021). Such dissimilarities make humans preferred over AI for tasks that require a subjective approach to making decisions (e.g., Castelo et al. 2019). Hence, the more dissimilar an AI is in its strategies to solve a problem, the more unfamiliar and obscure it gets for users. Realizing this, users will expect AI to think or behave similarly to humans, and a high dissimilarity can hint at the inferiority of AI and increase users' AI aversion. Hence,

**H1: Perceived human dissimilarity is positively associated with AI aversion**

Perceived AI bias is the degree to which a user perceives an AI to "compound existing inequalities in socioeconomic status, race, ethnic background, religion, gender, disability or sexual orientation to amplify them and adversely impact inequalities in" the applied field (Panch et al., 2019: p1). Recent studies on machine learning and AI have paid much attention to the vulnerabilities of algorithms to bias because of their severe social ramifications (e.g., Challen et al., 2019; Ntoutsi et al., 2020; Parikh et al., 2019). For example, the COMPAS system that predicts a person's reoffending risk is found to be biased against black defendants. Compared to
whites, it predicted higher risks for blacks than their actual risks, exposing an inherent racial bias in the model (Angwin et al., 2016). Similarly, Google’s ads tool, initially designed as gender-neutral, was found to be gender biased because it populated high-paid jobs less for women than men (Lambrecht & Tucker, 2019). Because the output quality of an AI is highly dependent on the quality of input data and the AI's output represents the results of calculations set by humans who developed the algorithms, an AI may inherit biases from the human developers (Ntoutsi et al., 2020). As a result, the AI might exhibit or amplify the existing discriminations and inequalities in its decisions (Karimi et al., 2018).

Extant literature indicates that people have lower confidence if they perceive someone as biased (Yeung, 2019). This is further supported by studies on tourism, recommendation systems, and artificial intelligence, which indicate that users perceive a biased service or a system as unreliable (Shang et al., 2022; Chau et al., 2013). This unreliability leads to anxiety and alters people’s behaviors negatively (Dwork & Minow, 2022). An unreliable, discriminant AI can mistreat certain groups of people and go against the principles of equity and justice. Consequently, an AI will be abandoned if people perceive its output as biased (Hong et al., 2020). Therefore,

**H2: Perceived AI bias is positively associated with AI aversion**

Perceived lack of autonomy is the degree to which a user perceives no control over using information technology (Liang et al., 2015). Based on self-determination theory, autonomy is one of the three basic human needs underlying motivated behaviors (Deci & Ryan, 1987; Deci & Ryan, 2000). Prior studies identified user autonomy as an essential antecedent of technology acceptance and exploratory use (Boonstra et al., 2010; Lowry et al., 2012; Walter & Lopez, 2008). Extant studies indicate that autonomy and independence are fundamental concerns regarding deploying robotic and autonomous systems in health care (Tan et al., 2021). Users may feel they need more autonomy because they cannot adjust the decision-making process of an AI (Kuncel, 2018) or modify an AI's design (Neumann et al., 2021). If users believe that they have little control over an AI when using it (e.g., setting hyperparameters, selecting models or features), their motivation to use that AI can be dampened (Burton et al., 2020). In contrast, research shows that users demonstrate less aversion if they have the autonomy to modify an algorithm (Dietvorst et al., 2020). People perceive autonomy as an essential factor because it delivers a sense of being in control. People generally prefer to be in the driver's seat to ensure
that the decisions reflect their interests and ego. Therefore, users may feel threatened if they do not control an AI’s decision-making procedure (Hebler et al., 2022). Moreover, a lack of autonomy yields lower satisfaction (Jung, 2011), and users are likely to use an AI if they are satisfied with the user experience. Therefore,

**H3: Perceived lack of autonomy is positively associated with AI aversion.**

Social influence is the degree to which users perceive that people important to them believe they should perform a specific IT behavior (Venkatesh et al., 2003). In our research context, the specific IT behavior refers to AI aversion. Extant studies find social influence responsible for changing individuals' thoughts, feelings, attitudes, or behaviors (Walker, 2007). If a large portion of a user's social group holds a particular opinion, the user will likely accept that view and change behaviors accordingly (Walker, 2007). In terms of the application of an AI, three parties can be involved: the first party, who interacts with an AI's decision that affects other people; the second party, who are directly affected by an AI's decision; and a third party, who are not affected at this moment but can become the second party sometimes in the future (Langer & Landers, 2021). Research indicates that the second and third parties can reform the first party's perception regarding AI (Healy et al., 2020). They can provide feedback regarding different aspects of an AI (e.g., accuracy, efficiency, consistency) and negatively influence the first party's opinion (Lee et al., 2015).

Furthermore, third parties can engage in social media and word-of-mouth discussions to shape others' opinions regarding AI (Langer & Landers, 2021). Social influence might accrue explicitly in a person's social setting, whereas others suggest AI aversion. It can further happen implicitly when users fear that others will perceive them as less capable or intelligent because they take help from an AI (Arkes et al., 2007; Diab et al., 2011; Eastwood et al., 2012). For example, Liang and Xue (2022) demonstrate that physicians' reluctance to accept a clinical decision support system (CDSS) is increased by their fear of losing face. Thus, when users receive direct feedback or suggestions from others or feel afraid that others will find them incompetent and lacking necessary expertise, they could exhibit an aversion behavior. Therefore, perceived social influence has the potential to convince users to engage in AI aversion. Hence, we propose,

**H4: Perceived social influence is positively associated with AI aversion**
Task complexity is defined as the amount of attentional capacity and cognitive processing that a task requires from a user (Bonner, 1994). It can increase with the number of instructions users need to follow or the dimensions they need to consider while performing a task. Because complex tasks require a higher level of skills and expertise, users' responses will differ for tasks with varying levels of complexity (Liu & Li, 2012). Thus, task complexity will likely moderate the effects of all independent variables in our model on AI aversion.

First, a complex task is more cognitively challenging for users than an easy task because it requires more information to be processed and introduces a higher cognitive load (Plass et al., 2012). Users are only boundedly rational because they have a limited capacity for working memory (Kalyuga, 2012). AI can remember old data without errors and process many real-time inputs more quickly and accurately to deliver a solution. This is especially effective and useful for solving complex tasks which require high information precision and intensive data analysis. For example, in solving highly complex clinical tasks, clinicians can employ AI to reduce cognitive load to minimize uncertainties and optimize performance (Adler-Milestein et al., 2021). Moreover, we contend that when people face a complex task, they are inclined towards someone or something they perceive as different and superior. The literature on superheroes supports our contention (Rosenberg & Canzoneri, 2008). Superheroes are portrayed as larger-than-life (i.e., different from humans) characters with extraordinary abilities, providing hope and inspiration for people, particularly in difficult situations (Morrison, 2011; Russel, 2013). The notion is that when people realize that they cannot overcome difficulties by solely relying on their abilities, they are more willing to resort to help from superheroes that are different from humans. Hence, in more complex situations, human dissimilarity of AI can be seen as a signal of possessing a superpower and becoming desirable. Therefore,

**H5a:** Perceived human dissimilarity has a stronger negative effect on AI aversion for complex tasks than for simple tasks.

Second, the possibility of AI bias is more substantial for complex tasks because solving complex tasks requires considering more significant amounts of variables and data and, thus, more chances to introduce biases. Sophisticated models involved in solving complex problems usually include a higher number of parameters. For example, OpenAI trained the GPT-3 language model on 175 billion parameters, while the large language model PaLM was trained on 540 billion parameters (Chowdhery et al., 2022). Google developed its latest language model
based on the switch transformer algorithm and trained it on 1.6 trillion parameters (Fedus et al., 2021; Wiggers, 2021). As more data are processed, and more parameters are configured during model training, there are more opportunities for biases to creep into the system through training data or the algorithm that determines the weights of different parameters. A study (Gichoya et al., 2022) finds that AI can identify race from medical images where humans cannot, highlighting the potential for AI to perpetuate existing biases and racism even when AI developers are unaware of such potential. Because an AI is trained on data that reflects the biases present in society, when the AI is used to make crucial decisions, such as in the medical diagnosis, these biases can seriously harm marginalized groups. For example, if an AI is trained on data that reflects racial discrimination in the criminal justice system, it may lead to biased sentencing and parole decisions against African Americans. For example, in 2018, researchers from MIT published a study that revealed that Amazon's facial recognition technology, known as "Rekognition," exhibited significant biases in identifying individuals with darker skin tones, particularly women (O'Brien et al., 2019). Similarly, in 2016, ProPublica investigated and determined that the risk assessment algorithm COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) was significantly more likely to incorrectly flag black defendants as being at high risk of reoffending compared to white defendants (Angwin et al., 2016). These incidents highlight the potential for bias and discrimination in using AI systems in essential decision-making processes.

Additionally, if a complex AI is found to be biased, it will be challenging to conduct a root cause analysis to determine the source of biases. Extensive evidence suggests that AI can carry human and societal biases (Silberg & Menyika, 2019). For example, employee recruitment is a complex business practice, and many factors are considered when hiring. Amazon developed an AI for this purpose, but it turned out to be gender biased. Unable to identify the root cause, Amazon had to terminate using it altogether (Dastin, 2018). From a user's perspective, compared to simple tasks, complex tasks typically have more significant ramifications if anything goes wrong. For example, a movie recommendation agent that recommends a movie (a simple task) based on biases would lead to less severe damage than a medical diagnostic AI system that makes inaccurate diagnoses (a complex task) based on biases. The inability to identify the source of biases will exacerbate the situation even more for complex tasks (Challen et al., 2019). As a
result, users will exhibit a stronger aversion if they perceive an AI to be biased in the case of a complex task than a simple task. Hence,

**H5b:** Perceived AI bias has a stronger positive effect on AI aversion for complex tasks than for simple tasks.

Third, the outcome of complex tasks is more uncertain than simple tasks (Avgerinos & Gokpinar, 2017). The higher number of factors that need to be considered for a complex task contributes to uncertainty (Schroder et al., 1967). Uncertainty can be defined as the lack of predictability associated with inputs, processes, or outputs of the task being completed (Griffin et al., 2007). Users would like higher autonomy when working on a complex task to reduce such uncertainties. Hence, the higher the uncertainty due to task complexity, the more users would like to be autonomous and have control over the task (Osman, 2010). Increased autonomy will reduce barriers and constraints in applying one's understanding, knowledge, and skills and enable a user to identify and respond to variances or uncertainties more effectively (Cordery et al., 2010). For example, Dietvorst et al. (2018) show that in an uncertain decision domain, users show less aversion to an AI if they can modify it, i.e., they have better control over it. Similarly, user autonomy studies suggest that people with higher autonomy feel more productive and satisfied (Langford, 2013; Shobe, 2018). Because a complex task demands higher efforts, skills, and knowledge from users, maintaining personal motivation or feeling productive is usually tricky. Hence, a complex task scenario where users further suffer from a lack of autonomy because of an AI will exacerbate the issue. For example, surgery is a highly complex task, and Gumbs et al. (2021) find that surgeons are skeptical of autonomous agents in such applications. Similarly, studies show that people are reluctant to trust autonomous cars because these cars reduce drivers’ autonomy (Kaur & Rampersad, 2018). Contrarily, users will more likely delegate a simple task to an AI agent because of less uncertainty in the outcome. Even if they have no control over how the task is performed, they know that the outcome will be the same. Hence,

**H5c:** Perceived lack of autonomy has a stronger positive effect on AI aversion for complex tasks than for simple tasks

Finally, as H4 posits, social influence occurs when a user’s social referents approve of AI aversion (Walker, 2007; Liang & Xue, 2022). Cognitive science research shows that people’s opinions and behaviors get reshaped during social interactions because of uncertainties in their judgments (Verma et al., 2016). There are two forms of social influence: 1) the expert effect:
when the opinion comes from authoritative individuals in the group, and 2) the majority effect: when most people in the group share a similar opinion (Moussaid et al., 2013). A robust and confident opinion or a consensus must be formed for social influence to affect one's opinion successfully. Usually, AI relies on complex mathematical models that make it difficult for users to comprehend the rationale behind an AI's decisions (Garcia-Magarino et al., 2019).

Understanding AI and identifying its strengths and weaknesses will be more difficult for a complex-task AI than a simple-task AI because the former relies on more complex models and is usually trained on more features (Fedus et al., 2021; Wiggers, 2021). When solving a complex task, AI performance can be evaluated from many different perspectives, and spectators can develop different and conflicting opinions from specific perspectives. As a result, achieving consensus regarding AI performance in complex tasks is not easy. The user, confused by the diverse and inconsistent voices, is unlikely to figure out what to do. Studies on opinion formation suggest that users commonly underestimate, if not ignore, contradictory feedback from their social group (Moussaid et al., 2013). This is in line with other findings that contradictory beliefs have an almost non-existent influence on individuals (Lorenz, 2007; Hegselmann & Krause, 2002). In contrast, it is easy for people to agree on the AI application for simple tasks. Thus, for a simple task, the social influence of the user's social environment is potent and self-reinforcing, which can substantially affect the user's behavior.

**H5d: Perceived social influence has a stronger effect on AI aversion for simple tasks than for complex tasks**

**Methodology**

**Data Collection**

We developed a randomized online survey on Amazon's Mechanical Turk (MTurk) to test the research model. MTurk is an online platform that allows data collection from mass users and has long been used for experimental research (Paolacci et al., 2010). We administered the survey via Qualtrics, an online survey platform. Because the use of AI is not limited to a specific group of users and we wanted to study the perception and behavior of all types of AI users, we did not specify any inclusion criteria. The respondents were informed that the survey was anonymous, and that no personal information would be collected. Each participant was randomly assigned to one of two conditions: a simple task condition or a complex task condition. In both conditions, participants received one of the two AI tools to complete a task and completed the
rest of the survey, including quality control questions afterward. A total of 398 responses were collected. We rejected 83 responses because of failing the quality control questions, resulting in 315 valid responses. In our final dataset, we had 174 respondents in the simple-task group and 141 in the complex-task group. The respondents have an average age of 37 years (SD=11.7), ranging from 21 to 89 years. There were 39% female and 61% male participants in the sample. Most participants have at least an undergraduate degree or above (57.8 percent), and all participants have some computer experience. The demographics of the respondents are shown in Table A1 in Appendix.

**AI Tools and The Manipulation of Tasks**

We selected two AI tools for two different tasks in this study. The first AI tool is a movie recommender. It takes users' input by asking six questions and suggests movies based on their answers. The questions are i) how are you feeling today?; ii) What comes closest to your occasion (e.g., movie date, just watching for myself, etc.); iii) Please choose any genre you are interested in; iv) How old would you like the movie to be?; v) Is the age-appropriateness rating of the movie important to you?; and vi) Please select any other category you are interested in. Based on the user responses, the AI tool suggests movies for the user.

The second tool is a health advisor. It diagnoses participants' health conditions and suggests the proper care for them. The health AI tool asks users questions and uses the responses as inputs to diagnose health conditions and make care suggestions. Users can skip a question and continue to the next. Once done, the tool offers a prospective care suggestion.

We define movie recommendation as a simple task and the health diagnosis as complex. We ask the respondents to rate task complexity in each scenario and use their reported ratings to confirm that user perceptions are consistent with each task’s complexity level.

