Exploring the Role of Trust during Human-AI Collaboration in Managerial Decision-Making Processes

by

Serdar Tuncer

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Arts

in

Human-Computer Interaction

Carleton University

Ottawa, Ontario

© 2022

Serdar Tuncer
Abstract

Despite the growing popularity of using Artificial Intelligence-based (AI-based) models to assist human decision-makers, little is known about how managers in business environments approach AI-assisted decision-making. Thus, our research is guided by two questions: (1) What facets make the Human (Manager)-AI decision-making process trustworthy, and (2) Does trust in AI depend on the degree to which the AI agent is humanized? Our results show that (a) AI is preferred for operational versus strategic decisions and decisions that indirectly affect individuals, (b) the ability to interpret the decision-making process of AI agents would help improve user trust and alleviate calibration bias, (c) humanoid interaction styles were believed to improve the interpretability of the decision-making process, and (d) organizational change management was essential for adopting AI technologies. Our survey analysis indicates that when interpretability and model confidence are present in the decision-making process involving an AI agent, higher trustworthiness scores are observed.

Keywords: Human-AI interaction, Trustworthy AI, Managerial Decision-Making, AI-Infused Decision-Making, Model Interpretability, Model Confidence
Acknowledgements

I would like to express my special appreciation and thanks to my thesis supervisor, Professor Alex Ramirez; you have been a tremendous mentor to me. I would like to thank you for encouraging my research and for allowing me to grow as a researcher. I also want to thank all the professors from whom I had the pleasure of taking courses during the program. I learned so much from all of you and will continue to heed your advice in my future projects. Additionally, a special thank you to Gözde Uyumaz, who helped me to conduct various statistical analyses.

During this research, I was fortunate enough to receive assistance from several executives as well as business organizations, including CADSI, Invest Ottawa and the Canadian Society of Association Executives.

Finally, to my loving family, Pinar and Ada. It would not be possible to achieve any of this without your love and support.
# Table of Contents

1  Introduction .................................................................................................................. 1
   1.1  Motivation .................................................................................................................. 2
   1.2  Vision and Philosophy .............................................................................................. 3
   1.3  Research Questions ................................................................................................... 5
   1.4  Contribution .............................................................................................................. 6

2  Background ..................................................................................................................... 8
   2.1  Definition of AI .......................................................................................................... 9
   2.2  AI, Machine Learning and Deep Learning ............................................................... 10
   2.3  Narrow and Strong AI ............................................................................................. 11
   2.4  Sectors and Research Areas in AI ........................................................................... 12

3  Decision-Making in Organizations and AI ................................................................. 14
   3.1  Pioneers of Managerial Decision-Making Theories ............................................... 14
   3.2  Human-AI Collaboration in the Workplace .............................................................. 17
   3.3  Challenges with AI-Infused decision-making ......................................................... 20

4  The missing link in Human-AI Collaboration for Decision-Making in Organizations: Trust 21
   4.1  What is Trust? ............................................................................................................ 22
   4.2  Trust in Non-Human Decision-Makers .................................................................... 22
   4.3  Building Trust in the Workplace .............................................................................. 25

5  Methodology .................................................................................................................... 27
   5.1  Pilot and Quantitative Study .................................................................................... 28
   5.1.1  Ethics and Recruitment ......................................................................................... 28
   5.1.2  Consent and Compensation ............................................................................... 29
   5.1.3  Data Collection .................................................................................................... 29
   5.1.4  Measures .............................................................................................................. 32
   5.2  Qualitative Study ..................................................................................................... 32
   5.2.1  Ethics and Recruitment ......................................................................................... 33
   5.2.2  Consent and Compensation ............................................................................... 33
   5.2.3  Data Collection .................................................................................................... 34
   5.2.4  Data Analysis ...................................................................................................... 34

6  Quantitative Analysis and Findings ............................................................................. 35
   6.1  Demographics .......................................................................................................... 35
   6.2  Reliability ................................................................................................................. 36
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3</td>
<td>Confidence and Interpretability Effect</td>
<td>37</td>
</tr>
<tr>
<td>6.4</td>
<td>AI Interface/Dashboard Interface Effect</td>
<td>39</td>
</tr>
<tr>
<td>6.5</td>
<td>Authority Effect</td>
<td>40</td>
</tr>
<tr>
<td>6.6</td>
<td>Knowledge Effect</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>Qualitative Analysis and Findings</td>
<td>44</td>
</tr>
<tr>
<td>7.1</td>
<td>Operational Versus Strategic Decisions</td>
<td>44</td>
</tr>
<tr>
<td>7.2</td>
<td>AI Transformation and Organizational Change Management</td>
<td>45</td>
</tr>
<tr>
<td>7.3</td>
<td>Trust Versus Bias in AI-Infused Decision-Making</td>
<td>46</td>
</tr>
<tr>
<td>7.4</td>
<td>AI Design and Interaction Styles</td>
<td>47</td>
</tr>
<tr>
<td>8</td>
<td>Discussion &amp; Conclusion</td>
<td>48</td>
</tr>
<tr>
<td>8.1</td>
<td>Future Work</td>
<td>49</td>
</tr>
<tr>
<td>8.2</td>
<td>Limitations</td>
<td>49</td>
</tr>
<tr>
<td>9</td>
<td>Appendix</td>
<td>51</td>
</tr>
<tr>
<td>10</td>
<td>References</td>
<td>65</td>
</tr>
</tbody>
</table>
List of Tables

Table 1: Gender distribution of participants ................................................................. 35
Table 2: Seniority participation of participants .............................................................. 36
Table 3: Level of AI knowledge .................................................................................... 36
Table 4: Confidence and Interpretability Effect ............................................................ 37
Table 5: AI Interface/Dashboard Interface Effect on Subdimensions ............................ 39
Table 6: Authority Effect (Final/Some) ....................................................................... 40
Table 7: Knowledge Effect .......................................................................................... 42
List of Figures

Figure 1: AI Ecosystem ................................................................. 5
Figure 2: Timeline of AI Development ........................................ 9
Figure 3: Evolution of the use of AI ................................................. 12
Figure 4: The evolution of the Human-Machine-AI Relationship .......... 17
Figure 5: Organizational Decision-Making Structures Involving AI-Based Algorithms .......... 18
Figure 6: Calibrating Trust ........................................................... 24
Figure 7: Study Flow ................................................................. 28
Figure 8: SAM ........................................................................... 32
List of Appendices

Appendix A – CUREB-B Ethics Clearance ................................................................. 51
Appendix B – Online Questionnaire ........................................................................ 53
Appendix C – Oral Consent Script ................................................................. 59
Appendix D – Survey Consent Form ............................................................ 61
Appendix E – Interview Script ........................................................................ 63
1 Introduction

Technology has played an essential part in human society. Until now, technology has been a tool mastered and controlled by humans. However, now there may be a new tool that can think and challenge its human master. Artificial Intelligence (AI) holds great promise in helping solve human problems independent of its creator; “it can understand humans and communicate with them, and it can even challenge humans in their unique characteristic: natural intelligence” (Abbass, 2019, p. 160).

As complex social processes become increasingly automated, algorithmic decision-making is being proposed as a possible solution in various social contexts. As such, AI-infused decision-making is becoming increasingly pervasive. It is taking on more effective roles within society, from assisting in medical care to determining parole and sentencing in the courtroom, from providing financial suggestions to identifying suitable job candidates, running autonomous vehicles, robots and much more. Despite the growing enthusiasm to incorporate AI into the business sphere, various issues and questions regarding the role of AI still need to be addressed. When viewed through the lens of Human-Computer Interaction (HCI), one such critical need is to understand the decision-making process for AI technologies. Notably, it is vital to understand the role of trust in the context of Human-AI teaming and collaboration (Łapińska et al., 2021; Shneiderman, 2020a; Xu, 2019; Xu et al., 2022).

Some systems using AI to aid decision-making have led to 'bad' or even harmful decisions (M. K. Lee, 2018; Shneiderman, 2020b) due to a lack of understanding on the user's part. The challenge is to effectively utilize AI's power while empowering the people using it. Prior studies regarding AI-infused decision-making and the relationship with trust often sought to evaluate the decision's trustworthiness rather than the decision-making process's trustworthiness (Ashoori & Weisz, 2019). Therefore, it is critical that the people who use the technology feel in control by understanding how it affects them (Liao et al., 2020; Xu et al., 2021) and, thus, trusting in the technology as well as in the decision-making process. Therefore, the primary focus of this research is to explore the role of trust during Human-AI collaboration in managerial decision-making processes.
1.1 Motivation

In recent years, many successful applications of AI have been built thanks to the convergence of improved algorithms, vast computing power, and ever-increasing amounts of data. Notably, as the applications of AI increase, the presence of AI solutions in the business environment is becoming indispensable. AI may have the potential to offer a competitive advantage to businesses through cost savings (such as reducing employment and intelligent production quality control), increased effectiveness of business processes (such as improving accounting and facilitating employee recruitment) and eliminating human error (Łapińska et al., 2021).

Trust is one of the most critical factors in the interaction between humans and between humans and intelligent agents such as robots and algorithms (Floridi, 2019). Additionally, it is a key factor in people’s acceptance and use of new technologies. Conversely, lack of trust often leads to users' aversion to the technology (Miltgen et al., 2013). As such, trust is an important factor shaping an organization’s adoption of AI. For these Human-AI collaborations in decision-making to be effective in the business environment, managers need to know when to trust or distrust an AI’s predictions. To do so, decision-makers must understand how the AI reaches its predictions, in other words, the AI’s “thought process.” The need to improve the transparency of AI-based systems to aid businesses and managers has been shown in several research studies (Glikson & Woolley, 2020). Likewise, with the growing use of AI in the business realm, studying the potential negative impacts of AI in this sphere due to such challenges may effectively drive change in the AI industry (Rakova et al., 2021), thus helping to maximize the benefit from AI while also minimizing its risks.

A study conducted in 2018 (Review MSM, 2018) indicates that 94% of business executives believe AI adoption can help solve strategic challenges. On the other hand, another study conducted the same year (Intelligent Economies, 2018) found that only 18% of organizations are true AI “pioneers,” with extensive AI adoption into their business processes. Moreover, 70% of C-level leaders who work with AI could not explain how a specific AI model made decisions or predictions, and only 35% said their organization made an effort to use AI in a transparent and accountable way (Responsible AI Report, 2021). More work needs to be conducted in this particular area. Some scholars (Shneiderman, 2020b; Xu et al., 2021) suggest that this discrepancy indicates a usability problem related to the lack of trust, revealing the importance of usability for...
decision-makers. It is also worth noting that enhancing the user's trust does not always result in the most satisfactory human-AI collaboration outcomes.

Empirical research by both industry and academics has provided insights into explainable, ethical, responsible AI, but each has its limitations. For example, AI applications are generally designed so that those in the industry (programmers, data scientists etc.) can understand. Likewise, there is a strong indication that those building AI systems focus on themselves rather than following a human-centred approach in order to create systems for end users (Xu et al., 2022). Given HCI’s core interest in technology that empowers people, this gap is necessary to discuss the nature and dynamics of trust in the presence of Human-AI interactions.

Furthermore, research on the real-life use of AI, such as organizations using it in their management or decision-making processes, is becoming an ever more critical (Glikson & Woolley, 2020; Winfield & Jirotka, 2018). One such study (Rakova et al., 2021), involving business professionals who use AI applications in their work daily, showed a gap between the priorities of academic research and the needs of users. Additionally, because AI behaviour is not deterministic, researchers must investigate how it evolves due to human-AI interactions (Rahwan et al., 2019). Therefore, a critical area for further research is studying how Human-AI interaction can facilitate human-centred, ethical, and safe AI integration in organizations.

Our literature review shows that further research and exploration into the acceptance and trust of non-human decision-makers is critically needed to understand how Human-AI interaction can help facilitate a human-centred, ethical, and safe integration of AI within organizations. As such, the human decision-makers within organizations –namely managers– should be the target users in future research, with one of its goals being the design of better interfaces and organizational processes in Human-AI collaboration.

1.2 Vision and Philosophy

Human-AI collaborative approaches can be broadly grouped into three categories: 1. The technology-centric perspective: Humans are biologically constrained in their information processing and display many types of cognitive bias, while computers provide virtually endless opportunities, 2. The human-centric perspective: AI will not develop certain essential qualities
attributed to humans, such as moral reasoning or empathy. For this reason, AI may cause danger to the decision-making process and well-being, and 3. The collective intelligence perspective: According to this approach, true intelligence emerges when multiple entities collaborate over longer periods of time (Peeters et al., 2021). Several research studies suggest that humans and AI systems can be more effective than either one alone (Abbass, 2019; Bansal et al., 2019; Campbell, 2016; Ezer et al., 2019; Jarrahi, 2018). Some scholars seek to answer Human-AI collaboration from an HCI perspective. As Xu et al. (2021, p. 3) noted:

“Like the HCI field that was brought about by PC technology 40 years ago, history has brought HCI professionals to a juncture. This time the consequence of ignoring the “human-centered design philosophy” is obviously more severe.”

