McMaster University, DeGroote School of Business

Ph.D. Dissertation:

"The Willingness to Collaborate with Artificial Intelligence (AI) in the Workplace: The Role of AI Autonomy and Explainability"

<u>Proposed by:</u>

Esraa Abdelhalim

Ph.D. Candidate,

Business Administration, Information Systems Area

abdelhae@mcmaster.ca

Supervisory Committee:

Khaled Hassanein

(Supervisor)

DeGroote School of Business

McMaster University

hassank@mcmaster.ca

Milena Head

(Supervisor)

DeGroote School of Business

McMaster University

headm@mcmaster.ca

Brian Detlor

(Committee Member)

DeGroote School of Business

McMaster University

detlorb@mcmaster.ca

Abstract

The world has recently witnessed rapid developments in Machine Learning (ML) algorithms, computing infrastructure, and the vast amounts of data available and accessible anytime and anywhere. This has facilitated a boom in the field of Artificial Intelligence (AI) in terms of the applications and benefits AI potentially brings to everyday life. AI is a technology that is able to interact, function on its own, explain, and learn from past experiences to inform its future actions and decisions. Generative AI is now changing the way we live, interact, and work, affecting all sectors of society. It is becoming mature enough that AI is transitioning from being a mere assistive technology to being an actor similar to humans with whom we can collaborate.

Many organizations, thus, seek to leverage AI and utilize its capabilities to achieve greater efficiency and effectiveness. Such organizations integrate AI in the workplace to handle many routine and repetitive tasks and free employees for more complex work. This does not mean that humans and AI work in isolation, however. Rather, it means that humans and AI can work together as collaborators to reach better decisions and overcome each other's weaknesses and deficiencies.

Although there is extensive literature on how new information technologies are adopted and accepted, there is little empirical work that studies how innovative technologies such as AI can become effective collaborators with humans in the workplace and the conditions under which humans are willing to collaborate with them. Therefore, this research proposes and empirically validates a new contextualized conceptual model that furthers our understanding of the factors that influence humans' willingness to collaborate with AI in organizational settings. Specifically, this work focuses on the role of AI autonomy and AI explainability (as AI contextual characteristics) in shaping people's beliefs about having AI as collaborators in the decision-making process. This study leverages the Actor-Network Theory and Net-Valence Theory as foundations to understand this phenomenon.

Perceiving AI as a collaborator in the workplace is a nascent phenomenon that has both concerns and benefits. Such concerns and benefits are not fully understood in the existing literature. Therefore, the proposed study employs a two-stage sequential mixed-methods approach to investigate this phenomenon. First, a qualitative study was conducted using one-on-one interviews to understand the top concerns and benefits of individuals in collaborating with AI in the workplace. Findings from the qualitative study were then used to fine-tune the proposed conceptual research model. The model was validated through a 2x2 factorial design scenario-based survey study using consistent Partial Least Squares (PLSc) as a Structural Equation Modeling (SEM) technique. Contributions to theory and practice are discussed, and study limitations and future work are outlined.

Acknowledgement

I would like to express my deepest gratitude to my thesis supervisors, Dr. Milena Head and Dr. Khaled Hassanein, whose unwavering guidance and support from day one were instrumental in the successful completion of this journey and reaching this milestone.

Dr. Head, your smile is a trademark. Whenever your name is mentioned, your smiley face is the first thing that comes to my mind. Your passion, positivity, and dedication inspired me in very different ways. Thank you from the bottom of my heart.

Dr. Hassanein, “thank you” will never be enough. Despite your busy schedule, you were always there to talk and help. Both your and Dr. Head’s invaluable insights, constructive feedback, and endless patience with me have not only shaped this thesis but also enriched my overall academic experience. I am truly fortunate to have had the opportunity to learn from both of you and be mentored by such dedicated and knowledgeable professors.

I extend my heartfelt thanks to Dr. Brian Detlor, whom I had the privilege to have on my supervisory committee and whose course and advice instilled in me a love for qualitative research. Your invaluable contributions, insightful suggestions, and critical evaluations helped refine and strengthen this research. Your collective expertise and diverse perspectives greatly enhanced the quality of this work, and I am grateful for the time and effort you invested in evaluating and improving my thesis.