**Measurements**

To identify articles related to AI aversion, we searched the Web of Science database using the following keywords: "Artificial intelligence aversion," "Algorithm aversion," "AI aversion," "Algorithm appreciation," or "AI appreciation" or "Artificial intelligence appreciation" for English articles from 1965 to 2020. This search results in 20 journal articles. Three articles are related to advantageous inequality (AI) aversion and are excluded from our analysis. Next, we searched "Scopus" (Elsevier's abstract and citation database) with the six keywords. We kept articles, reviews, or conference papers and excluded books or book chapters, resulting in another
20 articles. After combining articles from these two sources and removing 13 common articles, we obtained 22 journal articles that we used to determine antecedents of AI aversion.

We first identified 58 factors from the selected articles. Thirty-two are related to algorithm aversion, and 26 are to algorithm appreciation. Many factors have common theoretical meanings, with their names being synonyms or antonyms. We merged similar constructs based on their definitions, which resulted in a final list of 31 constructs (see Table A3, Appendix). Based on their definition, we collapsed these constructs under the four dimensions identified in our theoretical framework: AI, user, environment, and task.

We adopted the measurement instruments from validated measures in the extant literature. All the items are reflective, and the instruments can be found in Appendix, Table A3. Specifically, the three items of perceived human dissimilarity are adapted from Longoni et al. (2019). The three items of perceived AI bias are adapted from Pacheter et al. (2010). Perceived lack of autonomy is measured using three items adapted from Liang et al. (2015) and Ahuja and Thatcher (2005). We measure perceived social influence using a three-item scale from Chi et al. (2020). Similarly, AI aversion is measured following Bhattacherjee and Hikmet (2007) using a three-item scale. Finally, we measure the task complexity score by asking participants to indicate their perception of that task’s complexity using a seven-point Likert scale following Wang et al. (2014).

To rule out alternative explanations, we control for age, education, and gender (Dietvorst et al., 2015; Dietvorst et al., 2020). Because the time required for completing the task can differ, which could influence user perceptions, we also control for the total time spent completing the task. We pretested the measures on 100 undergraduate students, and the results confirmed that the measures have sufficient reliability and validity. The student responses were not included in the final data set for analysis.

**Results**

The covariance-based structural equation modeling (CBSEM) is employed to test the research model based on the following considerations. First, this method is widely used to investigate complex relationships between latent constructs in the social science research domain (Astrachan et al., 2014). Second, it can effectively evaluate complex measurement models and structural paths involving a multitude of variables and levels of constructs (Wang & Wang, 2019; Astrachan et al., 2014). Our model includes latent constructs and a complex structure,
including mediation and moderation of relationships. Thus, CB – SEM will be appropriate for analyzing data in this study. We use IBM AMOS 27 for this purpose.

Measurement Model

We evaluate the measurement model by examining each construct's reliability, convergent validity, and discriminant validity (Liang et al., 2015). Table 1 below shows that all values of composite reliability and Cronbach's alpha are more significant than 0.7, indicating acceptable construct reliability (Hair, 2009). All the items have a loading above 0.7, and only four are below 0.7 but above the acceptable threshold of 0.6 (Chin, 1998), suggesting satisfactory convergent validity. Regarding discriminant validity, we check if the factor loadings on their construct are more significant than the loadings on other constructs (Fornel & Larcker, 1981). The result shows satisfactory discriminant validity. Moreover, each construct’s square root of average variance extracted (AVE) is greater than its correlations with all other constructs in the model, confirming sufficient discriminant validity (Liang et al., 2015; Fornell & Larcker, 1981). We further check for common method bias (CMV) by conducting Harman’s one-factor test. The test generated five principal constructs, and the unrotated solution shows that the first construct explains only 44 percent of the variance. This value is below the accepted threshold of 0.5 and indicates that our data does not suffer from high CMV. (Liang et al., 2015).

Table 1.1: Statistics and correlations

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>CR</th>
<th>Alpha</th>
<th>SD</th>
<th>AVE</th>
<th>HD</th>
<th>AB</th>
<th>LA</th>
<th>SI</th>
<th>AA</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD</td>
<td>4.69</td>
<td>0.89</td>
<td>0.83</td>
<td>1.41</td>
<td>0.72</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td>4.15</td>
<td>0.90</td>
<td>0.90</td>
<td>1.70</td>
<td>0.76</td>
<td>0.41 **</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LA</td>
<td>5.06</td>
<td>0.90</td>
<td>0.86</td>
<td>1.46</td>
<td>0.76</td>
<td>0.24 **</td>
<td>0.38 **</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>4.00</td>
<td>0.89</td>
<td>0.94</td>
<td>1.81</td>
<td>0.74</td>
<td>0.44 **</td>
<td>0.65 **</td>
<td>0.35 **</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>3.88</td>
<td>0.90</td>
<td>0.94</td>
<td>1.89</td>
<td>0.75</td>
<td>0.29 **</td>
<td>0.71 **</td>
<td>0.40 **</td>
<td>0.75 **</td>
<td>0.87</td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05; p** < 0.01
Notes: CR: Composite Reliability; Alpha: Cronbach’s Alpha; S.D.: Standard Deviation; HD: Perceived Human Dissimilarity; AB: Perceived Algorithmic Bias; LA: Perceived Lack of Autonomy; SI: Perceived Social Influence. The bold values on the diagonal are the square roots of AVE.

Manipulation Check

We conduct a t-test to check if our manipulation of task complexity works. The result shows that the health condition diagnosis group has a significantly higher complexity score than
the movie recommendation group (p < 0.01; F = 42.033, df = 313), indicating a successful manipulation.

**Structural Model**

As indicated in Figure 2, our research model explains 74 percent of the variance for AI aversion. We assessed the hypotheses by checking the direction and significance of path coefficients in the SEM model. To test H1-H4, we used a bootstrapping procedure of 2000 resamples to analyze the entire sample of 315 responses. The SEM model’s goodness of fit (Chi-squared/df = 1.90, CFI = 0.98, NFI = 0.95, and RMSEA = 0.05) suggests an excellent fit with the data (Suki, 2014). Table 2 shows our model testing results. Specifically, both perceived AI bias (β = 0.42, p < 0.01) and perceived social influence (β = 0.54, p < 0.01) positively affect AI aversion, supporting H2 and H4. Similarly, perceived human dissimilarity (β = -0.15, p < 0.01) negatively affects AI aversion, supporting H1. Because the perceived lack of autonomy (β = 0.08, p > 0.05) is not found to be significant, our findings do not support H3.

![Figure 1.2. Model Testing Results (Base model)](image)

**Notes:** *p < 0.05; **p < 0.01; nonsignificant paths are in dashed lines

For moderating effects, we find that task complexity increases the effect of human dissimilarity (β = -0.27, p < 0.01), and the effect of perceived AI bias (β = 0.51, p < 0.01) on AI aversion, supporting H5a and H5b. Similarly, the effect of perceived social influence (β = 0.46, p
< 0.01) on AI aversion decreases for a complex task, supporting H5c. Again, we do not find the effect of perceived lack of autonomy ($\beta = 0.08, p > 0.05$) to be significant. Hence, H5c is not supported.

Table 1.2. Model testing results

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Hypotheses</th>
<th>Full Sample</th>
<th>Hypotheses</th>
<th>Simple Task</th>
<th>Complex Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H1</td>
<td>-0.15 (0.06)**</td>
<td>H5a</td>
<td>-0.04 (0.07)</td>
<td>-0.27 (0.09)**</td>
</tr>
<tr>
<td>AI Bias</td>
<td>H2</td>
<td>0.42 (0.07)**</td>
<td>H5b</td>
<td>0.39 (0.08)**</td>
<td>0.51 (0.12)**</td>
</tr>
<tr>
<td>Lack of Autonomy</td>
<td>H3</td>
<td>0.08 (0.06)</td>
<td>H5c</td>
<td>0.08 (0.07)</td>
<td>0.09 (0.08)</td>
</tr>
<tr>
<td>Social Influence</td>
<td>H4</td>
<td>0.54 (0.05)**</td>
<td>H5d</td>
<td>0.53 (0.06)**</td>
<td>0.46 (0.09)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td>0.74</td>
<td></td>
<td>0.78</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Note: *$p < 0.05$; **$p < 0.01$. Standard errors are in parentheses. Significant & supported hypotheses in bold.

**Moderating Effect of Task Complexity**

Following Floh and Treiblmayer (2006) and Bakker et al. (2003), we conducted a multi-group analysis to test H5a – H5d. First, we compared the focal, freely estimated (i.e., unconstrained) model with a competitive model in which structural parameters were constrained to equality between the simple and the complex groups. We then check the subsequent Chi-square test result and find that the two groups are significantly different (df = 17, CMIN = 30.28, $p < 0.05$). The result indicates a significant difference in the relationships between the two task groups.

Figure 1.3. Model testing results for a simple task (numbers on top) and complex task (numbers at the bottom) conditions
As Figure 3 shows, the effects of perceived human dissimilarity and perceived algorithmic bias are more substantial for complex tasks, perceived social influence is more substantial when the task is simple and perceived lack of autonomy is insignificant for both task groups. In summary, we find that H1, H2, H4, H5a, H5b, and H5d are supported, but H3 and H5c are not.

**Discussion**

The findings on the effect of perceived AI bias, perceived human dissimilarity, and perceived social influence on AI aversion suggest that these factors play a significant role in determining how users will react and interact with AI systems. Our findings have important implications for research on AI use in different contexts.

**Implications for Research**

First, although prior studies identified important antecedents of algorithm aversion (Dietvorst et al., 2015; Dietvorst et al., 2018; Dietvorst et al., 2020), they typically examined each antecedent individually in isolation from other antecedents. This individual experimentation makes it difficult to conceive how these antecedents’ influences would change when they are considered simultaneously. Our study integrates four possible dimensions of AI application and highlights the importance of considering social and psychological factors. Social influence turns out to be the most potent influencing factor for AI aversion. Social influence has long been recognized as a significant determinant of technology acceptance (e.g., Venkatesh & Morris, 2000; Hsu & Lin, 2008). While there is nothing new to finding social influence to be a significant antecedent of an IT-related behavior (in our case, AI aversion), it is worth noting that social influence being the most vital antecedent in our research model suggests the unique nature of the artifact of AI. Unlike traditional information technologies, AI is intended to augment human intelligence, and its working mechanisms are usually opaque, creating many doubts and uncertainties in users’ minds. When people are unable to acquire sufficient information to decide whether they should embrace or avert AI in order to reduce risks, they refer to their social circles and the important ones who influence their behaviors. As a result, social influence becomes a powerful determinant of AI aversion. This finding suggests that users are inclined to make sense
of AI jointly because of the complexities of this advanced technology. This highlights the importance of further scrutinizing social influence when theorizing AI aversion.

Second, the moderating role of task complexity suggests that IS research can get valuable insights by focusing on task characteristics. It is crucial to understand task complexity for its own sake and potential implications for research as a boundary condition (Campbell, 1988). Few studies on algorithm aversion focused on task characteristics, except Castelo et al. (2019), who investigated the direct effect of task objectivity on algorithm aversion. By investigating task complexity as a moderator, we demonstrate that users’ AI aversion behavior is contextual. Because task properties (e.g., complexity) are not a direct property of an AI, it helps to offer novel insights when used as a moderator. We find significant differences between our study’s simple and complex task groups. These findings suggest that the complexity of the task being performed by an AI system plays a role in determining people’s responses to AI. It is possible that when people are faced with tasks with different levels of complexity, they have specific expectations regarding what types of assistance they need and how they should be assisted. For example, we find that when facing complex tasks, users prefer to be helped by AI that does not simulate how humans solve problems, suggesting their willingness to defer to superpower. However, this deference disappears in the context of simple tasks. Our finding suggests that besides task complexity, other task characteristics might also play interesting moderating roles in users’ decision process for AI aversion. Therefore, future studies should consider various task characteristics to investigate how user behaviors differ in different task contexts. This will help develop a deepened and holistic understanding of AI aversion.

Third, our findings underscore the importance of AI bias in designing and developing AI systems. Despite researchers’ efforts to avoid embedding biases into AI systems (e.g., Sun et al., 2019; Noseworthy et al., 2020), AI biases seem to be inevitable, and there are many high-profile cases of biased AI systems that led to unfair or controversial decisions (e.g., Bellamy et al., 2019; Challen et al., 2019). Our study shows that people exhibit aversion if they perceive an AI system as biased. People dislike biased AI because it can lead to a lack of trust in the system and its outcomes. If users perceive the AI system as biased, they will hardly trust its outputs or recommendations and will be less likely to use it or act accordingly. This can be especially problematic if the AI system is used for critical decision-making (e.g., financial investments, jurisdictions, health diagnosis), as our findings suggest that the effect of AI aversion is higher.
when the task is complex. Thus, future studies should consider if trust as an underlying psychological construct plays any significant role in developing AI aversion.

Finally, it is surprising that lack of autonomy is not a factor influencing AI aversion. However, this does not mean that the concept of autonomy concerning AI is not essential, rather it highlights the complexity of human attitudes towards AI. One possible explanation for this nonsignificant result is the nature of the tasks used in the study. Both the movie suggestion task and the health diagnosis task in this study were low-stake tasks, meaning that they did not have an immediate threat to the users. For example, in the movie suggestion task, the AI's suggestions did not directly affect the user physically or financially. In the health diagnosis task, the users recalled a recent health issue and analyzed the response from the AI, and they did not have to follow the advice from the AI. Thus, users possibly did not feel threatened or anxious even though they lacked control over the AI. Therefore, the perceived lack of autonomy did not significantly affect AI aversion. It further suggests a new direction for future studies where researchers can simulate a high-stake scenario (e.g., riding in an autonomous vehicle or making financial decisions with the help of AI), to understand better the effect of perceived lack of autonomy on AI aversion. These high-stakes scenarios can provide a more realistic representation of how people perceive and react to AI systems that directly impact their safety, security, or financial well-being. The results of these studies could provide valuable insights into how people make decisions in high-stake situations and how the perceived lack of autonomy in AI systems affects their attitudes and behaviors toward AI.

**Implications for Practice**

Our study offers several practical implications. First, it is essential to note that while AI’s ability to learn and adapt may lead to the perception of its higher trustworthiness and lower error rates than traditional algorithms, this does not guarantee that it will not face user aversion. Our research findings inform industries implementing AI systems what factors contribute to the potential AI aversion. To ensure the successful adoption of AI technology, developers and businesses must proactively address and understand issues through user-centered research, testing, and feedback. Moreover, it is crucial to understand that user aversion towards AI is not limited to the AI system itself but can also be contingent on factors related to the task and environment in which it is being implemented. Thus, to effectively address and prevent user
aversion, businesses must adopt a holistic approach, considering all relevant factors and taking a comprehensive view of the phenomenon.