Moreover, as trust influences all aspects of a user's relationship with technology, measuring it requires a multifaceted approach. Previous research on trust has focused mainly on either end-user or employee interactions, with little attention dedicated to evaluating trust in C-Level decision-making interactions. Therefore, this research project blends the fields of business and human-computer interaction by considering AI applications’ design from both a social and a technological angle. We draw on two major theoretical directions as a foundation to critically examine design factors that affect the decision-makers' trust AI applications. The first is the Bounded Rationality Theory (Simon, 1996), which underlines that people make decisions that are rational, but within the limits of the information available to us and our mental capabilities. Therefore, Simon’s theory offers to extend capabilities for complex organizational decision-making by leveraging AI's increased computational information processing capacity and analytical approach. Herbert Simon is one of the founders of artificial intelligence and has directly influenced research on intelligent agents, decision-making and knowledge management (Pomerol & Adam, 2005).

Secondly, we draw upon Human-Centered User-Centred design, which is an iterative design method, often used in user experience (UX) research and is rooted in the fields of software engineering and human-computer interaction. Human-Centred Artificial Intelligence (HCAI) is an interdisciplinary approach with a “Human-Centered Design” to AI technologies to help them meet user needs while also developing human models for AI to monitor human status (Xu et al., 2021). HCAI research may help understand what AI can and cannot do (Inkpen et al., 2019). Likewise, (Shneiderman, 2020b) underlines that HCAI not only aims at “measuring human performance” (p.
113), but also “advances the goals of AI, while ensuring human control” (p.114), focusing on enhancing human performance and making systems reliable, safe, and trustworthy.

And finally, we believe that decision-makers in the business environment are “as the central player in the Decision-Making with AI-system in the loop cycle” (Ferreira & Monteiro, 2021, p. 4). We assume that the decision-maker is the main collaborator for those who are designing and developing AI systems as they provide the necessary feedback to improve this collaborative ecosystem (see Figure 1). When we turn our focus on the business environment, our decision-makers are managers who are responsible for short- and long-term strategic business decisions.

![Figure 1: AI Ecosystem](image)

### 1.3 Research Questions

A mixed-method study was chosen as the design for this research to enable an in-depth exploration aimed at identifying the role of trust during Human-AI collaboration in managerial decision-making processes. From this point of view, we examined the following three justification factors: (1) Model Interpretability, (2) Model Confidence, and (3) Humanoid Interface. Based on the conceptual framework and methodology outlined above, the following research questions were formulated:

- **RQ1** What facets make the Human (Manager)-AI decision-making process trustworthy? (Turner et al., 2020).
- **RQ2** Does trust in AI depend on the degree to which the AI agent is humanized? (Shneiderman, 2022).

Additionally, to delve further, we used the following hypotheses:
H1: Interpretability of the AI agent’s decision-making process has no significant effect on the user’s overall trust in the process (Rudin, 2019).
H2: Presence of decision-making process confidence score by AI agent has no significant effect on the user’s overall trust in the process (Turner et al., 2020)
H3: A humanoid interface for an AI agent has no significant effect on the user’s overall trust in the process. (Robert, 2017)

To answer these questions, we used a mixed research methodology mobilizing the mixed approach’s fundamental proposition, together with quantitative and qualitative approaches, allowing us to better understand research problems than each approach individually (Ivankova et al., 2006).

1.4 Contribution

We contribute mainly to areas where, in our literature review, we have identified a research gap.

(1) Our comprehensive quantitative analysis will focus on several factors to understand their impact on four aspects of trust and statistically demonstrate how interpretable models that divulge information about their development and design affect the decision-makers' trust in them.

(2) We will also test how anthropomorphism, or humanizing, of AI agents affects the decision-making process in the business environment.

(3) Finally, our study will add insights to AI-infused decision-making in an area that has been largely overlooked –managerial decision-makers in organizations. Thus, it is expected to help with the development of trustworthy AI systems through input, not only from the data scientists and engineers who construct these systems but from people who will use or be impacted by them in real-world situations.

This thesis is organized as follows: Section 2 presents related work and background information, Section 3 explains the decision-making process in organizations, Section 4 outlines the missing link in human-AI collaboration for decision-making in organizations, Section 5, 6 and 7 present
and analyze mixed methodology results, Section 8 discusses these results and suggests possible future work.
2 Background

Alan Turing, who laid the foundation of the computer in 1937, claimed that machines would act as intelligently as human beings one day. “The Turing Test” is a way of evaluating whether machines can think. According to Turing, if a person cannot distinguish whether the answers to questions come from a human or a network system, AI has been achieved. Turing called this process “the imitation game” (Turing, 1950). Then, in 1955, McCarthy et al. introduced the new term “Artificial Intelligence” in a Dartmouth summer research project proposal to study intelligence applied by machines (McCarthy et al., 2006). The goal was to achieve the simulation of intelligence by machines by describing that intelligence so precisely that it could be simulated. AI has been around for decades, with the appearance of AI algorithms in the late 1950s (see Figure 2). Since then, AI has experienced its ups (“AI springs”) and downs (“AI winters”) (Duan et al., 2019). Likewise, AI research and development experienced two significant setbacks around the winters of 1980 and 1993. However, much progress in AI springs from strategy games such as Go and Chess, which were triumphant moments for AI over its 60 years (Jarrahi, 2018; Pan, 2016). Because there were no commercial applications, machine-learning systems were limited until the 1990s. However, especially after the first “AI winter,” thanks to expert systems (a type of computer system that emulates the decision-making ability of a human expert), interest in AI applications increased once again.
2.1 Definition of AI

Currently, there are many variations of AI, but the concept can be defined from the business lens broadly as “AI aims to design algorithms to provide computers with cognitive skills and competencies for sense-making and decision-making” (Abbass, 2019, p. 168). Indeed, definitions of AI in the literature mainly stem from the complexity of expressing the concept of “intelligence” and “artificial” in human languages (Abbass, 2019). From the perspective of management research, AI can be defined as a “new generation of technologies capable of interacting with the environment by (a) gathering information from outside (including from natural language) or other computer systems; (b) interpreting this information, recognizing patterns, inducing rules, or predicting events; (c) generating results, answering questions, or giving instructions to other systems; and (d) evaluating the results of their actions and improving their decision systems to achieve specific objectives” (Ferrás-Hernández, 2018, p. 260). It should be noted that the definition of AI has evolved (Brachman, 2006), and recent definitions focus on mimicking tasks that humans perform (Bolander, 2019, p. 851; Huang et al., 2019, p. 44) and “simulate human intelligent behaviors” (C. Zhang & Lu, 2021, p. 2).

Figure 2: Timeline of AI Development (Reprinted from: https://www.actuaries.digital/2018/09/05/history-of-ai-winters/)
2.2 AI, Machine Learning and Deep Learning

AI is an umbrella term that arises from Machine Learning (ML) and Deep Learning (DL). Without ML and DL, AI can be defined as “Traditional AI,” based on collections of rules. The number of rules poses significant limitations on Traditional AI, such as image recognition which does not work well as a rule-based system.

In the late 1950s, Arthur Samuel introduced the term machine learning (ML), which relies primarily on data rather than rules to learn and optimize tasks. This data often uses algorithms, such as decision trees (a single tree), random forests (multiple trees), kNN (k Nearest Neighbor), linear regression, Naïve Bayes, and SVM (Support Vector Machines) (Campesato, 2020). Since data (instead of rules) is so important in ML, it is typically one of the following types:

- Supervised learning (labelled data)
- Semi-supervised learning (partially labelled data)
- Unsupervised learning: (clustering)
- Reinforcement learning: (trial, feedback, and improvement)

ML provides insight into mimicking how humans learn. Over the past decade, ML has helped to enable “self-driving cars, practical speech recognition, effective web search, and an understanding of the human genome” (Bakshi & Bakshi, 2018, p. 16).

One important subfield of ML is Deep Learning (DL). While ML involves multilayer perceptrons (MLPs), DL introduces deep neural networks with new algorithms and new architectures (e.g., convolutional neural networks) (Campesato, 2020). In other words, DL uses a "neural network" – based on the human brain – to essentially teach itself by discovering data features. This network of interconnected artificial "neurons" communicates to adjust their inputs and outputs. On the other hand, classical ML, or "non-deep" ML, requires human intervention in the "learning" process to structure and understand the differences in data inputs.
Unstructured data contains vast amounts of valuable information that must be organized and analyzed in order to be used to guide decisions. Particularly in the business realm, this application of AI might be the most valuable.

2.3 Narrow and Strong AI

Despite the "hype" around AI, today's AI is often referred to as "narrow" or "weak" because of its limitation to well-structured (modular or "decomposable") decision objectives (Shrestha et al., 2019). AI is a valuable tool, but at the end of the day, it is just as wise as it has been programmed to be. Floridi (2019) warned that inflated expectations about AI might lead to distraction. Another obstacle for AI is Moravec's Paradox. For example, artificial intelligence can achieve highly advanced mathematics, but simple skills like perception, speech, and movement that humans learn as babies remain complicated for AI to replicate. As Moravec (Moravec, 1988, p. 15) put it, "It is comparatively easy to make computers exhibit adult level performance […] and difficult or impossible to give them the skills of a one-year-old." In other words, for AI, the complex is easy, and the easy is complex.

As can be seen in Figure 3, AI is still in its infancy. Natural language processing, computer vision, pattern recognition optimization, and related AI projects are primarily concerned with satisfactory performance. Thus, setting realistic expectations may be a valid warning for the business community.
Today, there are different terms to describe possible future AI, such as “broad,” “strong,” or Artificial General Intelligence (AGI), which indicates that a machine can successfully perform any mental task a human can do. Although both narrow and wide AI may match or surpass human performance, narrow AI is focused on a single domain and cannot learn to extend into other areas, whereas general AI can. (Wang & Goertzel, 2007). Some scholars have even declared AI the greatest danger to humanity (Bramer, 2015), while others mark this as the start of the posthuman era (Wolfe, 2010) or the rise of singularity (Kurzweil, 2005).

2.4 Sectors and Research Areas in AI

Various applications and techniques fall under the broad umbrella of AI, ranging from decision support systems and neural networks to pattern recognition and genetic algorithms to deep learning and computing vision (C. Zhang & Lu, 2021). In terms of labour, machines are increasingly a substitute for human skills and intelligence in medicine, psychological counselling, human resource management, banking, transportation, and legal counselling (Shrestha et al., 2019).
The total amount of data that humans create and consume is forecasted to snowball. On average, as of 2020, every human on the planet creates at least 1.7 MB of data per second. By 2020, the data is estimated to have reached 64.2 zettabytes and is projected to grow to over 180 zettabytes by 2025 (Data Volume Worldwide, 2021). Similarly, by 2030, the AI market –valued at $27.23 billion in 2019– is projected to reach $16 trillion GDP (Fortune Business Inside, 2020; PwC, 2020).

On the other hand, the adoption of AI by businesses remains low. This slow business adoption may be explained by reluctant management, a lack of skilled workers, and a lack of trust and confidence in AI (Thomas, 2020). Nevertheless, by 2022, the World Economic Forum Report (Jobs of Tomorrow, 2021) estimates a global job creation of 133 million, 16% of which will be in the data and AI sector. There does appear to be a consensus that AI will transform the business from the top down (Daugherty & Wilson, 2018).
3 Decision-Making in Organizations and AI

Efficient organizational decision-making structure designs are a crucial challenge for managers and organization scholars. Thus, AI-augmented human decisions may offer an opportunity to design organizational structures that maximize humans’ and machines’ advantages while minimizing risks (Shrestha et al., 2019). To achieve AI-augmented human decisions, theories that attempt to describe human decision-making have been developed. Several scholars have studied how society can benefit from AI at the maximum level.

AI-based decision-making processes generally can be grouped into two: (1) high-stakes domains such as prison sentence recommendations; (2) lower-stakes domains such as personalized shopping and music recommendation (Ashoori & Weisz, 2019). Regarding the strategic organizational decision-making process, there are two types of decisions in the literature: (1) those made under risk versus; (2) those made under uncertainty. In today's ever-changing business environment, uncertainty is more prevalent, so insight and experience predominate in decisions (Mintzberg, 2013). Regarding responsibility, many researchers suggest that, especially in high-stakes scenarios, the decision-maker is responsible for the final decision. Therefore, the final decision-maker should be a human (Metcalf et al., 2019; Trunk et al., 2020).