To my mother and father, you have always encouraged me to work hard and strive to do my best. You always encouraged and helped me to pursue my dreams. Thank you for your guidance, love, and always being there to help and support. I owe a debt of gratitude that words cannot adequately express.

To my husband, thank you for being my partner in my personal and academic journeys. Your support and encouragement during the challenging times meant a lot to me. I cannot forget the days and nights you spent with our kids to help me complete this piece of my academic journey. Thank you for your invaluable support.

I would also like to acknowledge the support and encouragement of my three sisters, Marwa, Shaimaa, and Alyaa, who have always been the shoulder that I lean on. Being in my life, believing in my abilities, and your constant encouragement motivated me to persevere through the challenges and uncertainties I went through. I would like to especially thank Alyaa, whose presence in the Ph.D. program supported me in every possible way. I am incredibly lucky to have you in my corner, and I want you to know just how much your support has meant to me. Your kindness and understanding have been a source of strength and comfort.

Last but not least, I want to thank Deb Randall-Baldry, Bani Rafeh, and every staff member at McMaster University who facilitated things during my Ph.D. journey and who were the knowledgeable references I ran to when I needed help.

This thesis would not have been possible without the collective efforts and support of these individuals and many others who have influenced and guided me throughout my academic journey. Thank you all for your ongoing support, encouragement, and belief in my abilities.

Table of Contents

List of Tables

List of Figures

List of Key Terms

Term	Definition	Adapted From
AI	- The ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting, problem-solving, decision-making, and even demonstrating creativity	(Rai et al. 2019, p.iii)
Machine Learning algorithms	- Programmed algorithms that receive and analyze input data to predict output values within an acceptable range. As new data is fed to these algorithms, they learn and optimize their operations to improve performance, developing 'intelligence' over time.	(Wakefield 2023)
Social Actor	- An entity exhibiting social norms and is capable of agentic communication with others.	(Gambino et al. 2020; Nass et al. 1994)
Collaboration	- Having two or more social actors engage in a joint activity to achieve a shared goal.	(Bedwell et al. 2012; D. Wang et al. 2020)
AI Collaborator	- An AI that can perform cognitive functions that we normally associate with human minds, can work autonomously, can interact with and learn from humans, can adapt to	(Rai et al. 2019, p.iii)

	different situations, and can make proactive, predictive, or personalized decisions.	
Willingness to Collaborate	- The readiness of a human to accept and jointly work with an AI as a collaborator to complete a task	(Rosas and Camarinha-Matos 2010)
Autonomy	- The extent to which an AI collaborator works independently and makes decisions on its own without human intervention.	(Nickerson and Reilly 2004)
Explainability	- The extent to which an AI collaborator provides explanations about its recommendations/decisions.	Self-defined
Efficiency	- The ability to save time and speed up the decision-making process.	Self-defined
Integrity	- The extent to which an AI's recommendation is perceived as being objective, unbiased, and fair.	(Przegalinska et al. 2019; Whang and Im 2018)
Lack of Human Interaction	- The concern of losing human interaction when collaborating with AI.	(Dabholkar and Bagozzi 2002)
Incompatibility	- The concern that the AI might be inconsistent with and is unable to provide recommendations that align with an organization's and/or humans' needs and values.	(Carter and Bélanger 2005)

Chapter 1: Introduction

In one of the most famous dog cartoons from the early internet age in 1993, two dogs were standing in front of an internet-connected computer. The dog who was operating the computer said to the other dog, "*On the Internet, nobody knows you're a dog!*."[1] Since then, this cartoon has been referenced whenever issues related to anonymity and identification are discussed in computer-mediated communication. However, when we think about the statement that the dog shared with his companion, we can also conclude that it does not only tackle anonymity issues. It challenges a taken-for-granted assumption that users of internet-connected computers at that time always assumed that the other party at the other end with whom they interact is another human like themselves and cannot be something else.