Second, contrary to the common belief that people are interested in AIs like humans, our study indicates this is only sometimes true. Industry practices suggest a high emphasis on replacing human agents in numerous fields, including but not limited to customer service (Xu et al., 2017), cognitive therapy (Go and Sundar, 2019), and managerial tasks (Yam et al., 2022). Additionally, developers try inventing human-like chatbots or agents, hoping to increase adoption and use (Go and Sundar, 2019; Yam, 2022). Our study finds that the perceived human similarity of an AI is only sometimes desirable to users, especially for complex tasks. Instead, users may be more interested in using a dissimilar AI in terms of its strategies. The analytic approach taken by AI might differ from the so-called "rational" approach familiar to humans. This procedure helps ensure its decisions are consistent and aligned with personal or organizational goals. This is particularly helpful in situations with many potential courses of action, and choosing the one that will lead to the desired outcome is essential. Moreover, despite humans being considered relatively high in intelligence than animals, it has a limited physical computing capacity (Korteling et al., 2021). Because AIs have different operating systems and possess cognitive qualities and abilities different than humans (Korteling et al., 2018; Shneiderman, 2020), they can beat humans in terms of processing information and computation (Korteling et al., 2021). In addition to having a limited processing capacity, humans' cognitive information processing is more susceptible to systematic distortion. On the other hand, AI is not limited by physical abilities or cognitive limitations and can process a large amount of data quickly and accurately. Hence, AI can be more precise and objective than human intelligence and better transform data-rich problems into solutions (Guo, 2020). Therefore, industries developing AI systems for complex tasks should focus on AI properties that increase acceptance, such as accuracy and consistency (Glikson et al., 2020), rather than overly emphasizing the development of a human-like AI.

Third, we found that social influence substantially impacts AI aversion for simple and complex tasks. This has important implications for businesses. Because developing AI tools and training employees are expensive for organizations, user aversion can result in high costs. For example, Microsoft's AI tool MT-NLG, which was trained on 530 billion parameters across 560 Nvidia DGX A100 servers, costed millions (Wiggers, 2021). Despite the cost of training AI tools
decreasing over the years, it often exceeds the budget of most companies. For example, the training of Alibaba's MT6-10T language model is estimated to cost around $300,000 per 10 days. Similarly, Google's DeepMind spent approximately $35 million on training a model to learn the Chinese board game Go (Wiggers, 2021). On the other hand, the overall organizational expenses for employee training in the U.S. were almost $100 billion in 2022 (Lakewood Media, 2020). To reduce AI aversion, creating a positive community among users is crucial, and key influencers can be helpful in this regard. Thus, businesses can address potential user aversion by identifying key influencers within the user group and involving them in the AI system development and implementation process. By utilizing the social influence effect inversely, organizations can leverage it to their advantage. For example, they can share their feedback and encourage others to share their experiences, and ideas, which can help build trust and reduce the potential for AI aversion among users. Overall, the finding that social influence increases AI aversion highlights the importance of considering the social context in which AI systems will be used and taking steps to build trust and address potential user concerns.

Our findings suggest that when it comes to users' perception of AI, the issue of control may not be as significant as previously thought. Instead of prioritizing the provision of extensive control options for users, AI developers should focus on other aspects of user experience that can improve the overall acceptance and usage of the technology. One such aspect can be making the application user-friendly and easy to use. By simplifying the interface and making the technology accessible to many users, developers may increase the likelihood of widespread adoption. However, it is essential to note that providing too many options or overly complex controls can have the opposite effect. Studies have shown that too many options can overwhelm users and reduce confidence in their choices, leading to poor decision-making (Iyengar & Lepper, 2000). Additionally, overly complex controls can make it challenging to use, leading to user frustration and decreased use (Chernev et al., 2015). Therefore, it is essential to balance providing enough options to give users a sense of control and ensuring that the controls are simple enough.

**Limitations and Future Research**

Despite the insights gained from our study, several limitations should be considered. First, our sample consists of respondents from the US, which means that cultural factors may limit the generalizability of our findings. Future research should consider the influence of
established cultural factors, such as uncertainty avoidance, on AI aversion to address this limitation. Second, our data set consists of cross-sectional data, meaning we cannot directly support a causal relationship between our proposed antecedents and AI aversion. Collecting longitudinal data in an experimental setting can be a solution to explore how these factors affect AI aversion over time. Third, while we have included representative factors from four possible domains of AI application (i.e., AI, user, environment, and task), other factors undoubtedly influence AI aversion. For manageability consideration, we could only consider some possible factors in our study. Future research should identify and consider other factors that may affect AI aversion. Fourth, we did not find a significant effect of users' perceived lack of autonomy on AI aversion. As mentioned earlier, this nonsignificant issue may be because our tasks were relatively low stakes and users were not particularly concerned about their autonomy. To explore this relationship further, future research could investigate the influence of perceived lack of autonomy in an experiment where the results of an AI pose a higher threat to users.

Conclusion

This study aimed to understand users' AI aversion behavior by conducting an online experiment with 315 participants. Our findings demonstrate the direct impact of several antecedents on AI aversion, including perceived human dissimilarity, perceived AI bias, and perceived social influence. Additionally, we identified intergroup differences between the simple and complex task groups in the hypothesized relationships. Overall, this study makes a valuable contribution to the IS literature by synthesizing existing research from different domains and developing an integrative model connecting four critical areas of AI application: AI characteristics, task characteristics, environmental characteristics, and user characteristics. Our findings provide empirical validation for the concept of AI aversion and offer insight into the factors that influence this behavior. In summary, this study sheds light on the underlying factors that contribute to users' aversion to AI systems, which contributes to both research and practice regarding applications of AI.

References

Boonstra, A., & Broekhuis, M. (2010). Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC health services research, 10(1), 1-17.


Dietvorst, B. J., & Bharti, S. (2020). People reject algorithms in uncertain decision domains because they have diminishing sensitivity to forecasting error. Psychological science, 31(10), 1302-1314.


Appendix

Table 1.A1: Demographics of Respondents

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Categorization</th>
<th>Number of respondents</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1. Male</td>
<td>192</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>2. Female</td>
<td>123</td>
<td>39</td>
</tr>
<tr>
<td>Age</td>
<td>1. &lt; 21</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2. 21 - 30</td>
<td>111</td>
<td>35.2</td>
</tr>
<tr>
<td></td>
<td>3. 31 - 40</td>
<td>118</td>
<td>37.5</td>
</tr>
<tr>
<td></td>
<td>4. 41 - 50</td>
<td>44</td>
<td>13.9</td>
</tr>
<tr>
<td></td>
<td>5. &gt; 50</td>
<td>42</td>
<td>13.3</td>
</tr>
<tr>
<td>Education</td>
<td>1. High school</td>
<td>44</td>
<td>14.0</td>
</tr>
<tr>
<td></td>
<td>2. College</td>
<td>88</td>
<td>27.9</td>
</tr>
<tr>
<td></td>
<td>3. Undergraduate</td>
<td>98</td>
<td>31.1</td>
</tr>
<tr>
<td></td>
<td>4. MS</td>
<td>66</td>
<td>21.0</td>
</tr>
<tr>
<td></td>
<td>5. PhD</td>
<td>18</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>6. Others</td>
<td>1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 1.A2: Goodness-of-fit indices for the structural model

<table>
<thead>
<tr>
<th>Fit Indices</th>
<th>Recommended Level of fit</th>
<th>Model Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute fit measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-squared</td>
<td></td>
<td>209.38</td>
</tr>
<tr>
<td>df</td>
<td></td>
<td>110</td>
</tr>
<tr>
<td>Chi-squared/df</td>
<td>&lt; 3</td>
<td>1.90</td>
</tr>
<tr>
<td>GFI (Goodness-of-fit Index)</td>
<td>&gt; 0.9</td>
<td>0.93</td>
</tr>
<tr>
<td>RMSE (Root mean square error of approximation)</td>
<td>&lt; 0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>Incremental Fit Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGFI (Adjusted goodness-of-fit Index)</td>
<td>&gt; 0.80</td>
<td>0.89</td>
</tr>
<tr>
<td>NFI (Normed Fit Index)</td>
<td>&gt; 0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>CFI (Comparative Fit Index)</td>
<td>&gt; 0.91</td>
<td>0.98</td>
</tr>
<tr>
<td>IFI (Incremental Fit Index)</td>
<td>&gt; 0.92</td>
<td>0.98</td>
</tr>
<tr>
<td>RFI (relative Fit Index)</td>
<td>&gt; 0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>Parsimony Fit Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCFI (Parsimony Comparative of Fit Index)</td>
<td>&gt; 0.5</td>
<td>0.70</td>
</tr>
<tr>
<td>PNFI (Parsimony Normed Fit Index)</td>
<td>&gt; 0.5</td>
<td>0.68</td>
</tr>
</tbody>
</table>
### Table 1.A3: Measures

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
<th>Items</th>
<th>Source</th>
</tr>
</thead>
</table>
| Perceived Human Dissimilarity (PHD)      | The degree to which one perceives an AI to be different from a human in the approaches or strategies it takes to complete a task. In practice, AI is highly data-dependent and decides based on its learning from data.                                                                                                         | -(AI) works in a way that is different from human  
- The characteristics of the strategy that (AI) follows are different from those of human strategies.  
- Humans can not imitate (AI) for (task) | Longoni et al., (2019)                                                                                                                                   |
| Perceived AI Bias (PAB)                  | The degree to which a user perceives an AI to “compound existing inequalities in socioeconomic status, race, ethnic background, religion, gender, disability or sexual orientation to amplify them and adversely impact inequalities in” the applied field                                                                 | -(AI) can generate biased results  
- (AI)'s decision can be discriminatory  
- (AI) can make unethical decisions | Pacheter et al., (2010)                                                                                                                                      |
| Perceived Lack of Autonomy (PLA)         | the degree to which a user perceives to have no control over using an AI                                                                                                                                                                                                                                                                                                                              | I can not control the way (AI) works  
- I can not change the decision model for (AI)  
- I do not have the authority to change the design of (AI) | Liang et al. (2015) from Ahuja and Thatcher 2005                                                                                                           |
| Perceived Social Influence (PSI)         | The degree to which users perceive that people important to them believe they should perform a specific IT behavior                                                                                                                                                                                                                   | People who influence my behavior would not like me to utilize (AI)  
- People in my social networks will devalue using (AI)  
- People who are important to me would not prefer that I apply (AI) for (task) | Chi et al. (2020)                                                                                                                                         |
| AI Aversion (AA)                          | The human assessment of an AI that manifests in a negative affective reaction with concomitant cognitions and behaviors towards that AI                                                                                                                                                                                                                                                               | I don’t want to use (AI) for (task)  
- I would not select (AI) for (task)  
- I would ignore the output of (AI) | Bhattacherjee et al. (2007)                                                                                                                                     |

All constructs are measured using a 7 points Likert scale (1=strongly disagree, 4=neutral, 7=strongly agree).

### Table 1.A4: Literature Review (Algorithm Aversion)

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Construct</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Inability to judge subjective issues</td>
<td>Niszczoa and Kaszás (2020), Castelo et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Perceived lack of mind</td>
<td>Bigman and Gray (2018)</td>
</tr>
<tr>
<td></td>
<td>Perceived effectiveness</td>
<td>Castelo et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Difficulty to understand</td>
<td>Yeomans et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Slow speed</td>
<td>Effendic et al (2020)</td>
</tr>
<tr>
<td></td>
<td>Inability to explain</td>
<td>Yeomans et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Perceived capabilities</td>
<td>Jussupow et al. (2020)</td>
</tr>
<tr>
<td></td>
<td>Higher accuracy</td>
<td>Pezzo and Beckstead (2020)</td>
</tr>
<tr>
<td>Environment</td>
<td>User</td>
<td>TASK</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Inability to learn</td>
<td>Berger et al. (2020).</td>
<td></td>
</tr>
<tr>
<td>Algorithm agency</td>
<td>Jago (2019)</td>
<td></td>
</tr>
<tr>
<td>Ability to modify (-)</td>
<td>Dietvorst et al. (2018)</td>
<td></td>
</tr>
<tr>
<td>Lack of incentivization (Solution: Context-specific behavioral design)</td>
<td>Burton et al. (2020)</td>
<td></td>
</tr>
<tr>
<td>Uniqueness neglect</td>
<td>Longoni et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>Perceived financial return</td>
<td>German and Merkle (2020)</td>
<td></td>
</tr>
<tr>
<td>Commonality (with humans)</td>
<td>Prahl and Swol (2017)</td>
<td></td>
</tr>
<tr>
<td>Seeing them perform</td>
<td>Dietvorst et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>High outcome ambiguity</td>
<td>Fuchs et al. (2016)</td>
<td></td>
</tr>
<tr>
<td>User's lack of decision control</td>
<td>Burton et al. (2020).</td>
<td></td>
</tr>
<tr>
<td>Combatting intuition</td>
<td>Burton et al. (2020)</td>
<td></td>
</tr>
<tr>
<td>Discomfort</td>
<td>Castelo et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>pre-(neg) bias</td>
<td>Liu et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>Excessive Reliance on personal knowledge (domain expertise)</td>
<td>Kim et al. (2016), Niszczota and Kaszás (2020), Logg et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>Prahl and Swol (2017), Dietvorst et al. (2015), Bigman and Gray (2018), Berger et al. (2020)</td>
<td></td>
</tr>
<tr>
<td>Appreciation of self-opinion (overconfidence)</td>
<td>Logg (2017), Logg et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>Unfamiliarity (-)</td>
<td>Logg et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>Perceived task difficulty</td>
<td>Effendic et al. (2020)</td>
<td></td>
</tr>
<tr>
<td>Task subjectivity</td>
<td>Castelo et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>Objectivity of decision</td>
<td>Logg (2017)</td>
<td></td>
</tr>
</tbody>
</table>

Note: A negative sign indicates that extant studies showed this factor to reduce algorithm aversion or to increase AI appreciation.
CHAPTER 2: WHAT TASKS LEAD TO AI AVERSION: THE MODERATED MEDIATION OF TRUST AND DISTRUST

Abstract

Effectively comprehending how users perceive and interact with artificial intelligence (AI) has become paramount in today's ever-evolving Information Systems landscape. While trust and distrust have been recognized as influential factors shaping users' attitudes towards IT artifacts, their intricate relationship with task characteristics and their impact on AI aversion remains largely unexplored. To bridge this critical research gap, we conducted an online randomized controlled experiment on Amazon Mechanical Turk. Our comprehensive analytic approach including structural equation modeling (SEM), ANOVA, and PROCESS conditional analysis allowed us to shed light on the intricate web of factors that influence users' AI aversion. We discovered that distrust and trust played interesting mediating roles between task complexity and AI aversion. Moreover, our study unveiled intriguing differences in these mediated relationships between subjective and objective task groups. Specifically, our findings demonstrate that, for objective tasks, task complexity can significantly increase aversion by reducing trust and significantly decrease aversion by reducing distrust. In contrast, for subjective tasks, task complexity has no effect on aversion, but can significantly increase aversion by enhancing distrust. By considering various task characteristics and recognizing trust and distrust as vital mediators, our research not only pushes the boundaries of the human-AI literature but also makes a significant contribution to the field of AI aversion.