3.1 Pioneers of Managerial Decision-Making Theories

James G. March is among the prominent scholars among decision-making theorists. According to March, organizations are groups of individuals, and through the bargaining process, they reach their goals. Decisions are dependent on the available information and the expectations of the individuals (March, 1994). Three variables are essential in this process: Goals, Expectations, and Choice.

(1) Goals: As organizations are made up of individuals, often there are multiple and sometimes conflicting goals of those individuals or coalitions of individuals. Most organizations comprised of many people have multiple goals, which raises the possibility of conflict.
(2) Expectations: This depends on information; how information is presented, gathered, and conveyed affects expectations.

(3) Choice: Everyday choices are a response to reducing uncertainty concerning the organization's problems.

Another prominent thinker in decision-making is Herbert Simon. Simon was the first to propose the notion of an intelligent problem-solving device. He is one of the founders of AI and has indirectly influenced research on intelligent agents in decision-making and knowledge management. According to Simon, a decision “is a matter of compromise,” and the decision-making process is as follows: (1) Identify all the possible alternatives; (2) Determine all the possible consequences of these alternatives; (3) Evaluate all the possible consequences (Simon, 1996).

Evidence shows that human decisions are not consistent, that they would change depending on the situation, the mood of the decision-maker and other factors. Additionally, in the business realm, both businesses and consumers do not possess the knowledge or the computational capability required to execute the rational expectations strategy (Simon, 1996). Sometimes called intuition, this is an unconscious activity that cannot be controlled (Kahneman, 2003). However, Simon states that intuition relies on information and experience that the decision-maker utilizes – in this case, unconsciously – to determine alternatives and probabilities (Trunk et al., 2020). Simon’s Bounded Rationality theory focuses on the departure of real-world behaviour from “rationality as consistency”. He formulates Bounded Rationality as follows:

“The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world or even for a reasonable approximation to such objective rationality” (Simon, 1996, p. 28).

In Simon’s model, the decision-maker does not require complete knowledge of events. Instead, the goal is satisfactory decisions. Thus, he proposes the concept of “satisfice” to help understand the rise of AI. According to Simon, AI can be used most often in the form of heuristic search
Representations, heuristic searches, “what-if” analysis, scenario development and outcome evaluation are all part of the decision-making process (Pomerol & Adam, 2005). Within organizations, managers are the decision-makers that deal with such situations. According to Mintzberg (2013, p. 6), managers become “the nerve center of the unit—its best-informed member or the energy center of the unit’s culture.” They make decisions based on designing, delegating, authorizing, allocating resources, and deeming. Managers also use soft information that “links the organization with its environment” based on an intuitive decision-making process.

March, Simon, and Mintzberg laid the foundation for the study of managerial decision-making. Simon's Bounded Rationality theory recognizes the limitations in time and mental computational power as well as the situation's complexity in decision-making. While Simon suggests that, given enough information, humans would make economically rational decisions, Daniel Kahneman and Amos Tversky state that humans make irrational choices despite knowing better (Tversky & Kahneman, 1992). On the other hand, in empirical observations and interviews with managers, Mintzberg noticed that decision-making under ambiguity usually begins with a vague idea that is then matched to an opportunity that happens to be at hand, which then helps shape the vague idea.

From the HCI perspective, having access to the relevant information and sense-making process is vital for users to make the right decisions at the right time. Likewise, the context in which decisions are made must also be constructed clearly by choice architects. With the rise of technology, persuasive technologies have also weighed in on decision-making, especially on the internet, and in health and financial decisions for users. Therefore, it is now critical to understand who or what is behind these emerging technologies –the choice architects– affecting individuals’ decisions. Ideally, “a choice architect has the responsibility for organizing the context in which people make decisions” (Thaler & Sunstein, 2009). For example, while prospect theory mainly focuses on loss aversion to predict the human decision-making process¹, Nudge Theory promises to help people

¹ According to Prospect Theory, humans often do not choose logically and are susceptible to cognitive biases. For example, when asked to choose between receiving $900 or a 90% chance of winning $1000 (therefore, a 10% chance of winning $0), most avoid
make better choices for themselves without restricting their freedom of choice. However, some scholars critique prospect or cumulative prospect theory as inaccurate, too normative, and derived from improper methods and calculations (Sugden, 2009).

3.2 Human-AI Collaboration in the Workplace

With AI applications increasingly entering workplaces, the rapidly changing Human-AI relationship in these situations needs to be re-evaluated. The place and partnership of technology in human life has a long and indispensable history. Technological innovations have both helped and hindered ancient and modern societies (see Figure 4). This relationship has intensified with the introduction of computers into everyday life. Here, non-AI computing systems are characterized by automation, a system’s ability to perform well-defined tasks. While AI is characterized by autonomy, specifically the ability to perform specific tasks independently (Xu et al., 2021).

![Figure 4: The evolution of the Human-Machine-AI relationship across eras. (Xu et al., 2021, p. 9)]

the risk and choose $900. This is because the expected utility of both choices is the same. However, if asked to choose between losing $900 and a 90% chance of losing $1000, most prefer the 90% chance of losing $1000, thus engaging in risk-seeking behaviour to avoid the loss.

Nudge theory is based on the paternalistic view. The theory claims it is legitimate for choice architects to try to influence people’s behavior to make their lives longer, healthier, and better. Liberal Paternalism attempts to manipulate the choices people make to encourage better decisions, but in a way that doesn't limit what decisions they can make. Please see: Thaler, R. H., & Sunstein, C. R. (2009). Nudge: Improving decisions about health, wealth, and happiness. Penguin.
Since IBM’s Deep Blue beat Gary Kasparov at chess in 1997, human and machine partners have consistently beaten the individual machine. Kasparov’s vision for a new chess league consisted of centaurs — essentially partnerships between humans and AI. This Human-AI collaboration proposes a marriage of different yet complementary capabilities for effective decision-making (Dear, 2019). Therefore, utilizing AI in decision-making models may be more effective and beneficial for the organizational decision-making process to augment, not replace, human contributions (Jarrahi, 2018). So far, some frameworks have been proposed to address this issue (2019): (1) Human as a supervisor over automation that acts as an aid or an assistant; (2) Humans and autonomy acting as collaborating teammates; (3) Automation that oversees and acts as a limit on human performance. Shrestha et al. (2019) propose an extension to this decision-making structure by adding Human to AI and AI to Human “low to high” approach, matrix search space, the interpretability of the decision-making process and outcomes, the size of the alternative set, decision-making speed, and the replicability of decisions (see Figure 5).

<table>
<thead>
<tr>
<th>Organizational Structure</th>
<th>Specificity of the Decision Search Space</th>
<th>Interpretability</th>
<th>Size of the Alternative Set</th>
<th>Decision-Making Speed</th>
<th>Replicability</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full human to AI delegation</td>
<td>High (required for AI to function)</td>
<td>Low (due to absence of human involvement)</td>
<td>Large (not restricted by human capacity)</td>
<td>Fast (not restricted by human capacity)</td>
<td>High (computationally standardized)</td>
<td>Recommender systems, digital advertising, online fraud detection, dynamic pricing.</td>
</tr>
<tr>
<td>Hybrid 1: AI to human sequential decision making</td>
<td>High → Low (high in the first phase, low in the second phase)</td>
<td>High (due to human involvement in the final decision)</td>
<td>Large (due to involvement of AI in the first phase)</td>
<td>Slow (due to human decision-making as a bottleneck)</td>
<td>Low (vulnerable to human variability)</td>
<td>Idea evaluation, hiring.</td>
</tr>
<tr>
<td>Hybrid 2: Human to AI sequential decision making</td>
<td>Low → High (low in the first phase due to human involvement, and high in the second phase for AI)</td>
<td>Low (due to AI involvement in the final decision)</td>
<td>Small (due to human involvement in the first phase)</td>
<td>Slow (due to human decision-making as a bottleneck)</td>
<td>Low (vulnerable to human variability)</td>
<td>Sports analytics, health monitoring.</td>
</tr>
<tr>
<td>Aggregated human-AI decision making</td>
<td>Low (for decisions allocated to humans)</td>
<td>High (for decisions allocated to AI)</td>
<td>Small (same set of alternatives are evaluated by both humans and AI)</td>
<td>Slow (due to human decision-making as a bottleneck)</td>
<td>Partial (replicability only guaranteed in decision elements allocated to AI)</td>
<td>Top management teams, boards.</td>
</tr>
</tbody>
</table>

Figure 5: Organizational Decision-Making Structures Involving AI-Based Algorithms. Reprinted from (Shrestha et al, 2019 p.71).
Thus, hybrid models may extend capabilities for complex organizational decision-making by leveraging AI's increased computational information processing capacity and analytical approach. Likewise, human decision-making skills include a holistic, intuitive approach to uncertainty and equivocality — often what is considered common-sense situations (Jarrahi, 2018). In hybrid models, while AI systems support an analytical decision-making approach, they are less capable of understanding common-sense situations. As AI continues to develop, it is still challenged to capture inner logic and subconscious patterns, thus making it less likely to mimic human problem-solving that involves human intuition (Jarrahi, 2018) — giving humans an edge over AI in evaluating subjective, qualitative matters, such as norms, intangible political interests, and other complicated social, contextual factors.

When it comes to organizational decision-making, “aspects of both analytical and intuitive thinking are necessary but not sufficient for optimal business performance” (Martin & Martin, 2009, p. 6). A key question in the discussion above is the ability to assess the level of autonomy. When automation is applied to the decision-making process, it is crucial to choose appropriate levels and stages of automation, which lead to differential system performance benefits and pitfalls (Parasuraman et al., 2008). Some extremists are in favour of solely AI-driven decision-making. This view often cites human bias as the reason for excluding humans (as much as possible) from the decision-making process.

Despite AI’s increasing ability to learn and even exceed human performance for specific narrowly defined tasks, humans still vastly outperform AI with regards to generalizing knowledge to novel situations, creative problem solving and ambiguous situations due to our capabilities in nondeterministic analysis and understanding context and unstructured data (Jarrahi, 2018). Training AI requires vast amounts of data, but data is constantly changing and complex. Therefore, keeping humans in the loop is necessary to ensure the quality of information and interpretation. Even autonomous and “clever” AI will need to interact with humans and other AI systems in a social system (Abbass, 2019). The Computers Are Social Actors (CASA) paradigm demonstrated that humans often interact with computers like social agents (J.-E. R. Lee & Nass, 2010); thus, AI must become socially integrated.
3.3 Challenges with AI-Infused decision-making

Exponential increases in the application of AI have created a new type of collaboration: Human-AI teaming. This new collaboration is likely to cause massive disruption in the business environment. The speed of technological development often exceeds the linear nature of human development. Moreover, human decision-making can already be severely flawed in many situations (Kahneman, 2003; Pomerol & Adam, 2005; Simon, 1996; Tversky & Kahneman, 1992). Therefore, the aim is for AI to be a complementary tool to offset human shortcomings.

However, there is still a long way to go before Human-AI cooperation can be effectively implemented. For example, AI systems are currently situationally aware, but much less effective when it comes to the detection and explanation of anomalies as well as generally in improvised solutions. A shared mental model may thus be essential for Human-AI collaboration. While the AI should be able to understand and predict their human counterparts’ behaviours, they must also be selective in the information they present (Dietterich, 2020). Moreover, when decisions impact businesses and people, those decisions must be justified by the decision-maker. If part or all the decision is made by an AI or based on an AI-system outcome, then that too must be understood and justified. Some researchers insist that in developing the Human-AI decision-making process, justification of the decisions should be a central issue (Ferreira & Monteiro, 2021; Hoffman et al., 2018a). However, there is no consensus on what this justification should look like (Doran et al., 2017)
The missing link in Human-AI Collaboration for Decision-Making in Organizations: Trust

The aim of understanding and measuring trust is vital for a variety of reasons. For example, it helps users adopt and maintain a gradual and steady engagement with the system by decreasing risk, creating good and meaningful technological experiences, and assisting users in adopting and maintaining a progressive and steady relationship with the system (Gulati et al., 2019). Another vital issue from the perspective of HCI is that the decision-makers should trust the AI output. Otherwise, they are less likely to use AI output in their decision-making process. Therefore, if people understand how the AI reached its output, they may be more likely to trust the final decision. “Trust is the primary way to enhance the confidence of users with a system” (Ferreira & Monteiro, 2021). However, researching trust in Human-AI relations encounters the challenge that AI currently lacks the ability to explain its reasoning or justify its decisions in a human-like “thinking” process (Xu et al., 2021). Regardless, it is known that AI transparency can positively impact decision-making and overall trust (Trunk et al., 2020).