Today, with all the technological advancements in ML algorithms[2] and the evolution of intelligent machines that can imitate humans, this presumption has changed. As Norbert Wiener, the father of the modern theory of cybernetics, predicted in his book "*The Human Use of Human Beings*" that interactions in the future will not only be between humans but will be extended to include humans with machines and machines with machines (Wiener 1950).

We can now see how Artificial Intelligence and smart technologies like robots and chatbots have revolutionized the world around us. There is no one unified definition of AI. In the online-version Dictionary of Merriam Webster, AI is defined as "*the capability of machines to imitate intelligent human behaviour*" (Merriam Webster 2023). Another definition of AI that I utilized in

[1] Peter Steiner, "Dog Cartoon," The New Yorker (1993): 61

2 ML algorithms in this study are defined as "Programmed algorithms that receive and analyze input data to predict output values within an acceptable range. As new data is fed to these algorithms, they learn and optimize their operations to improve performance, developing 'intelligence' over time" (Wakefield 2023)

this study is "*the ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting, problem-solving, decision-making, and even demonstrating creativity*." (Rai et al. 2019, p.iii). AI affects the quality of our lives, from the time we wake up to the moment we sleep. There is AI in our navigation apps, the products and services recommended for us, keeping spam out of inboxes, monitoring our investments, detecting fraudulent transactions, and much more. In addition, Generative AI is now changing the way we live, interact, and work, affecting all sectors of society (Hacker et al. 2023). For instance, ChatGPT (i.e., an AI developed by OpenAI) was the hottest topic at the end of year 2022. GPT stands for Generative Pre-trained Transformer Chat, and GPT-3 can generate text and process any text or natural language commands without the need to modify or fine-tune them (Teubner et al. 2023). It was trained on more than 570 GB of data, such as books, blogs, articles, and others (i.e., around 300 billion words in total) (Hughes 2023; Teubner et al. 2023). These applications employ both simple and complex ML algorithms that help AI learn and be smarter over time. AI is working behind the scenes with or without our consciousness. Experts predict that by 2030, AI will add around $15.7 trillion to the global economy (de Cremer and Kasparov 2021).

ML algorithms that learn by example allow machines and humans to understand and interact with each other. The evolution of such technologies has opened the opportunity for new business capabilities such as automation of repetitive tasks, engagement with customers and employees, making decisions, and innovation (Benbya et al. 2021). For example, natural language processing capabilities enable chatbots to chat with patients and provide basic diagnoses like real doctors. In customer service, chatbots help agents focus on more complicated queries and tasks by answering simple systematic questions in real-time to improve customer experience.

Developments in speech-to-text technology enable us to speak to IBM Watson, Alexa, and Siri, and they can talk back to us.

Computer vision is another type of AI that empowers self-driving cars to find their way on the streets and avoid collisions with objects and people. Computer vision algorithms can also detect facial expressions and process images to authenticate identities through facial recognition, detect people and identify criminals in videos, diagnose cancer cells in skin images or find abnormalities in medical scans. AI can also analyze historical data to make informed decisions, such as approving loan requests, trading in the stock market, or recommending people for certain jobs or promotions. AI also plays a role in innovation (Benbya et al. 2021). Even though AI cannot yet develop complete innovative solutions on its own, AI can expand the search space of existing knowledge and offer a creative interpretation of data and recommendations that would support the innovation process (Wu et al. 2020). Statistics concerning internet traffic reported that, in 2021, non-humans (bots, hacking tools, etc.) made up 42.3% of the web traffic.[3]

Although AI dates back to the 1950s, AI boomed again in the 1990s and the 21st century for three primary reasons: 1) the big data that is now available and accessible, 2) the advances in ML algorithms, 3) and the enhanced computational power that is now cheaper and affordable. What is different about AI is that these systems can understand unstructured data, can reason to form hypotheses, can learn and get better over time, can have access to vast amounts of data that is beyond the ability of humans, can contextualize it, can interact with humans, and can function on their own (Murray et al. 2020). Traditional computing systems can capture, store, and process

---

[3] San Mateo,"42.3% of Internet Traffic in 2021 Wasn't Human As Account Takeover and Online Fraud Increases" Business Wire, January 11th 2023, https://www.businesswire.com/news/home/20220518005342/en/42.3-of-Internet-Traffic-in-2021-Wasn%E2%80%99t-Human-As-Account-Takeover-and-Online-Fraud-Increases

unstructured data, but they cannot understand it. They cannot reason contexts and situations and cannot learn from past experiences and interactions to get better. There is “X” and with AI, we have “AI + X”. This concept is applied in the “Soul Machines” project, which they name “Digital People.” The idea is that the technology is there, but if we add AI to the technology, it will enable it to interact and express itself in a human-way.