Keywords: AI aversion, Trust, Distrust, Task Complexity, Task Subjectivity, Structural Equation Modeling

Introduction

Artificial Intelligence (AI) has garnered widespread attention in recent years due to its potential to revolutionize various industries and applications (Topol, 2019; Adir et al., 2020; Johnson et al., 2021). The widespread popularity and trial of AI, like chatGPT, show the potential of AI applications to improve human lives and society (e.g., Biswas, 2023; Sallam, 2023; Surameery & Shakor, 2023). We define AI as an information processing system that can interpret external data correctly, learn from it, and use those learnings to achieve specific goals or complete tasks through adaptation to its environment (Haenlein & Kaplan, 2019). From automating repetitive tasks and improving existing tools to analyzing vast amounts of data and informing better decision-making, the capabilities of AI are numerous and varied (Buchanan, 2019; Goodell et al., 2021). Artificial intelligence is extensively utilized in the financial service industry, e.g., in fraud detection, banking robo-advisors, and algorithmic trading (Buchanan, 2019). AI's potential is immense, and a sub-domain has emerged in finance name Fintech (e.g., Qi & Xiao, 2018; Cao et al., 2021). The healthcare industry is another domain of many AI applications (Shaheen, 2021; Mintz & Brodie, 2019). Some emerging areas of AI application in healthcare include AI-led drug discovery (Chan et al., 2019), clinical trials (Woo, 2019), and patient care (Neill, 2013). Other areas of potential AI applications include customer service (e.g., chatbots), education (e.g., tracking student progress and making recommendations), transportation (e.g., route optimization), manufacturing (e.g., identifying bottlenecks and predicting maintenance), retail (e.g., analyzing customer data and making personalized product recommendations), and marketing (e.g., targeted marketing campaigns) and others (Chen et al., 2020; Huang & Rust, 2018; Abduljabbar et al., 2019; Chien et al., 2020).

AI has the potential to transform industries and improve our daily lives, but its success depends on users' trust in it (Glikson, 2020). Adopting and integrating AI into various industries
and societies could be impeded by people's reluctance to use it (Jussupow et al., 2020). This resistance, or "aversion," towards AI can be caused by mistrust of technology in general (Castelo & Ward, 2021). If this reluctance persists, it can prevent AI from being fully utilized and limit its potential to create value for society (Castelo & Ward, 2021). For example, the potential benefits of a fully autonomous vehicle (i.e., AI-driven driverless car), such as improved safety, increased efficiency, and environmental and economic benefits, depend on people's willingness to adopt and use it (Ye & Yamamoto, 2019; Talebpour, & Mahmassani, 2016; Greenblatt & Shaheen, 2015). Hence, understanding the factors contributing to AI aversion is crucial to promoting users' acceptance and adoption of AI systems. This study investigates the effect of task complexity on AI aversion and the mediating role of trust and distrust in this relationship. We further analyze the moderating role of task subjectivity to develop a comprehensive understanding of task-effect of AI aversion.

While AI is an algorithm, it differs from a standard algorithm in various ways. An algorithm is "any well-defined computational procedure that takes some value, or set of values, as input and produces some value or a set of values, as output in a finite amount of time" (Cormen et al. 2022, p5). Hence, there is no emphasis on learning from input data or changing its course of action in the future (i.e., it does not improve dynamically). Nevertheless, AI can do both and can potentially augment human intelligence (Haenlein & Kaplan, 2019). Interestingly, both algorithm aversion and AI aversion literature indicate that people generally prefer humans over algorithm and AI systems for decision-making, even when the AI performs better (Castelo et al., 2019; Dietvorst et al., 2015; Longoni et al., 2019). This tendency exists both when people observe the performance of AIs and when they consider them hypothetically (e.g., Dietvorst et al., 2015). Although a few studies have found that people sometimes prefer human agents over AI (Logg et al., 2019), these cases are considered exceptions (Castelo & Ward, 2021). Moreover, studies on algorithm aversion mainly focused on users (e.g., Castelo et al., 2019; Burton et al., 2020; Liu et al., 2019; Dietvorst et al., 2015) or the algorithm (e.g., Bigman & Gray, 2018; Dietvorst et al., 2015; Castelo et al., 2019). Some studies focused on situational factors like lack of incentivization (Burton et al., 2020); perceived financial return (German and Markle, 2022). While these studies add invaluable information to our understanding of algorithm aversion, they primarily focus on algorithms. As mentioned earlier, the differences between algorithms and AI warrant specific studies where users' perception towards AI is studied. Moreover, despite the growing interest in AI aversion, the role of task complexity in this phenomenon remained unexplained. Task complexity is a critical factor affecting users' perception and attitudes toward an IT artifact and is likely to affect their acceptance and use of AI systems. Moreover, the subjective nature of a task can alter their reactions to and trust or distrust of an AI (Castelo et al., 2019). Hence, investigating the effect of task complexity and subjectivity on AI aversion can provide valuable insights into designing and implementing more user-friendly and trustworthy AI systems.

This study aims to address the research gap by examining the effect of task complexity on AI aversion and the mediating role of trust and distrust in this relationship. Specifically, we propose that users' perceived task complexity is positively related to this AI aversion and that this relationship is mediated by their trust and distrust of an AI system. Additionally, we propose that these relationships will vary based on the subjectivity of a task. We employ an online experiment using Amazon Mechanical Turk (MTurk) to achieve these objectives. The survey collects users' perceptions of task complexity, trust and distrust of AI systems, and AI aversion. We manipulate the complexity and subjectivity of tasks by randomizing participants assigned to one of four task
groups and by offering four different AI tools that vary in complexity and subjectivity of the underlying tasks they perform. We analyze the data using structural equation modeling by AMOS. To study the moderating effect of task subjectivity, we utilize the multigroup analysis option available in AMOS.

The findings of our study are expected to make several contributions to the fields of Human-AI interaction in general and to AI aversion in particular. First, the study contributes to the literature by enhancing our understanding of the factors related to a task contributing to users' AI aversion behavior. Doing so offers insights into designing more user-friendly and trustworthy AI systems. Second, it empirically validates the importance of considering users' trust and distrust as independent constructs while studying the AI aversion phenomenon. Finally, the study provides a framework for future research on AI aversion and develops a path to investigate further the complex interplay among task complexity, trust, distrust, and AI aversion.

**Literature Review**

In this section, we first explain how AI differs from generic algorithms and why it is crucial to study AI independently. Moreover, we present a review of the current literature on AI aversion specifically. We further discuss the importance of studying distrust and trust as two separate constructs and finally develop our hypotheses supported by current literature findings.

**AI vs. Algorithm**

Despite people using algorithms and AI interchangeably, they are fundamentally different. An algorithm is a finite set of rules that solve a specific problem through operations (Knuth, 1997). It can also be defined as a sequence of computational steps that transform inputs into outputs (Cormen et al., 2022). Hence, algorithms are a set of logical commands that take some inputs (either user-provided or automatically generated) and generate some outputs. Examples of algorithms can be a sorting algorithm, a search algorithm, and so on. AI systems, on the other hand, are designed to learn and adapt from their inputs and experiences through machine learning models and mimic human cognitive processes like perception, reasoning, and decision-making (Jordan & Mitchell, 2015; Russel, 2010). These ML and deep learning algorithms enable AI systems to identify patterns and make predictions or suggestions based on those patterns. This ability sets it apart from traditional generic algorithms that do not learn or adapt and generate the same output given a constant input every time. Unlike traditional algorithms, AI systems can refine their outputs by learning from their experiences, making them more accurate and reliable over time (Castelo et al., 2019). Some examples of AI systems are a self-driving car that can drive a car without any human intervention, a chess-playing computer that can play chess at a superhuman level, and so on. In simple terms, algorithms are like recipes that provide step-by-step instructions to solve a problem, while artificial intelligence is like a chef that can learn and adapt its cooking techniques based on experience and feedback. For example, an algorithm for image recognition could involve a set of predefined rules to analyze pixel patterns, identify edges, and match shapes to classify objects in an image. This algorithm would follow a fixed set of instructions and would not be able to improve its accuracy over time. In contrast, an artificial intelligence system for image recognition would utilize machine learning algorithms. Initially, it might receive a large dataset of images labeled with corresponding object categories. The AI system would analyze the data, learn patterns and features, and build a model to recognize objects in images. As the AI system encounters new images, it can continue to learn and improve its recognition accuracy through feedback mechanisms. The AI system's performance would evolve as it gains more experience and adapts its internal representations (e.g., Basu et al. 2010).
The availability of large volumes of data, low data storage costs, and high computing power have enabled companies to implement data-driven approaches for different purposes. However, traditional algorithms have limitations regarding their ability to adapt and learn from their inputs. For example, AI-based recommendation systems, such as those used by Amazon and Netflix, learn from a user's previous behavior, such as their past purchases or viewing history, to suggest new products or content they are likely interested in. The ability of AI systems to learn and adapt from their inputs makes them different from simple algorithms. While algorithms always produce the same output for a given set of inputs, AI systems can generate different outputs based on input data and previous experiences (Castelo et al., 2019).

Moreover, the adaptability and flexibility of AI systems make them highly useful in many applications, such as speech recognition, natural language processing, and image and video analysis (Jordan & Mitchell, 2015). However, because it learns from its input data, there are concerns about the use of AI, such as privacy, bias, and transparency (Calo, 2017). As AI systems can learn from their inputs, there is a risk that they may perpetuate biases or make unfair recommendations based on incomplete or biased data (Castelo et al., 2019). Therefore, it is vital to examine how these factors shape the opinion of users regarding its application and acceptance.

**AI Aversion**

We conducted a literature review using the Scopus database to identify studies that address AI aversion. We searched within article title, abstract, and keywords using keywords "artificial intelligence aversion" OR "AI aversion," which resulted in only twelve articles. Four of these twelve articles were related to 'advantageous inequality' (AI) aversion. After discarding these four, we investigated the remaining articles to identify their research questions, & key findings (table A1 in the appendix). We conducted a second search, including keywords "artificial intelligence appreciation" OR "AI appreciation," and that resulted in an additional article (total of thirteen). The other result is a book that discusses game AI (Lewis & Dill, 2015); hence, not included in Table A1.

Searching on Semanticscholars gave us the same result as Scopus. Finally, to broaden our search, we utilized an AI tool (Elicit from elicit.org) to identify relevant articles using the keyword "artificial intelligence aversion" OR "AI aversion." After several iterations and clearing irrelevant articles from the search result, we continued until the system stopped providing new suggestions. We did a similar search two times and collected three sets of articles. Because Elicit search includes all sorts (e.g., both peer-reviewed and non-peer-reviewed; ArXiv) of articles, we needed to clean the article list. Hence, after removing two ArXiv papers and a book and keeping one set of shared articles, we ended up with 43 articles from all sources combined. We investigated each article based on their research settings to determine if they address the AI aversion phenomenon. We removed any literature review article or non-peer review article, or articles that are not found using their DOI/title from our consideration, removed different versions of a single article, removed articles that studied generic algorithm aversion or utilized a general-purpose algorithm to study their research question but used the keywords artificial intelligence aversion and algorithm aversion interchangeably in their articles. That reduced our combined list to twenty-six articles (Table A1).

The literature on algorithm aversion suggests that individuals tend to exhibit aversion towards algorithms, whether consciously or unconsciously (Mahmud et al., 2022). Similarly, studies focusing specifically on AI reveal a general aversion towards this technology (Longoni et al., 2019; Niszczota & Kaszás, 2020; Wu et al., 2021; Filiz et al., 2022; Gherhes, 2018; Kim et al., 2022; Lanz et al., 2023). While online surveys are commonly used to investigate this
phenomenon, some studies employ experiments to explore how different factors contribute to aversion under specific conditions.

Perceived threat (Zhou et al., 2022; Stein et al., 2019), a lack of trust (Frank et al., 2021), and a perceived lack of transparency and fairness (Lopez & Garza, 2023) are common factors that increase aversion. Furthermore, users are more likely to avert AI if it makes errors, provides bad advice, or is perceived as erroneous (Dietvorst et al., 2015; Prahl & Van Swol, 2017; Kießling et al., 2021). Some studies focus on the nature of the task and suggest that users exhibit aversion towards AI when it comes to subjective tasks (Castelo et al., 2019; Raj et al., 2023), while others yield contradictory findings (Sohn et al., 2020). Additionally, people are less inclined to accept AI recommendations for experience or hedonic products than search or utilitarian products (Xie et al., 2022; Longoni & Cian, 2022). A few studies have explored non-traditional antecedents of AI aversion, such as political conservatism (Castelo & Ward, 2021) or anthropomorphizing AI (Cui, 2022), and have found positive effects on aversion.

Consequently, the existing literature on AI aversion needs to include comprehensive studies investigating the effects of different task characteristics and, notably, the role of trust and distrust in shaping users' aversion behavior.

**Distrust and Trust**

Trust and distrust are two cognitive constructs extensively studied over the years. Historically, these two concepts have been treated as direct opposites, but recent research has suggested that they are two distinct constructs with different antecedents and consequences (Dimoka, 2010; Liu & Wang, 2010; Rani et al., 2018). While trust refers to an individual's positive expectations of another's intentions and willingness to cooperate, distrust is characterized by an expectation of harm and a desire to protect oneself from adverse consequences (McAllister, 1995; Lewicki et al., 1998).

The proposition that trust and distrust are separate constructs has been supported by several studies (e.g., Lewicki et al., 1998; McKnight et al., 2002). These studies have suggested that trust and distrust affect behavior and decision-making differently. For example, Cho (2006) found that buyers' perception of online sellers' benevolence fostered trust, while their perception of competence reduced distrust. Other studies have demonstrated that trust is associated with positive affect, while distrust is associated with negative affect (Watson & Tellegen, 1985). Recent research has also examined the mediating effects of trust and distrust on various outcomes. Social exchange and affective events theories have been used to explain the relationship between trust, distrust, and their outcomes (Blau, 1968; Weiss & Cropanzano, 1996). For example, trust mediates the relationship between psychological contract fulfillment and organizational commitment, while distrust mediates the relationship between psychological contract breach and organizational disidentification (McAlearney et al., 2012). Similarly, trust and distrust and their interplay are essential to successful advanced information technology deployment (Baba, 1999). Both constructs have been studied in other domains and experimental settings, including but not limited to online banking (Benamati & Serva, 2007); e-commerce (Ou & Sia, 2009); honest learning-oriented assessment (Carless, 2009); social reality construction (Govier, 1994) and so on.

Moreover, neurological research indicates that trust and distrust are associated with two different brain regions (Benbasat et al., 2010). For example, Dimoka (2010) conducted a specific neuroimaging study where she applied functional neuroimaging (fMRI) tools and observed the location, timing, and level of brain activity associated with trust and distrust to understand their underlying dimensions. They manipulated seller profiles to manipulate their trust and distrust
and observed how the manipulation activated different brain areas. The study identifies that trust and distrust are associated with two different areas in the brain with varying effects. Therefore, it is crucial to differentiate between trust and distrust as it has significant consequences for behavior and decision-making, such as human-AI interactions.