“It is also important to note that maximizing the user’s trust does not necessarily yield the best decisions from a human-AI collaboration. When trust is at maximum, the user accepts or believes all the recommendations and outcomes generated by the AI system” (Asan et al., 2020, p. 8). This excessive amount of trust in AI may have serious consequences. Researchers have found that if this trust is broken after people see mistakes being made by the AI, an "algorithm aversion" can develop where people stop trusting altogether (Prahl & Swol, 2017). On the other hand, complete trust may lead to "automation bias," where people overly rely on AI (Bogert et al., 2021).

As mentioned earlier, the complex multi-layer process of AI decision-making is generally not transparent, making it hard to predict and understand. For this reason, the "black box" phenomenon has become one of the major concerns in developing AI systems. It is essentially a problem of "inherent uncertainty since an ML system can make mistakes in its prediction based on learned patterns, and such uncertainty often cannot be fully captured before deployment using testing methods" (Y. Zhang et al., 2020, p. 2). There is a growing body of research suggesting that overcoming all these fundamental problems builds trust in AI applications (Ashoori & Weisz, 2019; Glikson & Woolley, 2020; Parasuraman et al., 2008). Similarly, understanding the process
of trust restoration after an initial trust for the AI system has been damaged in a vital field of study in human-machine teaming design (de Visser et al., 2018).

### 4.1 What is Trust?

Researchers in various fields have studied the role of trust in decision-making in diverse contexts (ChoJin-Hee et al., 2015). A critical point of division in trust literature is definitions emphasizing its cognitive aspects versus non-cognitive ones (Ferrario et al., 2020). For this reason, the definition of trust has been widely debated in various research disciplines but often focuses on the context of expectations and interactions between individuals. The following definition of trust is one of the most widely used ones in the literature: “the attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability” (J. D. Lee & See, 2004, p. 94). Trust (and distrust) from the perspective of HCI can be defined “as a sentiment resulting from knowledge, beliefs, emotions and other aspects of experience, generating positive or negative expectations concerning the reactions of a system and the interaction with it (Hoffman et al., 2018b, p. 41). Trust provides a successful foundation for effective conflict resolution, problem-solving, and team performance, particularly with teamwork. Without trust in a team, learning and cooperation are often impaired, leading to negative consequences (Kiffin-Petersen & Cordery, 2003). Rather than focusing on interpersonal trust, a growing body of research has also begun to address trust in AI in the business sphere. In the context of AI-infused decision making, Madsen and Gregor define trust as “the extent to which a user is confident in, and willing to act on the basis of, the recommendations, actions, and decisions of an artificially intelligent decision aid” (2000, p. 1).

### 4.2 Trust in Non-Human Decision-Makers

Trust plays an essential role in all human interactions and one of the primary motivators for individuals to adopt and use new technology is trust. Therefore, users' resistance to technology is frequently caused by a lack of trust (Lancelot Miltgen et al., 2013). While trust in humans grows over time due to repeated contact, trust in technology erodes over time due to experiences with faults and malfunctions. In other words, trust towards intelligent agents can decrease over time
and, once lost, can be hard to build up again. As Łapińska et al. state, “trust in artificial intelligence seems to be one of the key factors determining the level of this acceptance, and thus influencing the scale and effectiveness of implementation and use of artificial intelligence solutions in companies” (Łapińska et al., 2021, p. 2). When people understand how AI reaches its output, they may be more likely to trust the final decision (Ferreira & Monteiro, 2021). In other words, trust is becoming part of the bedrock of AI’s deployment, and transparency can positively impact decision-making and overall trust.

The European Commission's report 'Ethics principles for trustworthy AI' establishes a clear benchmark for evaluating trustworthy AI research and encourages international support for AI solutions that benefit businesses and the environment.

- **Human agency and oversight**: AI systems should enable equitable societies by supporting human agency and fundamental rights and not decrease, limit, or misguide human autonomy
- **Robustness and safety**: trustworthy AI algorithms require to be secure, reliable, and robust enough to deal with errors or inconsistencies during all life cycle phases of AI systems
- **Privacy and data governance**: citizens should have full control over their own data, while data concerning them will not be used to harm or discriminate against them
- **Transparency**: the traceability of AI systems should be ensured
- **Diversity, non-discrimination, and fairness**: AI systems should consider the whole range of human abilities, skills, and requirements and ensure accessibility
- **Societal and environmental well-being**: AI systems should be used to enhance positive social change and enhance sustainability and ecological responsibility
- **Accountability**: mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes Adapted from (Floridi, 2019).

Studies involving cognitive trust in AI show that the machine intelligence of the AI and how it is presented affect the nature of the trust that is developed. Emotional and cognitive trust can change even based on task characteristics. “With respect to receiving advice, users have more cognitive
trust in AI than in human advice on technical issues, but not on issues requiring social knowledge
[...]. However, when engaging in self-disclosure, users tend to be more honest and open with AI than with other humans, disclosing more personal information and engaging less in impression management behaviors” (Glikson & Woolley, 2020, p. 635). Previous research on algorithmic decision aids suggests mixed results in people’s acceptance of AI recommendations (Lee, 2018). For Instance, humans rely more on algorithms as tasks become more difficult (Bogert et al., 2021). Nevertheless, it should be noted that when a machine is “wrong,” it can be wrong in a far more dramatic way, with more unpredictable outcomes, than a human (O’Heigeartaigh, 2013).

Another issue is calibrating trust in AI-infused decision-making. Past HCI research has suggested that the humanization of technology increases positive feelings toward the technology and its use. However, this does not mean that people perceive the technology as more trustworthy or persuasive (Szeli, 2020). As discussed earlier, trusting too much or too little in an AI system can negatively impact interactions between users and the AI. Ideally, users have the appropriate level of trust given what the system can and cannot do. For example, initially stating that AI’s outcome may not be correct can decrease confidence in that outcome. Overall, users may prefer that AI because they do not over-trust it, thus reducing the likelihood they will be disappointed (see Figure 6).

Figure 6: Calibrating trust. Reprinted from (Turner et al., 2020, p.4).
4.3 Building Trust in the Workplace

Human-AI interactions need to be well-thought-out to help solve critical societal challenges (Jordan, 2019). “In this regard, trust is particularly relevant to the human-AI relationship because of the perceived risk embedded in human-AI relations, due to the complexity and non-determinism of AI behaviors” (Glikson & Woolley, 2020, p. 10). Several experiments show how easily humans can be manipulated by “black box” algorithms (Ghorbani et al., 2019). It has been shown that "Enhancing the explanatory power of intelligent systems can result in systems that are easier to use, and result in improvements in decision-making and problem-solving performance” (Nakatsu, 2006, p. 575). Therefore, researchers are trying to solve how AI can be interpretable, ethical, explainable, and trustworthy to allow understanding of intent, mutual predictability, and shared understanding. This is a call for further development in AI to allow for such beneficial characteristics while also mitigating risks. In this regard, trustworthy, responsible and Human-Centered AI (HCAI) is raising much interest. Shneiderman and Xu have proposed a human-centered AI (HCAI) approach (Shneiderman, 2022; Xu et al., 2022). Emphasizing user experience design, HCAI focuses on human users as the center of design thinking. HCAI designers note that consequential decisions, such as rejection of a loan or job interviews, raise additional questions. Therefore, these AI systems must help people understand those decisions by providing explanations they can understand, challenge or use to affect change (Shneiderman, 2020b). As aforementioned, mental models may be a solution. In humans, mental models are inferred from empirical evidence. Though people often cannot thoroughly explain their understanding, “with adequate scaffolding by some method of guided task reflection, people can tell you how they understand an event or system, they can describe their knowledge of it, and the concepts and principles that are involved” (Hoffman et al., 2018a, p. 9).

To solve this problem, several governmental and private-sector initiatives have issued principles to guide AI systems’ ethical development and deployment in recent years. Google, IBM, and the European Union's High-Level Expert Group are just a few examples. However, these principles are largely clustered around several specific concepts, including fairness, accountability, transparency, and privacy applied to AI systems (Jobin et al., 2019). “Despite their popularity, the abstract nature of AI ethics principles makes them difficult for practitioners to operationalize” (Madaio et al., 2020, p. 2). Some frameworks and tools try to make AI easier to use, although
many target data scientists and machine learning engineers as their end-users. Often those working with AI are mainly focused on studying the algorithms, rather than exploring how to develop AI systems that meet user needs (Asan et al., 2020; Shneiderman, 2022). Ultimately, it is critical that the end-users, such as executives and managers in businesses, understand the outputs of these AI systems (Xu, 2019) to improve their trust in those systems, trust in the decision-making process and adoption of this disruptive technology.
5 Methodology

Our literature review shows that trust is one of the most critical variables in human-to-human and human-to-intelligent-agent interactions, such as robots and algorithms. Since Human–AI collaboration in the business environment and its relation to trust is a relatively novel topic, not much research has been conducted thus far; however, it was enough to formulate directional research questions. Therefore, the primary focus of this research project is to identify the role of trust during Human-AI collaboration in managerial decision-making processes. To this end, this current research uses a mixed research method that includes quantitative data obtained through questionnaires and qualitative research obtained using semi-structured interviews conducted via the Zoom video conferencing platform. This integrative approach is used as an effective way to obtain a detailed understanding of Human-AI interaction and collaborative techniques. Using the mixed approach, together with quantitative and qualitative approaches, allows us to better understand research problems than using each approach individually (Ivankova et al., 2006). Given HCI’s core interest in technology that empowers people, the nature and dynamics of trust in the presence of Human-AI interactions is a research gap that is necessary to explore.

This research work is based on a three-stage methodology that allows the researchers to investigate current attitudes toward AI from managers’ perspectives. The first phase of our research focused on desk research of studies already conducted on the issue, most of which were mentioned above. In the second phase of our research, we conducted a pilot study to refine our pre-prepared questionnaire. After the pilot phase, we conducted a quantitative study of managers’ preferences regarding whether they trust AI in their working environment, using scenario-based questionnaires aimed at verifying the assumed hypothesis. Finally, we employed a qualitative study with semi-structured interviews to dive deeper into insights generated during the third phase.

The flow of the study is presented in Figure 7:
5.1 Pilot and Quantitative Study

As Lazar et al., (2009) underscores, a pilot study may aid in determining questions that could be difficult to understand. Pilot testing may also give researchers a more precise estimate of the interview's length. Researchers benefit from this approach since it makes them feel more at ease when conducting interviews. For this research, we first conducted a pilot study targeting three decision-makers (managerial, c-suite level) in Canada via the Canadian Society of Association Executives and Invest Ottawa’s “Book of Lists” in which there are more than 800 C-Level executives.

This pilot research helped refine the scenarios we had developed for the business environment – based on the scenarios used by Ashoori & Weisz (2019) (please see Online Questionnaire in Appendix B). These interviews were conducted virtually via Zoom video conference at a pre-arranged mutually convenient time and lasted approximately 40 minutes each. We transcribed the notes regarding the sessions. Participants were coded as Participant P1, P2, P3, etc.

5.1.1 Ethics and Recruitment

Carleton University Research Ethics Board-B (CUREB-B) reviewed and cleared our research methodology with the clearance number 116777 (see Appendix A). Following the pilot phase, 142 decision-makers (managerial, c-suite level) in Canada were recruited mainly through the survey platform panel, SurveyMonkey Audience, as well as C-Level executive organizations, such as the Canadian Society of Association Executives and Invest Ottawa, and finally, LinkedIn advertising was employed to reach C-Level managers across Canada. This allowed us to select participants.
from Canada who are in a managerial or C-level suite position and fit the study's participant inclusion criteria. To ensure that the desired target audience is being reached through SurveyMonkey Audience, we initially approved 40 respondents. Once their level of seniority and decision-making authority had been assessed, we proceeded with using SurveyMonkey Audience to reach the remaining respondents. We excluded participants who did not answer the whole questionnaire. In the end, we discarded 8 participants, leaving us with a final data set of n=134 participants. Criteria for being eligible to participate in the study included being 18 years or older and comfortable communicating in English. Additionally, these participants were currently employed in a managerial position or at the c-suite level and residing in Canada.

5.1.2 Consent and Compensation
Participants were asked to provide their consent at the beginning of the questionnaire through the Consent Form (please see Oral Consent Script in Appendix C and Survey Consent Form in Appendix D). Participants had to check the “I agree” box at the bottom of the page (immediately following the Consent Form) to attest that they understood the details of the study and give their consent to participate in it to be able to proceed with the questionnaire.