This situates AI in a great position to collaborate with humans and augment humans' intelligence for improved efficiency and effectiveness. Industry 5.0 examines how humans and intelligent machines can complement each other's unique capabilities rather than replace one another (Welfare et al. 2019). At the World Economic Forum in 2017, IBM CEO Ginni Rometty said. "*For most of our businesses and companies, it will not be man or machine... it will be a symbiotic relationship. Our purpose is to augment and really be in service of what humans do*." (IBM 2017). For that reason, IBM refers to AI systems as cognitive systems rather than just AI (Sommer 2017).

AI entails the use of advanced ML algorithms and software programs that approximate human cognition and reasoning to interact with humans, autonomously work on tasks, analyze massive data, and learn from past experiences to inform future actions and decisions. These technologies can now digest information from diverse sources instantly and record every bit of information they come across to learn and augment their knowledge base over time. Despite the superior performance of AI in some well-defined tasks such as identifying objects, discovering patterns, or playing games, we are still far away from the development of an AI that can solve different tasks at the same time (Dellermann et al. 2019) or tasks that are deemed to be complex, requiring human judgment. As a consequence, Dellermann et al. (2019) introduced the term “Hybrid Intelligence,”; which the authors define as “*the ability to accomplish complex goals by* *combining human and artificial intelligence to collectively achieve superior results than each of them could have done in separation and continuously improve by learning from each other*" (Dellermann et al. 2019, p.3). Furthermore, AI technologies are now designed with human-like attributes (e.g., body, face, voice) to increase their acceptance as social actors.

Early research on media and computers uncovered that people attribute social norms to such technologies and tend to treat them like real people (Nass et al. 1994; Nass and Moon 2000). People seem to perceive computers as social actors that are capable of agentic communication with human users (Gambino et al. 2020; Nass et al. 1994). In an organizational context, social actors seek to communicate with others in socially legitimated ways (Lamb and Kling 2003). Social actors "*engage in social intercourse as a collectivity and possessing rights and responsibilities as if the collectivity were a single individual*" (Whetten and Mackey 2002, p.395). Hence, in this research, I define social actors as "*an entity exhibiting social norms and is capable of agentic communication with others*".

Traditional, non-intelligent computing systems, on the other hand, lack the capability to reason through bodies of knowledge, interact with humans, make and explain decisions, function on their own, and learn from the interaction. They do not even possess any human appeal or human-like mental models. AI technologies are thus different from traditional computer programs and are now transitioning from being a mere technology to actors that can interact with humans (Schuetz and Venkatesh 2020) and collaborate (Sowa et al. 2021). The evolution of such technologies with human-like and technology-like characteristics encouraged organizations to incorporate AI in the workplace with whom humans can collaborate.

D. Wang et al. (2020) suggest that most human tasks and activities nowadays are accomplished collaboratively, and it is important to understand how to incorporate AI as collaborators into the already-complicated human workflow and to have a plan for a Human-AI Collaboration future of work (p. 1). In organizational behaviour, collaboration involves having two or more stakeholders collaborate in an effort to solve a set of problems that neither can solve individually (Bedwell et al. 2012). It occurs when a group of autonomous stakeholders of a certain problem context engage in an interactive activity, using shared rules, norms, and structures, to act or decide on issues related to that context (Bedwell et al. 2012). When collaborating, stakeholders bring different resources to bear on the issue at hand. Thus, collaboration involves mutual goal understanding, preemptive task co-management and shared progress tracking (D. Wang et al. 2020). Consequently, collaboration in this research is defined as " *having two or more social actors engage in a joint activity to achieve a shared goal.*"