**Task Complexity**

Task complexity has been used as an essential variable in different research streams, including goal-setting studies (e.g., Campbell & Gingrich, 1986); decision-making (Larichev & Moshkovich, 1988). In the literature, there are three ways to define task complexity: i) structuralists: which defines task complexity by its structure; ii) resources requirements: which defines task complexity by its resource requirements; and iii) interaction viewpoints: which defines it as a product of human-task interaction (Liu & Li, 2012). According to the structuralist's view, a complex task can have many elements which can interconnect them (e.g., Campbell, 1988; Ham et al., 2012). On the other hand, according to the resource requirements view, task complexity is defined by the requirements during the human information processing stage. Resource requirements can be of many kinds, e.g., cognitive demand (Campbell, 1988; Park, 2009); physical and mental demand (Li & Wieringa, 2000); cognitive efforts (Bettman et al., 1990), etc. Moreover, the resources can be related to visual, auditory, cognitive, and psychomotor resources or to knowledge, skills, and time (McCracken & Aldrich, 1984; Gill, 1996; Bystrom & Jarvelin, 1995; Nembhard & Osothsilp, 2002). Therefore, a complex task will require a user to invest more resources to complete it (Liu & Li, 2012). Finally, the interaction view considers task complexity as the result of the interaction between a task and the characteristics of its performer. Such characteristics include the user's needs, prior knowledge, and experience). Advocates of this view consider task complexity as a relative term and suggest investigating task complexity as a subjective measure from the participant's standpoint (Liu & Li, 2012; Gonzalez et al., 2005). Influenced by these perspectives mentioned above and research contexts, various definitions of task complexity have been proposed (e.g., Byström & Jarvelin, 1995; Vakkari, 1999). In our study, we adopt a specific definition of task complexity; building on the works of interaction viewpoint, we define that a task is complex when its complexity exceeds the perceived capacity of the task performer (Liu & Li, 2012). Hence, complex tasks will require users to allocate higher levels of attention and engage in deeper cognitive processing (Bonner, 1994). Consequently, the complexity of a task can be intensified by factors such as increased instructions that users must follow or a more significant number of dimensions they need to consider.

It is worth noting that while task complexity and task difficulty are sometimes used interchangeably (Bell & Ruthven, 2004), they are different concepts (Jacko & Ward, 1996; Robinson, 2001). Researchers have recently proposed distinguishing between these concepts (Aula et al., 2010). Here, we differentiate between these concepts and measure task complexity from the participant's perspective. As task complexity is an individual's perception of a task, the complexity level may vary from one user to another. We accept Valcour's (2007, p. 1513) definition of task complexity as "the level of stimulating and challenging demands associated with a particular task." Therefore, a complex task will be associated with a more significant challenge and require a higher cognitive capacity use (Fried et al., 2002).

**Distrust and AI Aversion**

Distrust refers to a user's unwillingness to be vulnerable to an AI because the AI may act in a way that will negatively affect the user (Dimoka, 2010). In other words, distrust is the user's expectation of action from the AI that could be harmful (Luhmann, 1979). Distrust may arise
from a user's perception of an AI's negative motives, poor capabilities, or harmful acts (Ullmann-Margalit, 2004). When users believe that the AI does not intend to act in their best interest in each task, distrust may increase. This suspicion and skepticism can lead to distrust (Lewicki et al., 1998; Sitkin, 1993) and even generate intense negative emotions that impact users' acceptance of a service or an IT artifact (Prakash & Das, 2022). For instance, Benamati & Serva (2007) found that users who assume harmful motives of the online banking system are more likely to take defensive actions.

Distrust can significantly reduce users' confidence in AI and increase their fear of harm or harmful intentions or actions (Govier, 1994). As a result, distrust can be a significant obstacle to user adoption of AI technology. For instance, distrust has been shown to shape users' opinions regarding accepting recommendations (Lee & Cho, 2023), interactions with online personalization agents (Chau et al., 2013), or customer loyalty (Lee et al., 2015). Distrust can keep users passive or even convert an active user to a passive user (Lee et al., 2015). Both theories and practices indicate that distrust is vital in shaping users' acceptance behavior (Prakash & Das, 2022). fMRI studies suggest that distrust activates the insular cortex area in the brain (Dimoka, 2010). The insular cortex is associated with fear, worry, and similar feelings that help people avoid unwanted negative interactions (Dimoka, 2010; Wicker et al., 2003; Preuschoff et al., 2008; Rilling et al., 2008). Hence, distrusting someone or something will alert a user, causing psychological discomfort that makes them act cautiously, leading them to guard themselves by rejecting the use or acceptance of the IT artifact (McKnight & Chervany, 2001a; Ou & Sia, 2010).

In the case of AI, the prospective unreliability of an AI can be an obvious source of distrust in it (Dwork & Minow, 2022). While several factors can contribute to this unreliability, an unreliable AI will receive aversion from its users, apart from the source. Therefore,

**H1: Distrust in AI is positively associated with AI aversion**

**Trust and AI Aversion**

Trust is an essential psychological construct studied extensively in Information Systems, Marketing, and Management literature (e.g., Suh & Han, 2002; Brockner et al., 1997; Wang et al., 2016; Rahi et al., 2017). Similarly, most AI studies also concentrated on trust to study users' attitudes. Please see (Glikson et al., 2020) for a comprehensive list of articles. Trust in AI is defined as a user's willingness to be vulnerable to an AI, believing that the AI will act according to his/her confident expectations (Dimoka, 2010). It is the user's confidence in an AI's ability to act according to their expectations and intentions (Dimoka, 2010). A user may trust an AI based on its credibility regarding its perceived competency, honesty, reliability, and benevolence (Dimoka, 2010). Trust has been widely studied across many domains as an antecedent of positive actions. For instance, in e-commerce, trust is a crucial factor that determines a customer's potential to transact (Grabner-Kraeuter, 2002). According to Mayer et al. (1995), the primary motivators for a person's behavior are their beliefs about the trustworthiness of a trustee (Mayer et al., 1995; Kim et al., 2008; Meskaran et al., 2010). Trust plays a significant role in influencing the purchase intention of consumers (Tariq & Eddaoudi, 2009; Van Der et al., 2003), the intention to use e-government systems (Hooda et al., 2022), m-banking (Ramos et al., 2018), online games (Wu & Liu, 2007), and technology-enabled bank channels (Dimitriadis & Kyrezis, 2010).

According to the Social Exchange Theory (Cropanzano & Mitchell, 2005), individuals engage in social interaction based on a cost-benefit analysis, and trust plays a vital role in this process. Users are highly likely to engage in social exchange when they trust the other party will
act in their best interest. In the case of AI, users will evaluate the AI's willingness to act in their best interest and form inherent beliefs based on those evaluations, which are represented as trust and influences a user's attitude towards that AI. Trust can address the uncertainty of using an AI and clarify its competence, reliability, or potential behavior (Gulati & Gargiulo, 1999). Therefore, trust will determine users' intentions and behaviors, including their willingness to be vulnerable to AI's decisions (Benamati & Serva, 2007). If trust is high, users will demonstrate less aversion towards an AI and vice versa.

**H2: Trust in AI is negatively associated with AI aversion**

**Task Complexity and Distrust**

As mentioned earlier, task complexity refers to the cognitive processing and attentional capacity that a task demands from a user (Bonner, 1994). It is influenced by factors such as the number of instructions that need to be followed and the number of dimensions a user needs to consider for executing a task. Therefore, users require a higher level of skill and expertise to perform a complex task, and their opinion varies based on the task's complexity level. An AI performing a complex task will receive a different reaction than an AI performing a simple task (Liu & Li, 2012). Because AI is highly data dependent, an AI performing a complex task must be trained on a significantly higher number of features (i.e., parameters) and a larger dataset than a simple one. For example, GPT-3.5, a large language model released in 2022 by OpenAI, was trained on 175 billion parameters and used neural networks to generate human-like text on numerous topics. However, its capability is limited to text data only. GPT-1 had 0.12 billion parameters, and GPT-2 had 1.5 billion parameters. While the number of parameters for GPT-4 is unknown, it is predicted to be more than a trillion (Lutkevich, 2023; Chowdhery et al., 2022). Hence, the possibility of accruing bias during model training is relatively higher for a complex-task AI than a simpler one.

For example, identifying a patient's race from medical images is a complex task for a human agent. However, AI can identify participants' races from medical images (Gichoya et al., 2022). Amazon's "Rekognition" facial recognition technology displayed significant biases in identifying individuals with darker skin tones, especially women (O'Brien et al., 2019), and the risk assessment algorithm COMPAS was more likely to incorrectly flag black defendants as high-risk compared to white defendants (Angwin et al., 2016). Hence, task complexity increases the possibility of bias in an AI and reduces its reliability. Moreover, a complex-task AI's high number of features and complex nature make conducting a post-hoc root cause analysis difficult. Because as the number of features increases, it introduces more potential sources of error or noise in the data, lower explainability behind its decisions, and higher obscurity in its decision-making processes (Avgerinos & Gokpinar, 2017). Hence, increasing task complexity affects users' opinion regarding an AI that performs that task and creates a negative feeling about the AI's conduct (Lewicki et al., 1998). Users fear that the AI might not care about them or might hurt them with its decisions or actions (Govier, 1994). Therefore, we propose-

**H3: Task complexity is positively associated with distrust in AI.**

Trust is a critical factor in the user's acceptance of an AI system, and it is determined by their belief and expectation that the AI will act in a manner consistent with their expectations (Pavlou & Gefen, 2004). Users' trust is based on an AI's credibility and competence, and it is dependent on their knowledge of the AI and their ability to predict the outcome of its actions (Dimoka, 2010). As such, the better users understand an AI and its capabilities, the higher their trust will be in it (Johnson & Grayson, 2005). Trust is not a simple feeling or emotion that can be activated by affection; it is initiated by cognition and involves rational evaluation of the AI's
competency and reliability based on available knowledge (McAllister, 1995; Jeffries & Reed, 2000). Trust also activates the part of the brain that deals with reward anticipation, behavior prediction, and uncertainty calculation (Dimoka, 2010). Therefore, building and increasing trust in an AI requires users to rationally evaluate the AI's capabilities and reliability based on available knowledge and reasoning.

As a task becomes more complex, the difficulty for users to understand how and why an AI system arrives at its recommendations or conclusions increases. This issue is called the black-box problem, where the AI system uses complex algorithms and data sets that are difficult for humans to interpret or explain. With a complex task, an AI system must consider many more features, making it even more challenging for users to comprehend and leading to opacity or the inability to investigate the internal mechanism of work (Adadi & Berrada, 2018). As a result, when users need help understanding an AI's decision-making process or how it works, they may become skeptical of its effectiveness and outcomes. This reduces their trust in AI, particularly for complex tasks commonly perceived as higher risk (Reason, 1990). Hence,

\[ H4: \text{Task complexity is negatively associated with trust in AI}. \]

**Moderating Role of Task Subjectivity**

A subjective task is one the output of which is open to the interpretation of participants for evaluation and relies on their personal feelings, opinions, or intuition (Castelo et al., 2019). On the other hand, evaluating the output of objective tasks is based on factual information and is not subject to personal interpretation. Users' reactions to an AI will likely depend on their perception of the AI's abilities. For example, technology adoption theories suggest that the perceived usefulness of the technology in completing a task is a fundamental determinant of technology adoption (Davis et al., 1989). Generally, people believe AIs cannot handle tasks requiring factual factors and emotional dimensions (Purdy et al., 2019). For instance, when judging the quality of a picture or a song, a human will consider both objective measures (e.g., resolutions) and subjective qualities (e.g., feeling good, matching context, emotion). However, an AI will consider only facts, making it unsuitable for tasks that require human abilities such as intuition and warmth (Haslam, 2006; Haslam et al., 2008). Moreover, AIs are considered to have agency (the ability to plan and action) but not experience (the ability to experience emotion subjectively) (Castelo et al., 2019; Gray et al., 2007). Hence, users may perceive an AI as more suitable to perform an objective task than a subjective task. Hence, their opinions regarding AI will vary based on the nature of the work (subjective vs. objective). This phenomenon can further be explained using the theory of constructivism, which suggests that people actively construct their knowledge and understanding of the world based on their experiences and perceptions (e.g., Bonder, 1986). Hence, people's interpretations of the subjectivity of a task can vary widely, leading to different responses to an AI performing that task.

Capturing the nuances of human judgment and preferences for subjective tasks is relatively tricky for artificial intelligence (AI) because such tasks are open-ended and ambiguous. For instance, evaluating artwork, a complex subjective task, requires a subjective assessment from users. People's opinions on such tasks can differ based on various factors, including experience, tastes, and personal preferences. According to Cheek et al. (2021), individuals feel confident in their personal beliefs and judgments and perceive those who hold different opinions as biased, harmful, or incorrect. However, because AI operates based on the data it is trained on and is limited by the quality of the training data (Parikh et al., 2019), differences in personal judgments can lead to undue biases or errors in the AI's recommendations (Ntoutsi et al., 2020; Daneshjou et al., 2021). Given the high uncertainties associated with complex tasks and their inherent
subjectivity, users will be highly skeptical of an AI intended for similar tasks. Therefore, people will perceive such an AI as incapable or unreliable of completing a complex subjective task, resulting in increased distrust (Dwork & Minow, 2022).

In contrast to subjective tasks, objective tasks have a clearly defined goal or outcome and do not depend on feelings or personal interpretations (Castelo et al., 2019). This makes them less ambiguous and more easily measured than subjective tasks. Users can assess the performance of an AI for objective tasks by comparing them to established benchmarks, which can increase their confidence in the AI. Although the possibility of making errors still exists, the absence of personal judgments and biases, a clear benchmark for performance evaluation, and a lack of ambiguity can mitigate users' distrust in an AI for objective tasks, even when those tasks are complex. Therefore,

\[
H5a: \text{Task subjectivity moderates the relationship between task complexity and distrust such that the relationship is more positive for subjective tasks than objective tasks.}
\]

Subjective tasks require a higher degree of personal judgment and interpretation, which can increase the perceived uncertainty and unpredictability of the outcome (Phillips-Wren & Adya, 2020). When a task is subjective, there may be multiple valid perspectives or criteria for evaluation, and the outcome may be influenced by personal biases, values, or preferences (e.g., Luo et al., 2020; Díaz et al., 2018; Waseem, 2016). As a result, individuals may find it more difficult to accurately evaluate the competence and reliability of the system responsible for completing the task, and this can lead to lower levels of trust. When a subjective task is also complex, it can exacerbate these challenges by increasing the cognitive effort required to process and interpret the information necessary to complete the task. This can make it more difficult for individuals to understand and evaluate the outcome, leading to more significant uncertainty and ambiguity (Mishel 1984, 1988). As a result, individuals may feel less confident in the outcome and the system responsible for completing the task, leading to lower levels of trust.

In contrast, objective tasks typically have more explicit criteria for evaluation and require less subjective interpretation. This can make it easier for individuals to understand and evaluate the outcome and the person or system responsible for completing the task, leading to higher levels of trust. While objective tasks can still be complex, the concrete, tangible nature of the task can make it easier to evaluate the outcome and reduce uncertainty (e.g., Inbar et al. 2010), mitigating the negative impact of complexity on trust. Overall, the subjective nature of the task can increase the perceived uncertainty and unpredictability of the outcome, exacerbating the negative impact of complexity on trust.

\[
H5b: \text{Task subjectivity moderates the relationship between task complexity and trust in a way that the relationship is more negative for subjective tasks than for objective tasks.}
\]

While distrust is known to be associated with AI aversion, the effect of distrust may vary depending on the subjectivity of the task. In particular, the level of risk involved in a subjective task is often ambiguous. Hence, users' perceived risk will vary (Weber & Hsee, 1998). The ambiguity in perceived risk for a subjective task may stem from two sources: uncertainty from the unpredictability of the task object or uncertainty from the unpredictability of the acceptance of that task's result. For instance, in a recidivism risk assessment task, where the likelihood of a convicted criminal reoffending is assessed, uncertainty primarily arises due to the object of the task—the convicted person—who depends on the subjective evaluation of another person. Despite gathering information about the convicted person's history, social, demographic, mental health, and other factors, the final decision relies on a judge's subjective decision regarding the likelihood of the convict's future actions (Porter & Brinke, 2009). On the other hand, tasks such
as composing music or drawing a picture rely on the subjective evaluations of others, such as consumers, and their results are thus unpredictable and uncertain (Jordanous, 2012).