In the online questionnaire, participants who completed it in its entirety received compensation directly from SurveyMonkey as part of their agreement with their Audience Panel. All compensation payments were administered by SurveyMonkey. Following SurveyMonkey’s standard procedures, participants who withdrew from the study before completing the survey did not be compensated. This situation was explained clearly in the consent form. Once participants completed the questionnaire by pressing “submit” at the end of the survey, they could no longer withdraw from the study.

5.1.3 Data Collection
Through the online questionnaire, demographic information was collected. This data included age, gender, level of seniority, decision-making authority, and level of knowledge of AI systems in business environments.
The scenario elicitation technique has been used in both Human-Computer Interaction (Rosson & Carroll, 2002) and business (Schoemaker, 1995) for several decades. It is a proven tool to investigate people’s opinions, beliefs, and attitudes towards a subject in a controlled environment. Particularly for decision-making, the scenario technique has been widely studied in economics and strategic management (Borgonovo & Peccati, 2011).

We utilized the scenario-based trust scale developed by Ashoori and Weisz (2019) to analyze factors that influence the trustworthiness of the AI-infused Decision-Making process. This scale includes the most relevant dimensions of trust: (a) Overall trustworthiness, (b) Reliability, (c) Technical Competence, and (d) Personal Attachment.

We are particularly interested in trustworthy AI in the business environment. Some researchers insist that in developing the Human-AI decision-making process, justification of the decisions should be a central issue (Ferreira & Monteiro, 2021; Shneiderman, 2022). From this point of view, we measured overall trust with the following three justification factors as suggested:

(1) **Model Interpretability:** Some AI models are interpretable, for example, decision trees or rule-based scoring systems. For interpretable models, it is possible for users to examine and comprehend the interaction by which the model achieves the suggestion. Other models are viewed as "black boxes," such as deep neural networks, whose inner workings do not provide detailed information on how to make recommendations (Adadi & Berrada, 2018). Though recent discussions in AI literature debate the use of "explicability," "interpretability," "transparency," and “responsible” as counterarguments to the black box issue, our study uses interpretability as a more desirable feature of a model as explanations and transparency may incur additional trust problems (Rudin, 2019).

(2) **Model Confidence:** When producing decision-making suggestions, many AI models allow users to see the degree of confidence of the model (Chong et al., 2022). Higher confidence scores indicate greater probable accuracy, while lower confidence scores indicate less likely accuracy. Decision-makers must know when to trust or distrust an AI's prediction for these human-AI decision-making partnerships to be productive.
Thus, building a correct confidence model is essential to this process. Unfounded trust or distrust is not a desired outcome (Bansal et al., 2019; Y. Zhang et al., 2020). Therefore, confidence measures should be accurate and provide confidence values that are interpretable for users (Turner et al., 2020; Waa et al., 2020).

(3) **Humanoid Agent:** Another critical point is from the HCAI perspective that during the design process, humanizing computers can lead to problems that eventually affect decision-making (Shneiderman, 2020b). AI agents can be broadly grouped under three main categories: (1) Humanoid: There are general similarities between the AI agent design and the human anatomy, (2) Anthropomorphic: The AI agent design imitates some parts of the human anatomy, and (3) Non-humanoid: The AI agent design resembles any other living organism (Natarajan & Gombolay, 2020). Anthropomorphism or humanizing of computers may cause (1) mistaken usage based on emotional attachment to the systems, (2) false expectations of AI responsibility, and (3) incorrect beliefs about the appropriate use of AI (Robert, 2017).

From this perspective, our study explores three types of AI-assisted real-life HR, marketing/sales and finance decision-making scenarios:

1. **AI model is interpretable, and model confidence is presented,**
2. **AI model is interpretable, but no model confidence is presented,** and
3. **AI model is not interpretable, but model confidence is presented.**

Furthermore, participants were divided into two groups to conduct A/B testing.

**Group A:** 71 decision-makers were randomly selected from the primary participant pool. This group was presented with an anthropomorphized AI agent called SAM. SAM was created using the AI face generator, www.generated.photos (see Figure 8). The image that was chosen was selected to be neutral regarding nationality, gender and age.
Group B: 71 decision-makers were randomly selected from the primary participant pool. This group was presented with a non-anthropomorphized AI/dashboard agent.

5.1.4 Measures
Measuring trust is important in emerging technologies is important as it can help create positive experiences with the new technology, help users adopt it, and reduce uncertainty and anxiety related to its use (Sollner & Leimeister, 2012). The scale we utilized has several dimensions of trustworthiness, particularly in the concept of AI:

- Overall trustworthiness: the process ought to be trusted,
- Reliability: the process results in consistent outcomes,
- Technical competence: AI is used appropriately and correctly,
- Personal attachment: participants like the process.

5.2 Qualitative Study
For the qualitative part of our study, we recruited eight decision-makers across Canada. Each session took 45 minutes on average. We started by asking demographic questions to make the participants comfortable talking with us and get some information about the participants' current status (e.g., if they are currently using AI agents, if not, whether they have a plan to use it). We
continued by asking more in-depth questions about the participants' experience or perception of possible AI usage in their business. Participants were asked whether they felt the questions were ambiguous or if they could think of any questions that had not been asked. By asking these questions, we aimed to understand the participants' mindset and what they were looking for if they wished to use AI applications. The participants were willing to share their personal experiences and/or plans regarding "what would be the essential criteria for AI applications to be more trustworthy" as well as "would they prefer to interact with a humanoid AI agent, or a computer interface/dashboard" (see Interview Script in Appendix E).

5.2.1 Ethics and Recruitment
Carleton University Research Ethics Board-B (CUREB-B) reviewed and cleared our research methodology with the clearance number of 116777. As the quantitative study's initial results became computable, semi-structured interviews were begun with 17 participants who had completed the quantitative portion to gather more in-depth data. Due to the ongoing COVID-19 pandemic and the cross-Canada nature of the project, interviews were conducted via Zoom video conferencing platform. Five participants preferred to submit written responses to the interview questions. Additionally, three of the participants were from the pilot study and were, thus, removed from the results, leading to 14 total respondents for this portion of the study. Due to the ongoing COVID-19 pandemic as well as the cross-Canada nature of the project, these interviews were conducted via Zoom video conferencing platform.

5.2.2 Consent and Compensation
For the qualitative study, at the beginning of the interview, participants were asked to provide their oral consent after we read through the Oral Consent Form Script (please see Appendix D and E). The reason for using the Oral Consent Form is because, due to the COVID-19 pandemic, all research for this project was conducted virtually through video conferencing. In the consent form, they were reminded that the session would not be audio recorded. To make sure the participants understood their rights, the researcher again verbally explained the purpose of the study as well as the participants' right of withdrawal if they wished at any time prior to the beginning of the session. These participants did not receive any compensation.
5.2.3 Data Collection
We asked questions verbally and took notes in Microsoft Word. Each session's notes were recorded individually in separate files and transcribed by the primary researcher. After transcription, the recording files were deleted.

5.2.4 Data Analysis
For this study, we gathered data from notes taken by the primary researcher from our 14 participants. Before initiating the analysis, to get a better understanding and have the bigger picture of the result of the analysis, we first focused on the transcribed responses. After analyzing the overall result, we imported the analysis into a software called ‘NVivo R1 Pro provided by the MacOdrum Library at Carleton University. This process helped us to conduct thematic analysis more precisely. The thematic analysis provides a systematic element to data analysis (Ibrahim, 2012, p. 40). The phases of a thematic analysis are:(1) Becoming familiar with the data, (2) Generating initial codes, (3) thematically combining them, (4) Reviewing, (5) Defining, and (6) Writing. By using this methodology, we were able to identify the participants’ thematic mind map and how they justified whether they could trust an AI agent in their work environment.
6 Quantitative Analysis and Findings

6.1 Demographics

In terms of demographic distribution, women make up just over half of the Canadian population yet continue to be underrepresented in management positions. Only 9 percent of C-suite executives at Canada’s 100 largest publicly traded corporations are women (Women in Capital Markets, 2022). Therefore, it was much more challenging to reach female managers than to reach male managers. As seen in Table 1, in our study, 72 of the participants were male, while 52 were female. Three participants preferred not to state their gender, and one participant identified as non-binary. While 34.3 percent of the participants declared themselves as senior managers, 26.1 percent said they were owners, followed by 14.2 percent identifying as C-level executives, directors at 10.4 percent, President or CEO at 8.2 percent, and Vice-President at 6 percent (see Table 2).

In terms of knowledge of AI (see Table 3), those who say they are Not Knowledgeable and those who say they are Very Knowledgeable share the same percentage of 15.7. While 43.3 percent of the participants expressed themselves as Somewhat Knowledgeable, 25.4 percent stated that they were Knowledgeable.

Finally, the mean age of the participants was 46 years old.

Table 1: Gender distribution of participants

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>58</td>
<td>43.3</td>
<td>43.3</td>
<td>43.3</td>
</tr>
<tr>
<td>Male</td>
<td>72</td>
<td>53.7</td>
<td>53.7</td>
<td>97.0</td>
</tr>
<tr>
<td>Non-binary</td>
<td>1</td>
<td>.7</td>
<td>.7</td>
<td>97.8</td>
</tr>
<tr>
<td>Prefer not to answer</td>
<td>3</td>
<td>2.2</td>
<td>2.2</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>134</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
6.2 Reliability

The sample volume required for a similar quantitative analysis is in the range of 100-500, with a minimum of 100 (Kline, 2016). SPSS 26 and JASP programs were used to analyze the data. First, we conducted a Confirmatory Factor Analysis (CFA) to test the construct validity. Additionally, Cronbach’s Alpha coefficients were computed as $\alpha = 0.91$. Fit indices were used to investigate the fit between the expected and observed covariance matrices in CFA. The calculated score of SRMR = 0.049 shows that the SRMR ($\leq 0.08$) has good fit criteria. TLI, GFI, and CFI indexes of 0.90 and above indicate good fit, and above 95 indicate excellent fit (Brown, 2015; Tabachnick & Fidell, 2007). The following scores were computed on the scale: TLI: 0.91; GFI: 0.989; CFI: 0.940; 0.918.
Kolmogorov-Smirnov (KS) coefficients of the scores were obtained for each difference variable (Dashboard/SAM) use, Confidence (present/absent), and Interpretability (present/absent) to determine further analysis methods. Because our data does not show a normal distribution, the Mann-Whitney-U test was conducted to examine whether our independent variables, namely Interpretability and Confidence, differ from the scores obtained from the sub-dimensions of the scale.

6.3 Confidence and Interpretability Effect

The results of the Mann-Whitney U Test, in which the differentiation of the scores obtained from the sub-dimensions of the scale in terms of Confidence (Present/Absent) and Interpretability (Present/Absent) variables and the size of the effect were examined, are given in Table 4.

<table>
<thead>
<tr>
<th>Table 4: Confidence and Interpretability Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trustworthiness</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
0.8 [55, 56]. Trustworthiness mean score with Confidence ($\bar{X}=10.840$) is higher than without Confidence ($\bar{X}=9.776$).

The Reliability score shows a statistically significant difference in terms of Confidence ($U=15877.50$, $p<0.05$). The effect size of this difference has a very low effect (Partial $n^2=0.098$). Reliability mean score with Confidence ($\bar{X}=5.414$) is higher than without Confidence ($\bar{X}=5.187$). The Technical Competence score shows a statistically significant difference in terms of Confidence ($U=14597.00$, $p<0.05$). The effect size of this difference has a very low effect (Partial $n^2=0.159$). The Technical Competence mean score with Confidence ($\bar{X}=8.485$) is higher than the Technical Competence mean score ($\bar{X}=7.873$) without Confidence.

The Personal Attachment score shows a statistically significant difference in terms of Confidence ($U=13897.00$, $p<0.05$). The effect size of this difference has a very low effect (Partial $n^2=0.192$). Personal Attachment mean score with Confidence ($\bar{X}=8.485$) is higher than without Confidence ($\bar{X}=7.873$).

The Trustworthiness score shows a statistically significant difference in terms of Interpretability ($U=13520.00$, $p<0.05$). The effect size of this difference has a very low effect (Partial $n^2=0.206$). The average Trustworthiness score with Interpretability ($\bar{X}=10.929$) is higher than the average of Trustworthiness ($\bar{X}=9.597$) without Interpretability.

The Reliability score shows a statistically significant difference in Interpretability ($U=14963.00$, $p<0.05$). The effect size of this difference has a very low effect (Partial $n^2=0.141$). Reliability mean score with Interpretability ($\bar{X}=5.478$) is higher than without Interpretability ($\bar{X}=5.060$). The Technical Competence score does not show a statistically significant difference in Interpretability ($U=16556.50$, $p>0.05$).