AI is arguably considered the technology that will have the most influential impact on team outcomes (Seeber et al. 2018). This is in line with what Norman (2017) stated: "*[a]s automation and artificial intelligence technologies develop, we need to think less about human-machine interfaces and more about human-machine teamwork*" (Norman 2017, p.26). That said, integrating AI as collaborators with humans in the workplace to jointly work on tasks may also raise a debate about whether these smart technologies will replace humans in the future. This belief assumes that humans and AI are independent of each other, such that each works in isolation. However, thousands of new jobs will be created to develop new AI products and to ensure they work properly (Wilson and Daugherty 2018). Additionally, AI is not completely replacing humans but complementing and augmenting them in the workplace. For example, AI technologies cannot yet share feelings and emotions or sense others' intentions and expectations. They do not feel tired or bored and still have limited creative abilities. They cannot generate new ideas for new business models or solve new problems outside their domain of expertise without human input (e.g., training an AI and exercising judgment and moral values) (Claudé and Combe 2018). Benbya et al. (2021) differentiate between "automation" and "augmentation." They refer to automation as "*tasks that are performed by a machine without any human involvement*." In contrast, augmentation is the "*continuous close interaction between humans and machines, with machines learning from humans via training datasets and humans learning from the insights gained through machines*" (Benbya et al. 2021, p. 285). Humans and AI have complementary strengths, and collaboration can augment these strengths. What is natural to humans, such as telling a joke, might be challenging for machines, and what is easy for machines, such as analyzing huge datasets in seconds, might be beyond humans' capabilities.

In a study that involved 1,500 firms, Wilson and Daugherty (2018) find that organizations realize the most exceptional performance improvements when humans and machines collaborate. Rai et al. (2019) advocate that hybrid Human-AI collaboration can align the capabilities of AI (such as speed, accuracy, scalability, and reliability) with the strengths of human agents (such as creativity, judgment, and empathy) to yield better outcomes. So, the beauty of Human-AI collaboration is that humans will handle what they can do best, and machines will handle what they can do best (Daugherty and Wilson 2018). Some Swedish banks, for example, currently use AI virtual customer service assistants, referring to them as their "newest employees" and even giving them real names such as "Aida" or "Nina" (Rai et al. 2019). These AI assistants allow human employees to work on other tasks.

AI can be helpful collaborators in multiple contexts. For example, they can help doctors in hospitals and medical diagnostics and support human resources professionals in interviewing and recruiting candidates. In the latter context, AI can capture facial expressions, analyze gestures, and skim large amounts of resumes in seconds to determine which candidates are the best fit for a particular job. A recent survey found that 56% of candidates who experienced discrimination in hiring expect AI to be less biased than human recruiters, and 49% believe that AI might increase their chances of getting a job (Modern Hire 2018). There are a number of available AI solutions that are currently being used to help recruiters in hiring and interviewing people, such as Knockri, Seedlink, Gecko, Spark Hire, and HireVue. For example, Unilever (a consumer-goods company) uses an AI called "HireVue" to interview approximately 100,000 candidates in the UK (Hymas 2019). With this AI, the company follows a three-stage recruitment process. First, it asks applicants to play an online game that assesses qualities that might suit particular positions. Second, candidates are asked to record a video in which they answer a set of questions related to the job position they are interested in. The software analyzes the facial expressions, the tonality, and the speed at which a candidate speaks to assess whether they match the positions they are applying for. For instance, if they speak too slowly or quickly, they might be unsuitable to work at a call center or a job where they interact with consumers. Once the process is complete, human recruiters reach out to the selected pool of candidates to invite them for a personal interview in order to make a final decision.