Because there is an element of subjectivity involved in the process, the data used for training an AI system can also be insufficient or subjective (Hajian et al., 2020; Barocas & Selbst, 2016; Asan et al., 2020). Therefore, if users distrust an AI for any reason, they will likely exhibit a stronger aversion toward that AI (Dietvorst et al., 2015). As the literature on algorithm aversion suggests, users often prefer human agents for subjective tasks because they can go beyond facts if necessary and consider factors that might be absent from the data (Castelo et al., 2019). This is especially critical for subjective tasks, where users lack any set measures to re-evaluate the subjective performance of an AI and reduce their distrust. Thus, users may stay uncertain and confused, which might convince them to choose a safer option and avert the AI altogether.

The scenario significantly differs depending on whether the task is objective or subjective. When users harbor doubts and distrust toward an AI for various reasons, the ability to verify the AI’s performance becomes crucial in developing a comprehensive understanding of the associated risks and benefits. This verification process aligns with the well-established prospect theory proposed by Kahneman and Tversky (1979), which posits that individuals prefer specific outcomes over uncertain ones. In the case of objective tasks, individuals can leverage the presence of set measures or benchmarks to compare and evaluate the AI’s performance, enabling them to make informed decisions regarding the risks involved in utilizing the AI for that particular task. The existence of known standards for comparison instills a sense of certainty and empowers users to rely on established measures when assessing the effectiveness of the AI. Consequently, this evaluation process aligns with prospect theory, as individuals' preference for specific outcomes facilitates objective judgment of the AI's performance. However, the situation significantly differs when the task is subjective. In subjective tasks involving personal opinions, emotions, or aesthetic judgments, the absence of clear and universally agreed-upon evaluation measures poses challenges for users. As a result, individuals need help objectively assessing the AI's performance for subjective tasks since there is no objective standard against which to compare its output. In these instances, the uncertainty and perceived risks associated with employing AI for subjective tasks are heightened, as individuals cannot rely on the same calculated sense of risk provided by objective measures. Therefore,

\[ H5c: \text{Task subjectivity moderates the relationship between distrust and AI aversion in a way that the relationship is more positive for subjective tasks than for objective tasks} \]

Trust is negatively associated with AI aversion, and this relationship can vary based on the nature of the task, i.e., subjective versus objective. For subjective tasks, users need to rely on the subjective evaluations of an AI and cannot verify the outcome directly, while for objective tasks, they can do so (Burrell, 2016). Existing literature suggests that verification can function as a means of reassurance and will provide a sense of control for users, as the act of verification can increase the perceived credibility and trustworthiness of the AI agent by improving transparency and accountability for AI recommendations (Veale & Binns, 2017; Jobin et al., 2019). For example, when AI provides a patient with a health diagnosis and treatment recommendations, physicians can verify whether the AI output is legitimate. Because the diagnosis accuracy and treatment appropriateness are relatively objective, there is a universal standard regarding the quality of the AI output. It is feasible for physicians to cross-check the AI output, and different physicians will likely reach the same conclusion. This consensus can offer a feeling of control for objective tasks, reassuring users that their trust in AI is justified and further enhancing their trust's impact on AI aversion.
However, in the case of a subjective task, such as AI image generation, cross-checking with experts or others might result in different opinions. Users may receive differing and contradicting feedback because it depends on a person's taste, liking, preferences, experience, and so on. A universal standard for evaluating the AI output needs to be clarified for users, making them second-guess their trust in AI. As a result, the impact of trust on AI aversion may be weakened for such subjective tasks. Taken together, at the same level of trust in an AI, trust tends to reduce AI aversion to a greater extent for objective tasks (Castelo et al., 2019). Therefore, we propose,

$H5d$: Task subjectivity moderates the relationship between trust and AI aversion in a way that the relationship is more negative for objective tasks than for subjective tasks.

**Methodology**

**Procedure**

We conducted a 2-by-2 randomized experiment using Amazon MTurk (Mechanical Turk), where each participant was randomly assigned to use AI to complete one of four tasks. MTurk is an open platform that has been widely used in information systems and other research domains for running social experiments (e.g., Paolacci et al., 2010; Shin et al., 2020). The data was collected using Qualtrics surveys. Because we wanted to study the perception and AI use of general AI users, our inclusion criteria are that the participants need to be from the United States and aged 18 or older. We informed all participants that the study was anonymous, and that no personal information would be collected. Each participant received a randomly assigned task and completed the task using an AI tool. They then completed the survey to answer questions based on their experience of using the AI. We collected three hundred seventy-five valid responses from 500 participants after filtering out 125 responses that failed the quality control questions.

**Experimental Design**

Our 2-by-2 experimental design is based on two binomial factors: task complexity (simple vs. complex) and task objectivity (subjective vs. objective). This design yields four groups, as shown in Table 1. In each group, the users used an AI tool to complete the task. Active user engagement and hands-on experience are known to enhance user involvement (Bonwell & Eison, 1991), which allow users to provide a more comprehensive evaluation of an AI (Wu & Sankar, 2013). By doing so, our research design provides a higher level of realism to users. We randomly assigned the participants to these four task groups. We employed a between-subject design in which each participant was assigned to exactly one task group.

<table>
<thead>
<tr>
<th></th>
<th>Subjective</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>Movie Selection</td>
<td>Image Detection</td>
</tr>
<tr>
<td>Complex</td>
<td>High Quality Image Generation</td>
<td>Health Diagnosis</td>
</tr>
</tbody>
</table>

The Simple Subjective group features an AI movie recommender that personalizes movie suggestions for users based on their inputs. Users interact with the AI to receive suggestions by answering a series of questions. A deep learning-based AI classifies images into relevant categories in the Simple Objective group. After the user uploads an image, the AI tool identifies the principal object within the image, such as a forest or building. The Complex Subjective group includes a photo up-sampling AI that converts low-resolution images to high-resolution outputs. This tool leverages self-supervised deep learning to fill in missing pixels and improve image quality. Finally, the Complex Objective group contains a health advisor AI that assesses the participant's health conditions. By asking questions, the tool provides personalized health suggestions based on the user's answers. For example, the movie recommendation AI group is
asked to imagine that they want to watch a movie and find an AI tool that suggests a movie list based on their mood, occasion, genre, released year, etc. We provided step-by-step instructions on how to navigate the AI tool and how to use it. The healthcare tool is like the movie recommendation tool in that the user input is answers to some questions that the AI ask them. However, for image classification and high-quality image generation AI tools, the user must input images we provide through Google Drive.

After entering the study website, each participant was asked to provide consent to participate in the study. They were then randomly assigned to one of the four groups and asked to perform a task using a provided AI tool. A vignette asked them to imagine that they were in a situation where they needed to perform the task. After the participants completed the task, they were taken to an online survey to answer questions about the research constructs.

**Measurements**

We utilize validated instruments from the literature. All items are reflective and can be found in Table A3 in the Appendix. Specifically, we adapted three items to measure AI aversion from Bhattacherjee and Hikmet (2007). Three items for trust and three items for distrust are adapted from Dimoka (2010), McAllister (1995), and Lewicki et al. (1998). Finally, we measure task complexity and subjectivity by asking participants to indicate their perception of the task’s subjectivity and complexity following Wang et al. (2014) and Castelo et al. (2019). We also measure factors such as age, education, and gender following established literature in algorithm aversion (e.g., Dietvorst et al., 2015; Dietvorst et al., 2020). In addition, we record each task’s completion time, which may impact users’ evaluations. We validated our measures using a pretest on 100 undergraduate students before conducting the MTurk experiment. We found high reliability and validity of the measures.

**Analytical Approach**

Based on the following considerations, we assess the research model using Covariance-Based Structural Equation Modeling (CBSEM). First, CBSEM is widely used to examine intricate connections between latent variables in social science research (Astrachan et al., 2014). Second, it can simultaneously analyze complex measurement models and structural relationships involving multiple layers of constructs (Wang & Wang, 2019; Astrachan et al., 2014). Given our research model's inclusion of latent variables and complex structure with mediating and moderating relationships, CBSEM is suitable for data analysis. IBM AMOS 27 is used to perform the analysis. In addition, we complement our CBSEM analysis with ANOVA by comparing trust, distrust, and AI aversion across the four task groups. Finally, we supplement our analysis by testing a conditional moderated mediation effect utilizing SPSS PROCESS by Hayes (2017).

**Results**

**Measurement Model**

We assess the measurement model by examining each construct's reliability, convergent validity, and discriminant validity (Liang et al., 2015). Table 2 reveals that all composite reliability and Cronbach's alpha values are higher than 0.7, demonstrating acceptable construct reliability (Hair, 2009). The item loadings are above the acceptable level of 0.6 (Chin, 1998), indicating good convergent validity. To evaluate discriminant validity, we compare the factor loadings on each construct to the loadings on other constructs (Fornel & Larcker, 1981) to ensure that a construct's factor loading is higher than its loading on other constructs (Fornell & Larcker, 1981). Our results demonstrate adequate discriminant validity. Furthermore, the square root of
Table 2.2. Descriptive statistics and correlations

<table>
<thead>
<tr>
<th>Construct</th>
<th>Mean</th>
<th>SD</th>
<th>CR</th>
<th>Alpha</th>
<th>AVE</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Task Complexity</td>
<td>3.43</td>
<td>1.27</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Task Subjectivity</td>
<td>3.55</td>
<td>1.15</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.385**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Distrust</td>
<td>4.19</td>
<td>1.73</td>
<td>0.90</td>
<td>0.89</td>
<td>0.70</td>
<td>0.322**</td>
<td>0.385**</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Trust</td>
<td>5.56</td>
<td>0.94</td>
<td>0.89</td>
<td>0.77</td>
<td>0.61</td>
<td>0.006</td>
<td>0.001</td>
<td>-0.099</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>5. AI Aversion</td>
<td>3.98</td>
<td>1.86</td>
<td>0.90</td>
<td>0.94</td>
<td>0.80</td>
<td>0.287**</td>
<td>0.334**</td>
<td>0.788**</td>
<td>-0.244**</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Notes: * p < 0.05; ** p < 0.01; CR: Composite Reliability; Alpha: Cronbach’s Alpha; S.D.: Standard Deviation. The bold values on the diagonal are the square roots of AVE.

Manipulation Check

We conduct t-tests to identify differences in the demographic variables among different task groups. The result shows no statistically significant differences [task: simple vs complex (age: F = 2.22, p > 0.05; education: F = 1.02, p > 0.05; gender: F = 0.44, p > 0.05); task: subjective vs objective; (age: F = 0.63, p > 0.05; education: F = 1.67, p > 0.05; gender: F = 0.59, p > 0.05)]. This suggests that the random assignment of participants was successful.

We then performed t-tests to evaluate the effectiveness of our manipulation of task complexity and task subjectivity. The results reveal a significant difference between the complex task groups and the simple task groups (µ<sub>complex</sub> = 3.83, µ<sub>simple</sub> = 3.12, t(373) = -5.58, p < 0.01), and between the subjective task groups and the objective task groups (µ<sub>subjective</sub> = 3.95, µ<sub>objective</sub> = 2.66; t(373) = 11.31; p < 0.01). This suggests that our manipulation of the independent variable and moderating variable is effective.

Structural Model

We performed a bootstrapping procedure of 2000 resamples on a sample of 375 responses. Figure 2 shows the result of our base model. As indicated, the model explains 76% of the variance for AI aversion. The model shows an excellent fit with the data (Chi-squared/df = 2.71, CFI = 0.96; NFI = 0.93; RMSEA = 0.06) (Suki, 2014). Task complexity has no significant effect on distrust (β = -0.05, p > 0.05), failing to support H1, but has a negative effect (β = -0.12, p < 0.05) on trust, supporting H2. We also find support for H3, and H4 as distrust increases AI aversion (β = 0.85, p < 0.01) and trust decreases AI aversion (β = -0.18, p < 0.01).
To test hypotheses H5a to H5d, we conduct a multi-group analysis using AMOS, following Floh and Treiblmaier (2006) and Bakker et al. (2003). The analysis compares a model with freely estimated structural parameters to a competitive model where the parameters were equal between the subjective and the objective groups. The Chi-square test results indicated a significant difference between the two groups ($df=15; \text{CMIN} = 60.43, p <0.01$), implying a meaningful distinction between the two task groups. Table 3 shows the result of our base and subgroup model testing.

**Moderating Effect of Task Subjectivity**

Figure 3 shows the comparative results between the subjective task group and the objective task group. The effect of task complexity on distrust is more substantial and opposite for objective tasks. Similarly, task complexity's effect on trust is more substantial and significant for an objective task than a subjective one. Moreover, the effects of distrust on AI aversion are significant for both task groups with marginal differences in their strengths. In contrast, the effects of trust on AI aversion are also significant for both task groups but the effect for the objective task group is more than doubled than that for the subjective task group. Overall, our multigroup analysis results support H5 and H5d, but do not seem to support H5b and H5c.
Figure 2.2. Model testing results for a subjective task (top) and an objective task (bottom)

While the multigroup analysis suggests that the two models differ between the subjective and objective task groups, it is still unclear whether each specific conditional indirect effect significantly differ between the two groups. To provide more precise findings, we conducted conditional analysis by using the PROCESS macro (Hayes, 2017). Moreover, we conducted ANOVA by comparing trust, distrust, and AI aversion across the four task groups to supplement our findings from CBSEM. The ANOVA result revealed that users’ distrust in AI (Figure 4) was significantly affected by task subjectivity (subjective vs. objective): [F(1, 371) = 4.13, p < 0.05]; non-significantly affected by task complexity (complex vs. subjective): [F(1, 371) = 0.382, p > 0.05], and significantly affected by the interaction between these two factors: [F(1, 371) = 46.92, p < 0.01]. When the task was subjective, distrust was significantly higher for complex tasks than for simple tasks: (µ\text{simple} = 3.42, µ\text{complex} = 4.48, t(177)= -3.97, p < 0.01). Conversely, when the task was objective, distrust was significantly lower for complex tasks than for simple tasks: (µ\text{simple} = 4.93, µ\text{complex} = 3.67, t(194)= 5.91, p < 0.01).

Figure 2.3. Effect of task complexity on distrust for objective vs subjective tasks, error bars indicate 95% confidence interval.

The ANOVA result revealed that users’ trust on AI (Figure 5) was similarly affected by task subjectivity (subjective vs objective): [F(1, 371) = 34.54, p < 0.01]; significantly affected by
task complexity (complex vs subjective): \[ F(1, 371) = 7.12, p < 0.01 \], and significantly affected by the interaction between these two factors: \[ F(1, 371) = 13.90, p < 0.01 \]. When the task was subjective, trust was not significantly different between complex and simple tasks: (\( \mu_{\text{simple}} = 5.77, \mu_{\text{complex}} = 5.87, t(177) = -0.77, p > 0.05 \)). But when the task was objective, trust was significantly lower for complex tasks than for simple tasks: (\( \mu_{\text{simple}} = 5.57, \mu_{\text{complex}} = 4.98, t(194) = 4.41, p < 0.01 \)).