The Personal Attachment score does not show a statistically significant difference in Interpretability ($U=16156.50$, $p>0.05$).
Therefore, regarding overall Interpretability, we fail to reject $H_{1:0}$ that the Interpretability of the AI agent’s decision-making process has no significant effect on the user’s overall trust in the process.

Thus, regarding overall Confidence, we reject $H_{2:0}$ that the presence of decision-making process confidence score by AI agent has no significant effect on the user’s overall trust in the process. Table 5 shows the results of the Mann-Whitney-U Test, in which the difference between the scores obtained in the Scenarios.

### 6.4 AI Interface/Dashboard Interface Effect

**Table 5: AI Interface/Dashboard Interface Effect on Subdimensions**

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Trustworthiness</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{x}$</td>
<td>$N$</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>-----</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>SAM</td>
<td>12,103</td>
</tr>
<tr>
<td></td>
<td>Dashboard</td>
<td>12,061</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>24,164</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>SAM</td>
<td>9,485</td>
</tr>
<tr>
<td></td>
<td>Dashboard</td>
<td>10,076</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19,561</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>SAM</td>
<td>9,485</td>
</tr>
<tr>
<td></td>
<td>Dashboard</td>
<td>9,712</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19,297</td>
</tr>
</tbody>
</table>

The Trustworthiness, Technical Competence, and Personal Attachment scores do not differ statistically in Scenario 1, Scenario 2, and Scenario 3 regarding SAM or Dashboard interface.
usage. On the other hand, the Reliability scores show a statistically significant difference in terms of SAM or Dashboard interface usage, particularly in Scenario 1 (U=1709.50, p<0.05) and Scenario 2 (U=1743.00, p<0.05).

In Scenario 1, the SAM interface's average Reliability score ($\bar{X}$=6.015) is higher than the Dashboard's Reliability score ($\bar{X}$=5.515). On the other hand, in Scenario 2, the SAM interface's average Reliability score ($\bar{X}$=4.956) is lower than the Dashboard interface's average Reliability score ($\bar{X}$=5.424).

Therefore, we fail to reject $H_{3;0}$ that a humanoid interface for an AI agent has no significant effect on the user’s overall trust in the process.

6.5 Authority Effect

When the normality of the Scenario 1, Scenario 2 and Scenario 3 scores obtained as a result of the application of the scale in three scenarios are examined according to the Authority (Final/Some) variable, Scenario 1 (KS$_{\text{final}}$=0.217, KS$_{\text{some}}$=0.170, p<0.05), Scenario 2 (KS$_{\text{final}}$=0.159, KS$_{\text{some}}$=0.153, p<0.05) and Scenario 3 (KS$_{\text{final}}$=0.183, KS$_{\text{some}}$=0.122, p<0.05) do not show normal distribution according to Authority (Final/Some) variable. Based on this result, the Mann-Whitney U Test was conducted to examine whether the scale scores obtained from different scenarios differ according to Authority.

The Mann-Whitney U Test results, in which the differentiation of the scores obtained in Scenario 1, 2 and 3 from the overall scale according to Authority (Final/Some) status and the size of the effect are examined, are given in Table 6.

<table>
<thead>
<tr>
<th>Table 6: Authority Effect (Final/Some)</th>
<th>$\bar{X}$</th>
<th>N</th>
<th>Mann-Whitney U</th>
<th>p</th>
<th>Partial n$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenario 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>35.015</td>
<td>69</td>
<td>2162.500</td>
<td>0.720</td>
<td>-</td>
</tr>
<tr>
<td>Some</td>
<td>35.431</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>35.216</td>
<td>134</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Scenario 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>31.246</td>
<td>69</td>
<td>1739.500</td>
<td>0.025</td>
<td>0.031</td>
</tr>
<tr>
<td>Some</td>
<td>28.492</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>29.910</td>
<td>134</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>31.058</td>
<td>69</td>
<td>1860.50</td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td>Some</td>
<td>29.000</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30.060</td>
<td>134</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When Table 6 is examined, it is seen that the total score obtained from Scenario 1 does not show a statistically significant difference in terms of Authority (Final/Some) \((U=2162.50, p>0.05)\). It is seen that the scale total score obtained from Scenario 3 does not show a statistically significant difference in terms of Authority (Final/Some) \((U=1739.50, p>0.05)\). However, the scale total score obtained from Scenario 2 shows a statistically significant difference in terms of Authority (Final/Some) \((U=1860.50, p<0.05)\). Additionally, the mean score of the group with the final decision from Scenario 2 \((\bar{X}=31,058)\) is higher than the mean score of the group with the Some decision from Scenario 2 \((\bar{X}=28,492)\) with a low effect \(\text{Partial } \eta^2=0.031\).

### 6.6 Knowledge Effect

Normality of the Scenario 1, Scenario 2 and Scenario 3 scores obtained from the scale (Knowledgeable/Somewhat Knowledgeable/Knowledgeable/Very Knowledgeable) was analyzed according to the variables, Scenario 1 \(1 (\text{KS}_{NK}=0.278, \text{KS}_{SK}=0.171, \text{KS}_K=0.212, \text{KS}_{VK}=0.146 \ p<0.05)\), Scenario 2 \(\text{KS}_{NK}=0.213, \text{KS}_{SK}=0.148, \text{KS}_K=0.212, \text{KS}_{VK}=0.261 \ p<0.05)\), and Scenario 3 \(\text{KS}_{NK}=0.174, \text{KS}_{SK}=0.104, \text{KS}_K=0.189, \text{KS}_{VK}=0.198 \ p<0.05)\). Our analysis reveals that it does not show a normal distribution according to the variables. Therefore, the results of the Kruskal Wallis H Test, are given in Table 7.
Table 7: Knowledge Effect

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Knowledge Level</th>
<th>$\bar{X}$</th>
<th>N</th>
<th>Kruskal Wallis H</th>
<th>p</th>
<th>Differences</th>
<th>Partial $n^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>1 Not Knowledgeable</td>
<td>32.714</td>
<td>21</td>
<td>5.139</td>
<td>0.162</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2 Somewhat Knowledgeable</td>
<td>35.207</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 Knowledgeable</td>
<td>35.941</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 Very Knowledgeable</td>
<td>36.571</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>35.216</strong></td>
<td><strong>134</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 2</td>
<td>1 Not Knowledgeable</td>
<td>25.524</td>
<td>21</td>
<td>17.702</td>
<td>0.001</td>
<td>1-2</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>2 Somewhat Knowledgeable</td>
<td>29.966</td>
<td>58</td>
<td></td>
<td></td>
<td>1-3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 Knowledgeable</td>
<td>29.588</td>
<td>34</td>
<td></td>
<td></td>
<td>1-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 Very Knowledgeable</td>
<td>34.667</td>
<td>21</td>
<td></td>
<td></td>
<td>2-4</td>
<td>3-4</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>29.910</strong></td>
<td><strong>134</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 3</td>
<td>1 Not Knowledgeable</td>
<td>25.571</td>
<td>21</td>
<td>16.888</td>
<td>0.001</td>
<td>1-2</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>2 Somewhat Knowledgeable</td>
<td>30.362</td>
<td>58</td>
<td></td>
<td></td>
<td>1-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 Knowledgeable</td>
<td>29.029</td>
<td>34</td>
<td></td>
<td></td>
<td>2-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 Very Knowledgeable</td>
<td>35.381</td>
<td>21</td>
<td></td>
<td></td>
<td>3-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>30.060</strong></td>
<td><strong>134</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The total score obtained from Scenario 1 does not differ statistically in terms of Knowledge (Not Knowledgeable/Somewhat Knowledgeable/Knowledgeable/Very Knowledgeable) ($H=5.139$, $p>0.05$). It is seen that the scale total score obtained from Scenario 2 differs statistically in terms of Knowledge (Not Knowledgeable/Somewhat Knowledgeable/Knowledgeable/Very Knowledgeable) ($U=17.702$, $p<0.05$). It is seen that the scale total score obtained from Scenario 3 differs statistically in terms of Knowledge (Not Knowledgeable/Somewhat Knowledgeable/Knowledgeable/Very Knowledgeable) ($U=16.888$, $p<0.05$).

The mean score of the Not Knowledgeable group ($\bar{X}=25.524$) from Scenario 2, the Somewhat Knowledgeable ($\bar{X}=29.966$), the Knowledgeable ($\bar{X}=29.588$) and the Very Knowledgeable ($\bar{X}=34.667$) scores from Scenario 2 are found to be lower than their average scores. The Very Knowledgeable ($\bar{X}=34.667$) score averages obtained from Scenario 2 were obtained from the
Scenario 2 of the Not Knowledgeable ($\bar{X}=25.524$), Somewhat Knowledgeable ($\bar{X}=29.966$) and Knowledgeable ($\bar{X}=29.588$) scores higher than their average score. It is seen that the effect of this difference is low ($\text{Partial } \eta^2=0.120$).

The very knowledgeable ($\bar{X}=35.381$) group's score averages from Scenario 3 are higher than Not Knowledgeable ($\bar{X}=25.571$), Somewhat Knowledgeable ($\bar{X}=30.362$), and Knowledgeable ($\bar{X}=29.029$) mean score. It is seen that the mean score of the Not Knowledgeable ($\bar{X}=25.571$) group from Scenario 3 is lower than the mean score of the Somewhat Knowledgeable ($\bar{X}=30.362$) group from Scenario 3.
7 Qualitative Analysis and Findings

For this qualitative study, we gathered data from notes taken by the primary researcher from 14 participants. Before initiating the analysis, to better understand the bigger picture of the result of the analysis, we focused on the transcribed responses. After this initial review, we imported the analysis into a software called NVivo R1 Pro provided by the MacOdrum Library at Carleton University. This process helped us to a more precise approach as the thematic analysis provides a systematic element to data analysis (Ivankova et al., 2006). Particularly, we identified the participants’ thematic mind map and how they justified whether they could trust an AI agent in their work environment.

Based on this, our analysis allowed us to classify the codes under four themes: (1) Operational versus strategic decisions, (2) AI transformation and organizational change management, (3) Trust versus bias in AI-infused decision-making, and (4) AI design and interaction styles. These themes were defined based on the analysis of the interviews conducted with the participants and then refined by analyzing the open-ended questions within the questionnaire.

RQ1 What facets make the Human (Manager)-AI decision-making process trustworthy? (Themes A, B and C)

7.1 Operational Versus Strategic Decisions

Our qualitative research indicates that AI decision-making is often not interpretable, making it hard for users to predict and understand. For this reason, the “black box” phenomenon has become one of the key challenges in the decision-making process, according to our participants. Several participants indicated that they believe AI adoption may help solve critical strategic difficulties encountered in their business. At the same time, only a small portion of organizations identified themselves as genuine AI users with extensive AI adoption in their business processes. When asked what decision-making processes they would like to delegate to AI, many indicated that the use of AI should be evaluated on a per case basis. However, if they were to generalize, “automated planning” and “simplified decision-making” were the preferred AI uses. Should there be something outside of the routine, then participants expressed that there would be a need for
human (management) intervention. In other words, AI is still seen as a tool for operational decisions than strategic ones:

“From my perspective, AI strategy should not be completely delegated to the AI or even IT department where AI applications are created or implemented, because AI tools may go beyond simply enhancing productivity and instead may lead to changes in the strategy. My main role here is to decide our strategy” (P.11).

Almost all participants state that they are against the use of AI in human resources recruitment. They believe that the use of AI decision-making has so far not been beneficial in HR, even though HR managers are currently employing such screening tools. On the other hand, they indicted the use of AI in fields that rely on both incremental data and observations, such as finance and marketing.

While only two participants stated that they received AI assistance in strategic decisions, both participants noted they usually compare AI results with “non-AI” decision support systems. Adding that they still believe their “gut feelings” are as crucial as data-driven decisions in their business. Participants also underlined that data is their company’s most critical strategic asset. None of the executives we interviewed had bought data from suppliers such as data brokers. Instead, they were willing to create and maintain their own unique data.

Additionally, those who used AI to assist decision-making in their companies preferred to “see the whole picture” or “be involved in the process from the beginning” rather than just see the final report prepared by the technical department.