Based on the foregoing discussion, there is an evolving interplay between humans and AI in the workplace that requires further investigation. Gunkel (2012) suggested that to address the evolving challenges pertaining to interacting with intelligent machines and autonomous decision-making systems, researchers have to shift their focus from studying computer-mediated communication to human-machine communication. Schuetz and Venkatesh (2020) argue that developing AI with human-like capabilities challenges five main assumptions that Information Systems (IS) researchers have held for decades about users' perceptions and behaviour when interacting with Information Technology (IT) artifacts. These pre-held assumptions include: (a) the relationship between users and IS artifacts is unilateral; (b) IS artifacts are ignorant of and isolated from their environments; (c) IS artifacts are directed by their software owners and developers and ignore the fact that new intelligent systems are adaptive; (d) the results derived from IS artifacts are derived from deterministic (i.e., if-then clauses) rather than probabilistic and complex statistical models (e.g., neural networks) – hence, such artifacts can incorporate many contextual factors without the knowledge of developers and users; and (e) users are assumed to be aware of using the IS artifacts since they do not possess any human-like attributes as opposed to intelligent cognitive systems like AI that disguise the feeling of interacting with a non-human agent. Schuetz and Venkatesh (2020) recommend that research in this area is limited and that the five assumptions discussed above hinder the applicability of the current body of knowledge of the IS literature to the new emergent intelligent artifacts.

Most of the existing empirical work focuses on examining the factors that influence people's intentions to use or trust a technology from the perspective that they operate and work in isolation from humans. The extant literature also examines how technology can be a facilitator in traditional and virtual human-to-human collaboration. Baird and Maruping (2021) raise the concern that IS scholars, and theorizing of IS use tend to look at IS artifacts as *"passive tools"* by putting an emphasis on "*human agency*" rather than on *"IS artifact agency"* in the user–IS artifact relationship. Nonetheless, with the evolution of ML algorithms, cloud computing, and intelligent technologies, IS artifacts are no longer passive tools to achieve a certain goal. Rather, they can now take control and carry out tasks independently (Libert et al. 2020). They can even provide justification and explanation for their actions or decisions. Ignoring the agentic role of such IS artifacts provides an incomplete picture of their capabilities to achieve goals (Baird and Maruping 2021).

Therefore, this study is an attempt to explore the Human-AI symbiosis phenomena in organizational settings. The main objective of this research is to understand the factors that are likely to influence the willingness of humans to collaborate with AI in the workplace. The willingness to collaborate in this study is defined as "*the readiness of a human to accept and jointly work with an AI as a collaborator to complete a task*". In doing so, it is important to understand how individuals make a trade-off between the positive utilities (i.e., benefits) and the negative utilities (i.e., concerns) from collaborating with AI. In general, when humans contemplate integrating new technology into their work or life, they follow a "net valence" approach by weighing both the benefits and concerns of incorporating that particular technology (Breward et al. 2017; Cazier et al. 2008; Dinev and Hart 2006). Such benefits and concerns might vary by context and by the characteristics of the AI.

This research is, thus, novel as it views AI as a social actor that shares agency with humans and not just a mere piece of passive technology (Schuetz and Venkatesh 2020). It is also unique as it aims to understand individuals' beliefs regarding the perceived benefits and concerns of collaborating with AI. Hence, the main research question of this study is:

"What are the factors that influence humans' willingness to collaborate with AI in the workplace?"

Thesis Outline

The rest of this thesis is organized as follows. *Chapter 2* outlines a detailed elaboration of the body of knowledge relevant to the various aspects of this study. *Chapter 3* discusses the theoretical premise of this research. *Chapter 4* outlines a qualitative study that explores and identifies the benefits and concerns people have when collaborating with AI. Based on the results of this qualitative study, *Chapter 5* presents a research model and associated hypotheses. *Chapter*

6 describes a quantitative methodology to validate this research model and presents data analysis results. *Chapter 7* provides a detailed discussion of the study's results and elaborates on the practical and theoretical contributions of the research and its limitations.