Figure 2.4. Effect of task complexity on trust for objective vs. subjective tasks, error bars indicate 95% confidence intervals.

To examine the effects of task complexity on AI aversion, we run two sets of ANOVA tests. In the first test, task complexity was the independent variable, task subjectivity was the moderator, and AI aversion was the dependent variable, without covariates. The result indicated that users’ AI aversion (Figure 6) was significantly affected by task subjectivity (subjective vs objective): \[ F(1, 371) = 6.87, p < 0.01 \]; non-significantly affected by task complexity (complex vs subjective): \[ F(1, 371) = 0.778, p > 0.05 \], and significantly affected by the interaction between these two factors: \[ F(1, 371) = 13.02, p < 0.01 \]. When the task was subjective, AI aversion for complex tasks was marginally higher than AI aversion for simple tasks: (\( \mu_{\text{simple}} = 3.42, \mu_{\text{complex}} = 3.94, t(177) = -1.77, p < 0.1 \)). But, when the task was objective, AI aversion for complex tasks was marginally lower than AI aversion for simple tasks: (\( \mu_{\text{simple}} = 4.60, \mu_{\text{complex}} = 3.75, t(194) = 3.47, p < 0.01 \)).
In the second ANOVA test, we introduced distrust and trust as covariates. The effects of task complexity, task subjectivity, and their interaction on AI aversion disappeared after adding these covariates. As Figure 7 shows, for task complexity, \( F(1, 369) = 1.82, p > 0.05 \); task subjectivity \( F(1, 369) = 0.07, p > 0.05 \); task complexity x task objectivity, \( F(1, 369) = 2.20, p > 0.05 \). Trust and distrust both significantly affected AI aversion: for trust, \( F(1, 369) = 23.84, p < 0.01 \); for distrust, \( F(1, 369) = 557.94, p < 0.01 \). These findings further validate our proposal of trust and distrust as mediators in our model. Altogether, these ANOVA results provide additional evidence that the effect of task complexity on AI aversion is mediated by trust and distrust and is moderated by task subjectivity.

Supplemental Analysis

Additionally, we conducted a moderated mediation analysis using PROCESS (Hayes, 2017). We used 5000 bootstrap samples and the 95% confidence interval for the analysis. As Table 4 shows, the conditional indirect effect of task complexity is significantly positive when the task is objective, but insignificant when the task is subjective. The conditional indirect effect of task complexity on AI aversion via distrust is significant for both objective and subjective tasks, but with opposite signs. For objective tasks, the effect is negative, whereas for subjective tasks, the effect is positive. We further checked the index of moderated mediation (difference between conditional indirect effects for subjective vs. objective tasks) for both trust and distrust as mediators and find that in both cases, the difference is significant. As Table 4 shows, the indirect effect via trust for the subjective task is 0.25 lower than that for the objective task, whereas the indirect effect via distrust for the subjective task is 1.89 higher than that for the objective task.

Combined with Figure 3, these findings suggest that, for objective tasks, task complexity can significantly increase aversion by reducing trust and significantly decrease aversion by reducing distrust. For subjective tasks, task complexity has no effect on aversion by influencing trust but can significantly increase aversion by enhancing distrust.

Table 2.4: Conditional Indirect Effect of Task Complexity on AI Aversion for Two Task Groups

<table>
<thead>
<tr>
<th>Indirect Effect</th>
<th>Task</th>
<th>Effect</th>
<th>Bootstrap SE</th>
<th>Bootstrap 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC(\rightarrow)TR(\rightarrow)AA</td>
<td>Objective</td>
<td>0.22</td>
<td>0.09</td>
<td>[0.07, 0.42]</td>
</tr>
<tr>
<td>Subjective</td>
<td>-0.03</td>
<td>0.04</td>
<td>[-0.14, 0.04]</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------</td>
<td>------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Subjective – Objective</td>
<td>-0.25</td>
<td>0.09</td>
<td>[-0.47, -0.09]</td>
<td></td>
</tr>
</tbody>
</table>

**Objective**

- \[\text{Objective} = -0.98, 0.18, [-1.33, -0.63]\]
- \[\text{Subjective} = 0.92, 0.23, [0.46, 1.36]\]
- \[\text{Subjective – Objective} = 1.89, 0.29, [1.31, 2.48]\]

**Note:** TC: Task Complexity; DT: Distrust; TR: Trust; AA: AI Aversion. Significant indirect effects are in bold. Number of bootstrap samples: 5000.

**Discussion**

**Implications for Research**

This study examines the effects of task complexity on users' AI aversion, mediated by both trust and distrust and moderated by task subjectivity. We ran an experiment to compare four different AI tools that complete tasks with varying degrees of complexity and subjectivity. This study makes several significant theoretical contributions.

First, despite increasing research interests in users' response to AI, there is a dearth of relevant studies, particularly in the Information Systems (IS) domain, regarding AI aversion. Existing studies on AI aversion have primarily focused on user or technology factors, neglecting the impact of the characteristics of the task performed by AI. However, researchers have long noticed the crucial role of task characteristics in affecting human-computer interaction (Goodhue & Thompson, 1995; Dishaw & Strong, 1999; Fang et al., 2005; Castelo et al., 2019). In this research, we focus on how task complexity and task subjectivity jointly determine users’ AI aversion. Given the current widespread concerns about whether AI will possibly harm human society and even endanger human civilization, it is highly relevant and timely to explore what tasks humans feel comfortable to delegate to AI. The findings could help to prioritize or regulate AI development for specific tasks that invoke least resistance.

Second, current literature on algorithm aversion presents a contradictory findings regarding the effect of task subjectivity on algorithm aversion. Some studies indicate that for tasks where users subjectively (based on personal feelings, emotions etc.) assess the output, they prefer humans over AI (e.g., Xie et al., 2022; Raj et al., 2023). Contrary to this, others (e.g., Sohn et al., 2021) found opposite results. Utilizing users preferences for AI vs human generated pictures for t-shirts, they found that compared to human, users expressed their willingness to pay significantly higher for AI-generated products. Similarly, Logg et al. (2019) find that people prefer algorithms over humans for subjective tasks. They used the task of ‘forecasting the popularity of songs and romantic partners’ (a highly subjective task) to check users’ preferences. But none of these studies consider other task properties that might be of high importance to users. Hence, based on current findings, we can’t make a conclusive statement regarding the direct effect of task subjectivity on AI aversion. Instead, by putting task subjectivity as a moderator and by studying its interactions with other task characteristics (e.g., task complexity) and different psychological constructs (i.e., trust and distrust), we offer a holistic picture and uncover new insights into users’ AI aversion behavior. Our findings strengthen the importance of task subjectivity on AI aversion and provide a more nuanced understanding of the underlying mechanism shaping users opinion regarding AI. But more importantly, by placing task subjectivity as a moderator and empirically validating its role, we contribute to the AI aversion literature and offer a possible solution to the apparent aforementioned contradiction. We show that task subjectivity is crucial in understanding AI aversion, but it completes the picture better while functioning as a moderator rather than a stand-alone independent variable.

Third, recent studies highlight the important roles of trust and distrust in influencing individual behaviors (e.g., Dimoka, 2010; Benamati & Serva, 2007; Dwork & Minow, 2022).
Connelly et al., 2012; Lumineau, 2017). However, no research has considered both trust and distrust together and compared their relative importance on target behaviors related to AI. In this study, we empirically validated that trust and distrust are orthogonal constructs that are not necessarily associated with each other. The low correlation between trust and distrust (r = -0.099, p > 0.05) suggests that the two constructs are independent of each other. Moreover, we find that both constructs play important roles in shaping users’ AI aversion behavior and distrust has a much stronger effect on AI aversion than trust (0.85 vs. -0.18, see Figure 2). This suggests that when people make approach-avoidance decisions in a context characterized by potential risks, their distrust plays a dominant role. By considering both trust and distrust and their relative importance, we reveal users’ cognitive processes underlying their AI aversion decision.

Finally, based on current literature we predicted that task complexity will increase users’ distrust in AI (H3). Instead we find (non-significantly) that task complexity reduces distrust. This surprising effect gets clarified once we consider the moderating effect of task subjectivity. The interaction effect shows that while for subjective tasks complexity increases distrust, it reduces it for objective tasks (H5a). This moderating effect implies that task complexity can be an important source of distrust based on the subjective nature of a task. In other words, people will be more apprehensive towards relying on an artificial intelligence for a task if it is both complex and subjective. It further indicates that if the task is complex, but it has some set measures to cross-validate the output of an AI (i.e., the task is objective), being a complex task will rather reduce their hesitation in accepting an AI to complete that task. This finding adds a significant contribution to our current body of knowledge of how users evaluate distrust on AI by empirically presenting the importance of contextual factors. To the best of our knowledge, the existing literature on AI aversion has not considered how users’ evaluation of distrust (and trust) happens in different contextual conditions. Our study is among the first that considers this gap in the AI aversion literature, theorizes and empirically validates it.

Additionally, while task complexity for objective tasks increases AI aversion by reducing trust, it reduces aversion by decreasing distrust (table 4). Considering these two findings together, our study reveals an interesting paradoxical effect of task complexity on AI aversion for objective tasks. While we do not study the effect of specific objective matrices in forming trust or distrust on an AI, our findings provide valuable insights into the possibility of users’ distrust or trust towards AI being influenced by different objective measures of a complex task. The contrasting effects observed suggest that the influence of task complexity on AI aversion is intricately tied to individuals' level of trust or distrust in AI systems. By examining specific objective evaluations, individuals can either form judgements that either enhance their trust or increase distrust in AI. This is similar to the classic example of a half-glass experiment. While prioritizing objective matrices like the accuracy of the task output can reduce distrust, prioritizing the error might reduce their trust in the AI. Hence, the intricate interplay between task complexity, trust, and distrust that we revealed in our study adds a novel and significant contribution to the body of research by revealing a paradoxical path from task complexity to AI aversion.

**Implications for Practice**

Our findings have important practical implications for AI developers and business organizations. Specifically, the study highlights the significant impact of task characteristics on users' AI aversion behavior. This implies that a one-size-fits-all approach to AI development and marketing may not be feasible. Instead, developers and organizations should carefully consider...
the specific characteristics of each task and tailor their approach accordingly. For example, our findings indicate that a combination of task characteristics, such as task complexity and subjectivity, influences AI aversion. This implies that developers and organizations should take a more nuanced approach to understanding different task factors influencing users' acceptance of AI technology. For instance, they may consider the contextual relevance (i.e., how well AI technology aligns with the specific context in which it is being implemented); user expertise; and a user-centered design. By doing so, they can increase the likelihood of avoiding potential AI aversion.

Next, the findings of this study provide insights into the psychological mechanisms that users undergo to evaluate AI. Specifically, our study shows that users rely on both trust and distrust towards when deciding whether to avert AI and that the type of task determines how these mechanisms work. Considering the growing emphasis on developing AI for highly complex tasks and based on our findings that task complexity increases distrust and reduces trust, we suggest that organizations and developers need to allocate their resources strategically for both building trust and reducing distrust. Specifically, if the AI is intended for a complex subjective task, more resources should be directed towards understanding and addressing users’ distrust to tackle their AI aversion behavior. Alternatively, for complex objective tasks, a high emphasis should be channeled to understanding users’ trust on AI. Because task complexity for objective tasks also reduces distrust, they could alternatively conduct an online survey (Fang et al. 2014) to identify which objective features for a complex task affect trust and which affect distrust. It will further help efficiently prioritizing certain task features over others for marketing and advertisements.

**Limitations and Directions for Future Research**

This study has several limitations and calls for further research into AI aversion. First, our sample is extracted from the US population, and the cultural background of our sample may limit the generalizability of our findings. Future research should consider cultural factors, such as uncertainty avoidance, to understand better how AI aversion behaviors vary in different cultural contexts. A more diverse sample representing different countries and cultures can be employed to explore how cultural factors influence AI aversion.

Second, while our study is focused on two task characteristics (complexity and subjectivity), we acknowledge that other task characteristics may also be relevant to AI aversion. In addition, technological factors related to AI, human factors related to users may influence, and the complex interaction among all these factors could also influence AI aversion. Future research should identify and consider salient factors to yield a more holistic understanding of AI aversion.

Finally, we assessed trust and distrust through users’ self-reporting that may not fully capture their true beliefs. To improve measurement accuracy, future studies can adopt biometric methods that have been applied by neuroIS researchers, such as eye tracking (Jenkins and Jiang 2010), galvanic skin response (Khawaji et al. 2015), electrocentrography (Akash et al. 2018), and functional magnetic resonance imaging (Dimoka 2010). Compared with self-reports, these biometric methods are less influenced by method biases, which can help ascertain whether trust and distrust indeed play a role in influencing users’ AI aversion.

While our study provides valuable insights into the task-related factors that influence AI aversion, further research is needed to fully understand the complex psychological mechanisms that underlie users' attitudes toward AI technology. By addressing the limitations of our study and incorporating new insights from cognate disciplines, future studies can develop more
comprehensive models of AI aversion and create strategies to promote greater acceptance and adoption of AI technology.

**Conclusion**

This study provides valuable insights into why users exhibit aversion towards AI. It investigates the effect of task complexity on users' AI aversion behavior while considering task subjectivity as a contextual factor. Utilizing a randomized controlled experiment, we empirically validate that the impact of task complexity on AI aversion is mediated by both trust and distrust and the indirect effects are contingent on task subjectivity. For objective tasks, task complexity can significantly increase aversion by reducing trust and significantly decrease aversion by reducing distrust. In contrast, for subjective tasks, task complexity has no effect on aversion, but can significantly increase aversion by enhancing distrust. These intricate moderated mediation effects have potential to advance the field of human-AI interaction, thus making a significant contribution to IS research. They also suggest the need to consider task characteristics and their interplay with users' perceptions when designing and implementing AI systems.

**References**


Cui, Y. (2022). Sophia Sophia tell me more, which is the most risk-free plan of all? AI anthropomorphism and risk aversion in financial decision-making. International Journal of Bank Marketing, 40(6), 1133-1158.


Dietvorst, B. J., & Bharti, S. (2020). People reject algorithms in uncertain decision domains because they have diminishing sensitivity to forecasting error. Psychological science, 31(10), 1302-1314.