7.2 AI Transformation and Organizational Change Management

Most C-level managers we interviewed believe that significant technology changes cannot be implemented at a business level if management cannot clearly understand or describe what advances are coming. For this reason, many participants indicated that they participate in AI conferences regularly, even though 3 participants noted having received technical “crash courses” from experts to stay up to date. According to some of our participants, CEO and C-level management must provide a clear vision of AI that goes beyond buzzwords. Without a clear
articulation and vision of the AI transformation and its operational implementation, the change would be impossible to achieve. Participants believe that the IT or technical department will play an essential role in the overall AI transformation in their company as they have played a crucial role in digital transformation over the last decade. However, they are aware of the importance that this transformation must happen simultaneously with other departments. In other words, every business process may need to be optimized or disrupted with AI. No single accepted recipe will lead to an ideal AI strategy for organizations. Moreover, managers feel they have more responsibility in the current AI revolution, comparing previous digital or software-based transformations. Thus, many of the participants highlighted two terms:

- **Culture**: Company culture must be aligned with the new AI implementation. If the AI implementation or transformation is not complemented with cultural change, it will fail.
- **Diversity**: According to executives, inclusiveness and diversity are indispensable values and must not be undermined by raising a data-centric view of the world.

Finally, a significant portion of participants noted that because the AI systems are data-hungry, it requires “knowledge transfer” between management (who are not literate about analytics-related fields) and analytics experts.

### 7.3 Trust Versus Bias in AI-Infused Decision-Making

With regards to participants who have not started to use AI capabilities, one of the main reasons cited was, in part, a lack of trust and confidence in AI. Noting that AI can give managers means to make better decisions, one participant underlines:

“*However, there is a risk that technology will exacerbate human shortcomings, as we see on daily life while using devices such as Google, etc., like our tendency to have preconceptions about some people. And there is the reverse of the model; if you are not involved in the decision-making process and leave everything to AI, it most likely thinks what it is offering the right suggestion and perhaps will keep doing so for lack of feedback.*”

*(P.5)*
Many participants agree that it is imperative that time and effort be prioritized when building AI models for their organizations. As another participant states:

“The first iteration will never be one-size-fits-all, and effort needs to be made to encounter as many scenarios as possible for training. I think it’s also important to compare results with work being done by people to confirm its accuracy but also to see how it can simply enhance what is already being done now (and eventually shift to mainly AI automation). This would instill trust in the business and the people currently doing the work.” (P.14)

RQ2 Does trust in AI depend on the degree to which the AI agent is humanized?

7.4 AI Design and Interaction Styles

Participants are generally optimistic that AI may create efficient interactions in the business context. There is a consensus that AI can create better experiences for business managers if the technology is developed to be interpretable. All the executives we spoke to agreed that better design and usability are still indispensable when it comes to any use of technology. Furthermore, they agreed that UX might affect their decision-making capabilities dramatically in AI technologies. According to participants, when we questioned this preference, a better user experience is strongly related to “efficiency.”

In terms of AI interaction styles, conversation has already become increasingly prevalent in our daily lives, such as Alexa, Siri, and Google Maps. Perhaps, these devices or apps may have created an interaction idea for participants. As one participant says:

“I would love to be able to talk with AI about the result it has shown. In this way, I feel more comfortable with the result it suggests. To be honest endless layers of menus are not really helpful.” (P.1)

Likewise, another participant states, “The design of AI should be able to hint if the process is a black box” (P.7). Similarly, Participant 2 explains the importance of auditing AI:

“We are mainly emotional creatures prone to trust things or persons [...] another problem arises when you describe something with your design, you create this and trust it, right? So, the question is, who is going to audit this design?”
8 Discussion & Conclusion

AI is becoming increasingly pervasive in our everyday lives, both in the personal and business spheres. Like all disruptive technologies, this transformation is expected to bring new benefits to business solutions and raise some concerns. Explaining the process may be critical in improving Human-AI interaction by helping users accept, trust, and feel in control of the technology. Such methods are slowly being applied to widely used applications (Weld & Bansal, 2018). Though explanations have been shown to improve user understanding of AI systems, whether this helps improve user trust and acceptance are unclear. Potential gaps appear to exist between algorithmic explanations and end-user needs (Liao et al., 2020).

While AI can help make better and easier planning and decisions in organizations, our results show that managers still need to understand how AI reaches its output, and then they may be more likely to trust the final decision. Our mixed-method analysis indicates that, even though participants were eager to see AI-infused decision-making used in managerial decision-making, they were hesitant to do so if the decisions were strategic or directly impacted humans, such as in the HR scenario. On the other hand, if the decision were operational and indirectly affected humans, such as in the investments and marketing strategy scenarios that required analysis of large amounts of data, the participants were inclined to trust AI-infused more readily. This result was also demonstrated in the Reliability facet of the scale.

AI-based interfaces, such as speech recognition and gesticulation, require more human input. This viewpoint may open new possibilities for HCI professionals to create human-centered systems instead of traditional approaches, which are often criticized (Xu et al., 2022). Several participants noted their preference for a humanoid AI interface, stating that the use of conversation would be their preferred interaction method. When questioned further about this preference, participants’ expectations about interpretability and model confidence were intertwined. Many decision-makers equated conversations with an AI agent as a means to question and better understand how the final decision was made (Ferrario et al., 2020). This need to converse stemmed from a desire for more information about how the decision was being made (Interpretability) and assessing the level of
confidence in that decision (Model Confidence). An ideal experience may combine accessibility and inclusive design to make it not only compliant with principles or standards but truly reachable and interpretable to all. Furthermore, parallel to Shneiderman’s suggestions, users, in our case managerial decision-makers, did not prefer post hoc decision-making processes but implied that real-time access to several design metaphors was more desirable (2022). Nonetheless, it appears that AI is a powerful new technology that will continue to have more and more impact on our lives. To truly benefit from its potential, it will be critical to develop trusted systems for users and developers/programmers. A holistic and multidisciplinary approach that involves AI developers, users and policymakers is needed to build trusted AI systems and decision-making processes.

8.1 Future Work
Our study shows that further research and exploration into the acceptance and trust of AI-infused decision-making processes is critically needed within the business realm. Such research will help better understand what is needed in technology and process design to help facilitate an ethical and safe integration of AI within organizations. Furthermore, the human decision-makers within organizations –namely managers– should be the target users in future research. One of its goals is to design better interfaces and organizational processes for Human-AI collaboration, focusing on interpretability and paying attention to the advantages and disadvantages of humanoid design as it influences user trust. To this effect, our study also explored the relationship between demographics and trust as well as knowledge of AI and authority. Though no strong relationship was present in this study, these effects may be a subject for future studies. As Xu et al. (2021) suggest, HCI professionals need to proactively conduct applied research through cross-industry and cross-disciplinary collaborations to increase their influence. Academic scholars should actively do the same to increase AI-related collaborative projects between industry and academia. Furthermore, Morris (2020) points out that a larger pool of AI users, including underrepresented groups, such as those with disabilities and the elderly, is another critical area for both scholars and professionals to focus on as it pertains to a relationship of understanding and trust.

8.2 Limitations
Nevertheless, we are aware that our research has limitations. Due to the difficulty of reaching C-level businesspeople throughout Canada, our sample size of 134 participants may not represent all
C-level managers nationally. Additionally, our A/B testing used a static, visual image of a humanoid/anthropomorphic AI agent. A 3-D or live version of an AI agent with conversation capacity would provide a closer real-life experience.
Appendices

Appendix A – CUREB-B Ethics Clearance

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

The following research has been granted clearance by the Carleton University Research Ethics Board-B (CUREB-B). CUREB-B is constituted and operates in compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2).

Ethics Clearance ID: Project # 116777

Project Team Members: Mr. Serdar Tancer (Primary Investigator)
Alejandro (Alex) Ramirez (Research Supervisor)

Study Title: Human-AI Collaboration in Managerial Decision-Making Processes: Exploring the Role of Trust

Funding Source: (If applicable):

Effective: December 17, 2021 Expires: December 31, 2022

This certification is subject to the following conditions:

1. Clearance is granted only for the research and purposes described in the application.
2. Any modification to the approved research must be submitted to CUREB-B via a Change to Protocol Form. All changes must be cleared prior to the continuance of the research.
3. An Annual Status Report for the renewal or closure of ethics clearance must be submitted and cleared by the renewal date listed above. Failure to submit the Annual Status Report will result in the closure of the file. If funding is associated, funds will be frozen.
4. During the course of the study, if you encounter an adverse event, material incidental finding, protocol deviation or other unanticipated problem, you must complete and submit a Report of Adverse Events and Unanticipated Problems Form.
5. It is the responsibility of the student to notify their supervisor of any adverse events, changes to their application, or requests to renew/closed the protocol.
6. Failure to conduct the research in accordance with the principles of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2nd edition and the Carleton University Policies and Procedures for the Ethical Conduct of Research may result in the suspension or termination of the research project.

IMPORTANT: Special requirements for COVID-19:

If this study involves in-person research interactions with human participants, whether on- or off-campus, the following rules apply:
1. Upon receiving clearance from CUREB, please seek the approval of the relevant Dean for your research. Provide a copy of your CUREB clearance to the Dean for their records. See Principles and Procedures for On-campus Research at Carleton University and note that this document applies both to on- and off-campus research that involves human participants. Please contact your Dean's Office for more information about obtaining their approval.

2. Provide a copy of the Dean's approval to the Office of Research Ethics prior to starting any in-person research activities.

3. If the Dean's approval requires any significant change(s) to any element of the study, you must notify the Office of Research Ethics of such change(s).

Upon reasonable request, it is the policy of CUREB, for cleared protocols, to release the name of the PI, the title of the project, and the date of clearance and any renewal(s).

Please email the Research Compliance Coordinators at ethics@carleton.ca if you have any questions.

CLEARED BY: 

Date: December 17, 2021

[Signature]

Bernadette Campbell, PhD, Chair, CUREB-B

[Signature]

Kathryne Dupre, PhD, Co-Chair, CUREB-B
Appendix B – Online Questionnaire

Online Questionnaire through SurveyMonkey

PAGE 1:

Carleton Survey Consent Form script will be presented on the first page of this online questionnaire.

PAGE 2:

Q1: Which gender do you identify as?
- Woman
- Man
- Non-binary
- Gender-fluid
- Trans-woman
- Trans-man
- Two-Spirit
- Prefer not to answer

Q2: What is your age? (Single textbox – whole number)

Q3: What is your level of seniority? (Please select the closest match to your job title)
- Manager
- Director
- Vice President
- Senior Vice President
- C level executive (CIO, CTO, COO, CMO, etc.)
- President or CEO
- Owner

Q4: What level of decision-making authority do you have?
- Final decision-making authority (individually or as part of a group)
- Significant decision-making or influence (individually or as part of a group)
- Minimal decision-making or influence

Q5: What is your level of knowledge of AI systems in business environments?
- Very knowledgeable
- Knowledgeable
- Somewhat knowledgeable
- Not knowledgeable

PAGE 3:

Q6: Please read Scenario #1:
A company’s future real estate investments need to be determined. Here’s what you know about how these future real estate investments will be determined:

- The manager will use an AI model to help determine what real estate investments will bring the company the highest returns in 10 years, but the final decision will be made by the manager.
- The AI model was created by a team of data scientists, who trained it using historical data on real estate trends in the region.
- By examining the model, one is able to understand exactly why this decision regarding future real estate investments was made and what factors were used in making that decision.
- Information about the data used to train the model, including where that data came from, how much data was used for training, how much data was used to evaluate the accuracy of the model, and that model’s accuracy will be made available.
- Information about other recommendations made for other real estate investments will be made available.
- Information about how confident the AI model is in its 10-year real estate investment recommendation will be available.

PAGE 4:

Q7: Please rate the following questions regarding Scenario #1 from “Strongly disagree” to “Strongly agree.”: (All items below are rated on 4-point Likert scales: “Strongly disagree,” “Disagree,” “Agree,” and “Strongly agree.”)

a. This decision-making process is trustworthy
b. I would change one or more aspects of this decision-making process to make it trustworthy*
c. This decision-making process will produce a fair outcome for the person affected by the decision
d. The decision maker needs more information about how the AI model was trained and tested in order to trust the process*

e. This decision-making process would always make the same recommendation under the same conditions

f. The outcome of this decision will be consistent with other decisions made for other people

g. The use of an AI model is appropriate in this scenario

h. This decision will be made based on reliable information

i. I trust that the technical implementation of the AI model is correct

j. I am confident in this decision-making process. I feel that it works well

k. I am wary of this decision-making process*

l. I like this decision-making process

(*) Items with an asterisk will be reverse-coded.

PAGE 5:

Q8: Please read Scenario #2:

(In the A/B testing, Group A will be presented with an image of the following humanoid AI agent, while group B will not be presented with an image. Other than the image, both Group A and Group B will be shown identical scenarios and follow-up questions.)