69


Meyer, D. (2018, October 10). Amazon Reportedly Killed an AI Recruitment System Because It Couldn't Stop the Tool from Discriminating Against Women. Yahoo.com. Retrieved June 1, 2023, from https://finance.yahoo.com/news/amazon-reportedly-killed-ai-recruitment-100042269.html?guccounter=1&guce_referrer=aHR0cHM6Ly9hbmFseXRpY3NpbmRyYW1hZy5jb20v&guce_referrer_sig=AQAAAMlq-nf8kXOM4CYGj5A8rAkr3c3Gw2Il6jscwA6GxyRImJTUpx_tlRCrHPo_cjuiPUxDI7gYuh11RY ulpU_y3yCOR3gzw9u3nMjcmY084aFsHrXq-bMP-1bSV1VttM0e74WKPPSMHZYdJs3wnjRsXIDYrhMm_wxUFFELMFZB7


Sallam, M. (2023, March). ChatGPT utility in healthcare education, research, and practice:
systematic review on the promising perspectives and valid concerns. In Healthcare (Vol. 11, No. 6, p. 887). MDPI.


https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6139009/

### Appendix

Table 2.A1. literature review for AI aversion (organized by year)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Problem addressed</th>
<th>Key Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dietvorst et al.</td>
<td>2015</td>
<td>Reasons for algorithm aversion</td>
<td>• Seeing an algorithm perform affects aversion positively (seeing an AI making mistakes reduces confidence, that in turn increases aversion)</td>
</tr>
<tr>
<td>Prahl &amp; Van Swol</td>
<td>2017</td>
<td>Antecedents of algorithm aversion</td>
<td>• Participants find human to be more common than AIs • bad advice experience decreases utilization of AIs</td>
</tr>
<tr>
<td>Gherhes</td>
<td>2018</td>
<td>The consequences of AI emergence in the future</td>
<td>• Participants negatively associate AI to different outcomes. • The responses varied based on a responder’s specialization and gender (women are more concerned than men; humanities students are more concerned than technology students)</td>
</tr>
<tr>
<td>Castelo et al.</td>
<td>2019</td>
<td>Task dependent (task subjectivity) algorithm aversion</td>
<td>• People rely on algorithms less for subjective tasks • increasing task’s perceived objectivity and perceived affective human-likeness increase trust in algorithms for subjective tasks</td>
</tr>
<tr>
<td>Stein et al.</td>
<td>2019</td>
<td>Antecedents of users’ threat experience (UTE) and the effect of UTE on autonomous technology acceptance</td>
<td>• Perceived threat to immediate physical and human uniqueness are antecedents of UTE • UTE reduces autonomous tech. acceptance</td>
</tr>
<tr>
<td>Longoni et al.</td>
<td>2019</td>
<td>Consumer receptivity of AI in medicine</td>
<td>• Users show aversion to AI in both real and hypothetical healthcare application • A perceived inability to account for consumers’ unique characteristics and circumstances drives consumer resistance to medical AI • The relationship is mediated by uniqueness neglect • AI resistance is eliminated when AI service is framed as customized, or for someone else, or used to only supports a human provider.</td>
</tr>
<tr>
<td>Sohn et al.</td>
<td>2020</td>
<td>Consumers’ evaluations of GAN-generated products (based on GAN disclosure/non-disclosure)</td>
<td>• Functional, social, and epistemic consumption values positively affect willingness to pay in GAN-generated products. • Compared to non-GAN-generated products, willingness to pay is significantly higher for GAN-generated products. • Evaluations of functional value, emotional value, and willingness to pay are highest when GAN technology is used, but not disclosed.</td>
</tr>
<tr>
<td>Niszczota &amp; Kaszás</td>
<td>2020</td>
<td>Robo-investment aversion</td>
<td>• Robo-investment aversion exists in the financial realm (when considering investment in controversial and non-controversial industries)</td>
</tr>
<tr>
<td>Wu et al.</td>
<td>2021</td>
<td>Acceptance of a painful treatment plan from artificial intelligence (vs human)</td>
<td>• Patients accepted a human doctor more than an AI doctor when faced with the same treatment plan. • The nature of the treatment plan and the treatment subject had an interactional effect on the acceptance of the treatment plan. • Experience capacities mediated the relationship between treatment provider (AI vs. human) and treatment plan acceptance.</td>
</tr>
<tr>
<td>Frank et al.</td>
<td>2021</td>
<td>Drivers and social implications of AI adoption in healthcare</td>
<td>• 1 out of 10 individuals choose medical AI over human physicians in a hypothetical triage-phase of COVID-19 pre-hospital entrance. • Key predictors of medical AI adoption are people’s trust in</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>during the COVID-19 pandemic</td>
<td></td>
<td>medical AI and, to a lesser extent, the trait of open-mindedness.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Mistrust and perceived uniqueness neglect from human physicians,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>as well as a lack of social belonging significantly decrease people’s</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>medical AI adoption.</td>
<td></td>
</tr>
<tr>
<td>Castelo &amp; Ward</td>
<td>2021</td>
<td>Effects of conservatism on lay perception of AI and its variations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>based on a moral reframing intervention</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Political conservatism is negatively associated with comfort &amp;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trust in AI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Moral reframing reduces the negative effect</td>
<td></td>
</tr>
<tr>
<td>Kießling et al.</td>
<td>2021</td>
<td>Effect of algorithm aversion on fake news detection</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Fake news flag from AI has a weaker effect because participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>assume the AI to be more erroneous than human experts</td>
<td></td>
</tr>
<tr>
<td>Filiz et al.</td>
<td>2021</td>
<td>Effect of incentives and feedbacks on algorithm aversion</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Initial aversion reduced after receiving feedbacks and incentives</td>
<td></td>
</tr>
<tr>
<td>Xie et al.</td>
<td>2022</td>
<td>Effect of AI (vs. human) recommenders on consumers’ preferences for</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>search versus experience products</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Consumers show less avoidance of algorithms when recommending</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>search products compared to experience products.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Consumers are less likely to purchase experience products</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>recommended by AI, while there are no significant differences</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>between AI versus human recommenders when recommending search</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>products.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Consumers have a higher level of cognitive conflict when AI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>recommends experience products compared to human recommenders.</td>
<td></td>
</tr>
<tr>
<td>Kim et al.</td>
<td>2022</td>
<td>Effect of AI assistance to service employees on service outcomes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Providing AI-generated diagnoses significantly improves service</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>outcomes measured by academic performance.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Some tutors may not utilize AI assistance due to AI aversion.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Factors associated with unforeseen barriers to usage (i.e.,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>technology overload) can moderate its impact on outcomes.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Tutors who contribute significantly to the firm's revenue</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relied heavily on AI assistance but unexpectedly benefited little</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>from AI in improving service outcomes.</td>
<td></td>
</tr>
<tr>
<td>Cui</td>
<td>2022</td>
<td>Effect of anthropomorphizing an AI chatbot on risk preferences</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Anthropomorphizing AI chatbot strongly (positively) influence</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>risk aversion in financial decision making</td>
<td></td>
</tr>
<tr>
<td>Filiz et al.</td>
<td>2022</td>
<td>Effect of algorithm aversion on robo-advisors’ establishment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Algorithm aversion negatively affect robo advisors’ establishment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>even with high level of accuracy</td>
<td></td>
</tr>
<tr>
<td>Zhou et al.</td>
<td>2022</td>
<td>Impact of intergroup threat on AI acceptance in healthcare</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Perceived threat affects AI avoidance positively</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Individual level threats influence patients to rely on human</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>doctors more</td>
<td></td>
</tr>
<tr>
<td>Longoni &amp; Cian</td>
<td>2022</td>
<td>Effect of ‘word of machine’ (WofM) on AI recommendations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• People in general prefer AI when decision making is based on</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>utilitarian attributes and vice versa</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• WofM effect reverses when the recommendation needs to match a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>person’s unique preferences</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ‘Consider the opposite’ protocol attenuates the effect of WofM.</td>
<td></td>
</tr>
<tr>
<td>Lanz et al.</td>
<td>2023</td>
<td>Employees’ adherence to unethical instruction from AI (vs human)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Employees adhere less to unethical instructions from an AI than a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>human supervisor.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Individual characteristics such as the tendency to comply without</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dissent or age constitute important boundary</td>
<td></td>
</tr>
</tbody>
</table>
conditions.
• The perceived mind of the supervisors serves as an explanatory mechanism.

Lopez & Garza 2023 Consumers acceptance of AI’s that evaluate them • The lack of transparency anxiety impacts perceived fairness • Perceived lower fairness is evaluated negatively

Raj et al. 2023 Effect of AI disclosure on creative content evaluations • No effect on short stories, but negative evaluations of emotional poems • perceiving the content as distinctly human can result in the negative evaluation

Table 2.A2. Demographics of respondents

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Categorization</th>
<th>Number of respondents</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1. Male</td>
<td>227</td>
<td>60.5</td>
</tr>
<tr>
<td></td>
<td>2. Female</td>
<td>148</td>
<td>39.5</td>
</tr>
<tr>
<td>Age</td>
<td>1. &lt; 21</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2. 21 - 30</td>
<td>111</td>
<td>37.3</td>
</tr>
<tr>
<td></td>
<td>3. 31 - 40</td>
<td>118</td>
<td>35.5</td>
</tr>
<tr>
<td></td>
<td>4. 41 - 50</td>
<td>44</td>
<td>15.5</td>
</tr>
<tr>
<td></td>
<td>5. &gt; 50</td>
<td>42</td>
<td>11.7</td>
</tr>
<tr>
<td>Education</td>
<td>1. High school</td>
<td>44</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>2. College</td>
<td>100</td>
<td>26.7</td>
</tr>
<tr>
<td></td>
<td>3. Undergraduate</td>
<td>127</td>
<td>33.9</td>
</tr>
<tr>
<td></td>
<td>4. MS</td>
<td>81</td>
<td>21.6</td>
</tr>
<tr>
<td></td>
<td>5. PhD</td>
<td>22</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>6. Others</td>
<td>1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 2.A3. Goodness-of-fit indices for the structural model

<table>
<thead>
<tr>
<th>Fit Indices</th>
<th>Recommended Level of fit</th>
<th>Model Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute fit measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-squared</td>
<td>185.71</td>
<td></td>
</tr>
<tr>
<td>Df</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Probability level</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Chi-squared/df</td>
<td>&lt; 3</td>
<td>2.61</td>
</tr>
<tr>
<td>GFI (Goodness-of-fit Index)</td>
<td>&gt; 0.9</td>
<td>0.94</td>
</tr>
<tr>
<td>RMSE (Root mean square error of approximation)</td>
<td>&lt; 0.1</td>
<td>0.06</td>
</tr>
<tr>
<td>Incremental Fit Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGFI (Adjusted goodness-of-fit Index)</td>
<td>&gt; 0.80</td>
<td>0.89</td>
</tr>
<tr>
<td>NFI (Normed Fit Index)</td>
<td>&gt; 0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>CFI (Comparative Fit Index)</td>
<td>&gt; 0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>IFI (Incremental Fit Index)</td>
<td>&gt; 0.92</td>
<td>0.96</td>
</tr>
<tr>
<td>RFI (relative Fit Index)</td>
<td>&gt; 0.93</td>
<td>0.90</td>
</tr>
</tbody>
</table>
Parsimony Fit Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Threshold</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCFI (Parsimony Comparative of Fit Index)</td>
<td>&gt; 0.5</td>
<td>0.65</td>
</tr>
<tr>
<td>PNFI (Parsimony Normed Fit Index)</td>
<td>&gt; 0.5</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Table 2.A4. Measures

<table>
<thead>
<tr>
<th>Construct</th>
<th>Dimension</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust</td>
<td></td>
<td>(AI) has the ability for (task)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(AI) will be (task) according to my</td>
</tr>
<tr>
<td></td>
<td></td>
<td>expectations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(AI) is credible when it is used for</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(task)</td>
</tr>
<tr>
<td></td>
<td>Distrust</td>
<td>Dimoka (2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I am skeptical that (AI) is competent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>in (task)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I am worried that (AI) would not be</td>
</tr>
<tr>
<td></td>
<td></td>
<td>trustworthy in (task)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I believe (AI) will perform (task) in a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fraudulent way</td>
</tr>
<tr>
<td></td>
<td>AI Vehension (AA)</td>
<td>Chi et al. (2020)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I don’t want to use (AI) for (task)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I would not select (AI) for (task)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I would ignore the output of (AI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bhattacherjee et al. (2007)</td>
</tr>
</tbody>
</table>

All constructs are measured using a 7 points Likert scale (1=strongly disagree, 4=neutral, 7=strongly agree); AI: the AI tool a participant used for the experiment; task: the particular task a participant used that AI tool for during the experiment.

Table 2.A5: Literature Review for Algorithm Aversion

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Construct</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Inability to judge subjective issues</td>
<td>Niszczota and Kaszás (2020), Castelo et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Perceived lack of mind</td>
<td>Bigman and Gray (2018)</td>
</tr>
<tr>
<td></td>
<td>Perceived effectiveness</td>
<td>Castelo et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Difficulty to understand</td>
<td>Yeomans et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Slow speed</td>
<td>Effendic et al. (2020)</td>
</tr>
<tr>
<td></td>
<td>Inability to explain</td>
<td>Yeomans et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Perceived capabilities</td>
<td>Jussupow et al. (2020)</td>
</tr>
<tr>
<td></td>
<td>Higher accuracy</td>
<td>Pezzo and Beckstead (2020)</td>
</tr>
<tr>
<td></td>
<td>Inability to learn</td>
<td>Berger et al. (2020).</td>
</tr>
<tr>
<td></td>
<td>Algorithm agency</td>
<td>Jago (2019)</td>
</tr>
<tr>
<td></td>
<td>Ability to modify (-)</td>
<td>Dietvorst et al (2018)</td>
</tr>
<tr>
<td></td>
<td>Lack of incentivization (Solution: Context-specific behavioral design)</td>
<td>Burton et al. (2019)</td>
</tr>
</tbody>
</table>

78
<table>
<thead>
<tr>
<th>User</th>
<th>Factor</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uniqueness neglect</td>
<td>Longoni et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Perceived financial return</td>
<td>German and Merkle (2020)</td>
</tr>
<tr>
<td></td>
<td>Commonality (with humans)</td>
<td>Prahl and Swol (2017)</td>
</tr>
<tr>
<td></td>
<td>Objectivity of decision</td>
<td>Logg (2017)</td>
</tr>
<tr>
<td></td>
<td>Seeing them perform</td>
<td>Dietvorst et al (2015)</td>
</tr>
<tr>
<td></td>
<td>High outcome ambiguity</td>
<td>Fuchs et al. (2016)</td>
</tr>
<tr>
<td></td>
<td>User's lack of decision control</td>
<td>Burton et al. (2020)</td>
</tr>
<tr>
<td></td>
<td>Combatting intuition</td>
<td>Burton et al. (2020)</td>
</tr>
<tr>
<td></td>
<td>Discomfort</td>
<td>Castelo et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>pre-(neg) bias</td>
<td>Liu et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Excessive Reliance on personal knowledge (domain expertise)</td>
<td>Kim et al. (2016), Niszczota and Kaszás (2020), Logg et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Appreciation of self-opinion (overconfidence)</td>
<td>Logg (2017), Logg et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Unfamiliarity (-)</td>
<td>Logg et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>Perceived task difficulty</td>
<td>Effendic et al. (2020)</td>
</tr>
<tr>
<td></td>
<td>Task subjectivity</td>
<td>Castelo et al. (2019)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TASK</th>
<th>Factor</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perceived task difficulty</td>
<td>Effendic et al. (2020)</td>
</tr>
<tr>
<td></td>
<td>Task subjectivity</td>
<td>Castelo et al. (2019)</td>
</tr>
</tbody>
</table>

Note: A negative sign indicates that extant studies showed this factor to reduce algorithm aversion or to increase AI appreciation