Resumes for a new company position need to be reviewed, and top candidates selected for on-site interviews. Here's what you know about how the review and selection of top applicants will be determined:

- The manager will use an AI model to help review the hundreds of applications received and determine the top 10 candidates, but the final decision will be made by the manager.
- The AI model was created by a team of data scientists, who trained it using historical data on HR hiring procedures and job applications.
- By examining the model, one is able to understand exactly why this decision regarding the top 10 candidates for the position was made and what criteria were used in making that decision.
• Information about the data used to train the model, including where that data came from, how much data was used for training, how much data was used to evaluate the accuracy of the model, and that model’s accuracy will be made available.
• Information about other recommendations made for other hiring situations will be made available.
• Information about how confident the AI model is in its recommendation of the top 10 candidates will NOT be made available to you

PAGE 6:

Q9: Please rate the following questions regarding Scenario #2 from “Strongly disagree” to “Strongly agree.”: (All items below are rated on 4-point Likert scales: “Strongly disagree,” “Disagree,” “Agree,” and “Strongly agree.”)

a. This decision-making process is trustworthy
b. I would change one or more aspects of this decision-making process to make it trustworthy*
c. This decision-making process will produce a fair outcome for the person affected by the decision
d. The decision maker needs more information about how the AI model was trained and tested in order to trust the process*
e. This decision-making process would always make the same recommendation under the same conditions
f. The outcome of this decision will be consistent with other decisions made for other people
g. The use of an AI model is appropriate in this scenario
h. This decision will be made based on reliable information
i. I trust that the technical implementation of the AI model is correct
j. I am confident in this decision-making process. I feel that it works well
k. I am wary of this decision-making process*
l. I like this decision-making process

(*) Items with an asterisk will be reverse-coded.

PAGE 7:

Q10: Please read Scenario #3:

(In the A/B testing, Group A will be presented with an image of the following humanoid AI agent, while group B will not be presented with an image. Other than the image, both Group A and Group B will be shown identical scenarios and follow-up questions.)
Pricing needs to be determined for a new flagship product that will be launched. Here’s what you know about how this product pricing will be determined:

- The manager will use an AI model to help determine what the new product pricing will be, but the final decision will be made by the manager.
- The AI model was created by a team of data scientists, who trained it using historical data on new product pricing across multiple sectors.
- By examining the model, you are NOT able to understand exactly why this decision regarding the pricing for the new product was made and what factors were used in making that decision.
- Information about the data used to train the model, including where that data came from, how much data was used for training, how much data was used to evaluate the accuracy of the model, and that model’s accuracy will be made available.
- Information about other recommendations made for determining the price of other new product will be made available.
- Information about how confident the AI model is in its recommendation regarding the new product price will be available.

**Q11:** Please rate the following questions regarding Scenario #3 from “Strongly disagree” to “Strongly agree.”: (All items below are rated on 4-point Likert scales: “Strongly disagree,” “Disagree,” “Agree,” and “Strongly agree.”)

a. This decision-making process is trustworthy
b. I would change one or more aspects of this decision-making process to make it trustworthy*
c. This decision-making process will produce a fair outcome for the person affected by the decision
d. The decision maker needs more information about how the AI model was trained and tested in order to trust the process*
e. This decision-making process would always make the same recommendation under the same conditions
f. The outcome of this decision will be consistent with other decisions made for other people
g. The use of an AI model is appropriate in this scenario
h. This decision will be made based on reliable information
i. I trust that the technical implementation of the AI model is correct
j. I am confident in this decision-making process. I feel that it works well
k. I am wary of this decision-making process*
l. I like this decision-making process

(*) Items with an asterisk will be reverse-coded.

PAGE 9:
Thank you for participating in this survey.
Appendix C – Oral Consent Script

Research Consent Form Script for Oral Consent

1. Hello, my name is Serdar Tuncer and I am a Master’s student in the Computer Sciences Department, Human-Computer Interaction Program at Carleton University. I am working under the supervision of Prof. Alejandro Ramirez.

2. I would like to invite you to participate in a study titled “Human-AI Collaboration in Managerial Decision-Making Processes: Exploring the Role of Trust”. This study aims to explore the role of trust during collaborative interactions between humans and artificial intelligence agents in scenarios where decisions are made in a business environment.

The study involves one semi-structured interview conducted virtually via the Zoom video conferencing platform about trust in AI in business environments where there is Human-AI collaboration in managerial decision-making. There will be no audio or video recording of the interview. Notes from the interview will only be transcribed by the principal researcher.

We estimate that the interview will take about 30 minutes to complete. Your participation in this interview is voluntary, and you may choose not to take part, or not to answer any of the questions. If you decide to withdraw after the interview, your responses will be removed if you notify the researcher up until two weeks after your interview, or until February 28th, 2022.

We do not anticipate any risks from taking the survey, nor do we anticipate that you will derive any benefit.

We will treat your personal information as confidential, although absolute privacy cannot be guaranteed. No information that discloses your identity will be released or published without your specific consent. However, research records identifying you may be inspected by the Carleton University Research Ethics Board for the purpose of monitoring the research. The results of this study may be published, but the data will be presented so that it will not be possible to identify any participants. All research data will be password-protected, and any hard copies of data will be kept in a locked cabinet at Carleton University.

You will be assigned a pseudonym so that your identity will not be directly associated with the data you have provided. All data, including coded information, will be kept in a password-protected file on a secure computer.

“In-session” data, such as the audio, video, and chat transcript from the interview, will not be recorded. Only notes will be transcribed for research purposes. We will password protect any
research data that we store or transfer. Your de-identified data will be retained for a period of 5 years and then securely destroyed.

This research has been cleared by Carleton University Research Ethics Board-B (Clearance #116777). If you have any ethical concerns with the study, please contact the Carleton University Research Ethics Board – B by phone at 613-520-2600 ext. 4085 or by email at ethics@carleton.ca. During Covid, the Research Ethics Staff are working from home without access to their Carleton phone extensions. Accordingly, until staff return to campus, please contact them by email.

You can also reach me at serdartuncer@cmail.carleton.ca. You may contact my supervisor at Alex.Ramirez@carleton.ca.

3. **Statement of consent**

Do you have any questions about this study or need any clarification?  
Do you voluntarily agree to participate in the study?  
Yes_______ No_______

Date: ______________________  
Participant’s Name/Pseudonym/Initials (as appropriate): ______________________

**Research team member who interacted with the subject**

I have explained the study to the participant and answered any and all of their questions. The participant appeared to understand and agree. I provided a copy of the consent information to the participant for their reference.

__________________________  
Signature of researcher  
__________________________  
Date
Appendix D – Survey Consent Form

Research Consent Text for Online Survey

4. **Name and Contact Information of Researchers:**
   
   Serdar Tuncer  
   Computer Sciences Department, Human-Computer Interaction Program.  
   Tel.: 613-276-5751  
   Email: serdartuncer@cmail.carleton.ca  
   Supervisor and Contact Information: Prof. Alejandro Ramirez, Alex.Ramirez@carleton.ca, Information Systems (IS), Sprott School of Business, Carleton University

5. **Project Title**
   
   Human-AI Collaboration in Managerial Decision-Making Processes: Exploring the Role of Trust

6. **Carleton University Project Clearance**
   
   Clearance #: 116777  
   Date of Clearance: 2021

7. **Invitation**
   
   We are asking you to complete this survey because you are a decision-makers who currently employed in a managerial position or c-suite level, are 18 years of age or older, and reside in Canada. This survey is being conducted by Serdar Tuncer of the Carleton University Department of Computer Sciences (serdartuncer@cmail.carleton.ca, 613-276-5751) working under the supervision of Prof. Alejandro Ramirez (Alex.Ramirez@carleton.ca, 613-520-2600 ext. 2397).

   **Objectives and Summary:**
   
   The aim of this study is to better understand the role of trust during collaborative interactions between humans and artificial intelligence agents in scenarios where decisions are made in a business environment.  
   
   We estimate that the survey will take about 15 minutes to complete. Your participation in this survey is voluntary, and you may choose not to take part, or not to answer any of the questions. You may withdraw from this study prior to starting the questionnaire or during the questionnaire by closing the browser window at any time before completing and submitting it. Once you have completed the questionnaire by pressing “submit” at the end of the survey, you may no longer withdraw from this study.

   **Risks and Benefits:**

61
We do not anticipate any risks from taking the survey. You will receive a compensation directly from SurveyMonkey as part of your existing agreement with SurveyMonkey’s Audience panel.

**Confidentiality and Data Storage:**
We will treat your personal information as confidential, although absolute privacy cannot be guaranteed. No information that discloses your identity will be released or published without your specific consent. Research records may be accessed by the Carleton University Research Ethics Board in order to ensure continuing ethics compliance.

The results of this study may be published, but the data will be presented so that it will not be possible to identify you, unless you give consent. All research data will be password-protected, and any hard copies of data will be kept in a locked cabinet at Carleton University.

Your data will be stored and protected by SurveyMonkey (Parent Company SVMK Inc.) platform, on servers located in Canada, but may be disclosed via a court order or data breach. All data will be kept confidential, unless release is required by law (e.g. child abuse, harm to self or others).

We will password protect any research data that we store or transfer. Your data will be retained for a period of 5 years and then securely destroyed.

**REB Review and Contact Information:**
This research has been cleared by Carleton University Research Ethics Board-B (Clearance #116777). If you have any ethical concerns with the study, please contact the Carleton University Research Ethics Board – B by phone at 613-520-2600 ext. 4085 or by email at ethics@carleton.ca. During Covid, the Research Ethics Staff are working from home without access to their Carleton phone extensions. Accordingly, until staff return to campus, please contact them by email.

**Direct Consent:**
I voluntarily agree to participate in this study.

- [ ] Yes
- [ ] No
Appendix E – Interview Script

INTerview Guide

Human-AI Collaboration in Managerial Decision-Making Processes: Exploring the Role of Trust

Serdar Tuncer
Computer Sciences Department, Human-Computer Interaction Program.
Tel.: 613-276-5751
Email: serdartuncer@cmail.carleton.ca

Introduction:
Thank you [Insert Name] for participating in this study. I believe you have already read the email invitation. I will now go over some additional details about our “Human-AI Collaboration in Managerial Decision-Making Processes: Exploring the Role of Trust” study so that I may receive your oral consent prior to beginning the study. If you wish to receive a written copy, I can share that as well to your email address. This research has been cleared by Carleton University Research Ethics Board-B (Clearance #116777).

Oral Consent:
[Read through “Oral Consent Script” and receive consent]

Questionnaire:
If you are ready, shall we start? I will now send you a link via the Zoom chat to a questionnaire on SurveyMonkey that I would like you to fill out. This questionnaire presents three scenarios of collaborative interactions between humans and artificial intelligence agents where decisions are made in a business environment [share SurveyMonkey link]. Please let me know if you have any questions as you progress.

Now that you have completed the questionnaire, I would like to ask you a few additional questions:

1) Regarding the three scenarios presented in the questionnaire, did you find they reflect the current use of AI in the business environment?
   a. If no: What details would make these scenarios a more accurate reflection of the current AI – strategic decision-making environment?

2) Do you use Artificial Intelligence (AI) to make strategic decisions in your company?
   a. If yes: Are the decision-making AI applications you employ user-friendly?
      i. If yes: How and when are they used? What is the workflow?
ii. If no: What would make them more user-friendly?

b. If no: Do you intend to use them in the short or medium-term?

3) What is your opinion on the authority of AI in making strategic decisions within a business environment?
   a. Do you (or would you) trust AI to make strategic decisions in your workplace?
      i. If yes:
         1. In which circumstances?
         2. Are there any situations where you wouldn’t trust AI to make strategic decisions?
      ii. If no:
         1. Why?
         2. What would have to change for you to start trusting AI in strategic decision-making?

4) What are the essential criteria for AI applications to be more trustworthy?
   a. What criteria would make AI less trustworthy?

5) If given the option, would you prefer to interact with a humanoid AI agent or a computer interface/dashboard?
   a. Why do you prefer this interface type?

6) Is there anything you would like to add?

Conclusion:
Thank you for your participation!

If you have any questions after this session regarding this study, please contact me at 613-276-5751 or by email serdartuncer@cmail.carleton.ca. You may also contact my supervisor Prof. Prof. Alejandro Ramirez at Alex.Ramirez@carleton.ca or 613-520-2600 ext. 2397.
10 References


Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann,
I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A.
https://doi.org/10.1038/s41586-019-1138-y

https://doi.org/10.1145/3449081


75


Mathematical, Physical and Engineering Sciences, 376(2133), 20180085.
https://doi.org/10.1098/rsta.2018.0085


Conference on Fairness, Accountability, and Transparency, 295–305.

https://doi.org/10.1145/3351095.3372852