Chapter 2: Literature Review

2.1. The AI Evolution

In 1950, Alan Turing introduced the challenging Turing test: "*Can computers communicate in a way that persuades a human that they are communicating with another human as well?"*(Negnevitsky 2005). Following this, Marvin Lee Minsky co-founded an Artificial Intelligence Laboratory at the Massachusetts Institute of Technology (MIT). The first academic AI conference took place at Dartmouth College in 1955. In the same year, researchers at Carnegie Mellon University (previously the Carnegie Institute of Technology) created the first AI program called "*Logic Theorist*" – the first automated reasoning program that was called "the first artificial intelligence program" (Gugerty 2006). AI continued to evolve and led to the development of "Expert Systems" (i.e., software programs created with a database that are fed with a predetermined set of rules to mimic expert knowledge and recommend solutions to problems) (Negnevitsky 2005). Afterwards, Frank Rosenblatt introduced the notion of the Artificial Neural Network (ANN) that mimics how a neural system works in the human brain (Rosenblatt 1959). Rosenblatt introduced a digital "perceptron" that could detect images and identify whether, for example, they are men or women. Later, Geoffrey Hinton, argued that Rosenblatt's ANN consisted of only one layer, while a neural network consists of more than one layer: an input layer, output layer, and hidden layers (Hinton 1991). However, due to the limited computational power and capabilities, AI development faced a period of stagnation known as the "AI winter" (Hendler 2008).

As discussed earlier, AI boomed again in the 1990s and the 21st century. Searching the Web of Science using the keyword "Artificial Intelligence" alone for the period from 2015-2021 yielded an exponential increase in research totaling 235,000 publications (see Figure 1). Figure 2 shows that the top 10 research areas that tackled AI topics during this period were purely technical, with “Computer Science” dominating these areas.

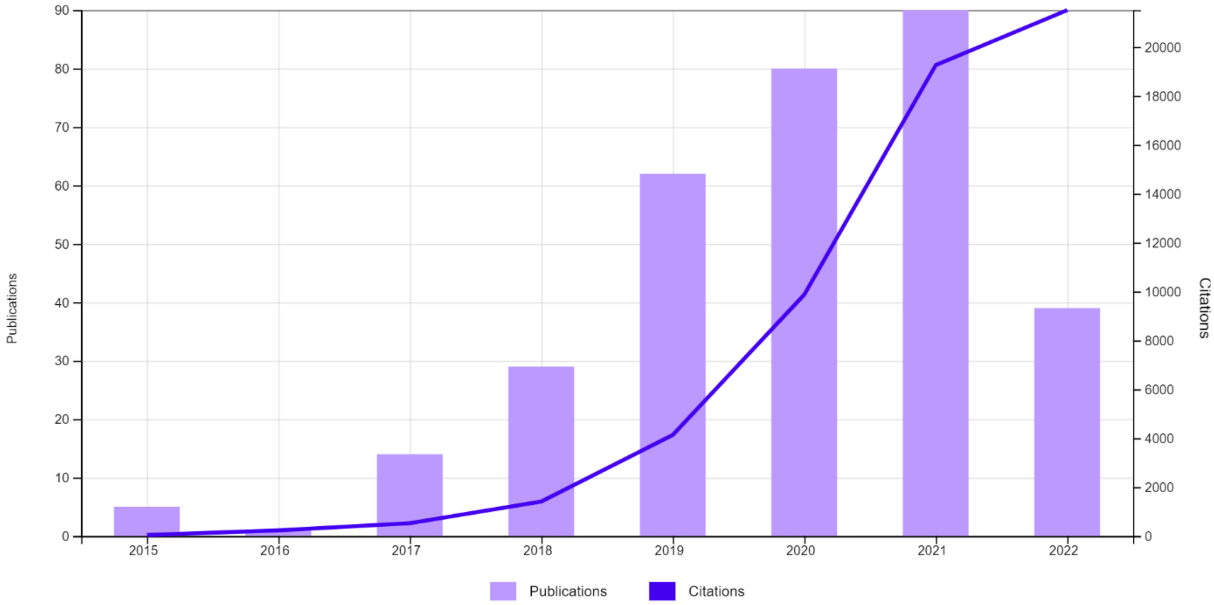


Figure 1: Number of Publications on Web of Science on the AI Topic from 2015-2022

Figure 2: Classifying AI Research by Research Area

However, when searching the Web of Science using the "Human-AI" keyword, results returned only 524 publications, with the "Computer Science" domain dominating other fields as well (see Figure 3). This asserts that the information systems literature lacks sufficient empirical research on Human-AI collaboration and understanding humans' perception when collaborating with AI as a social actor.

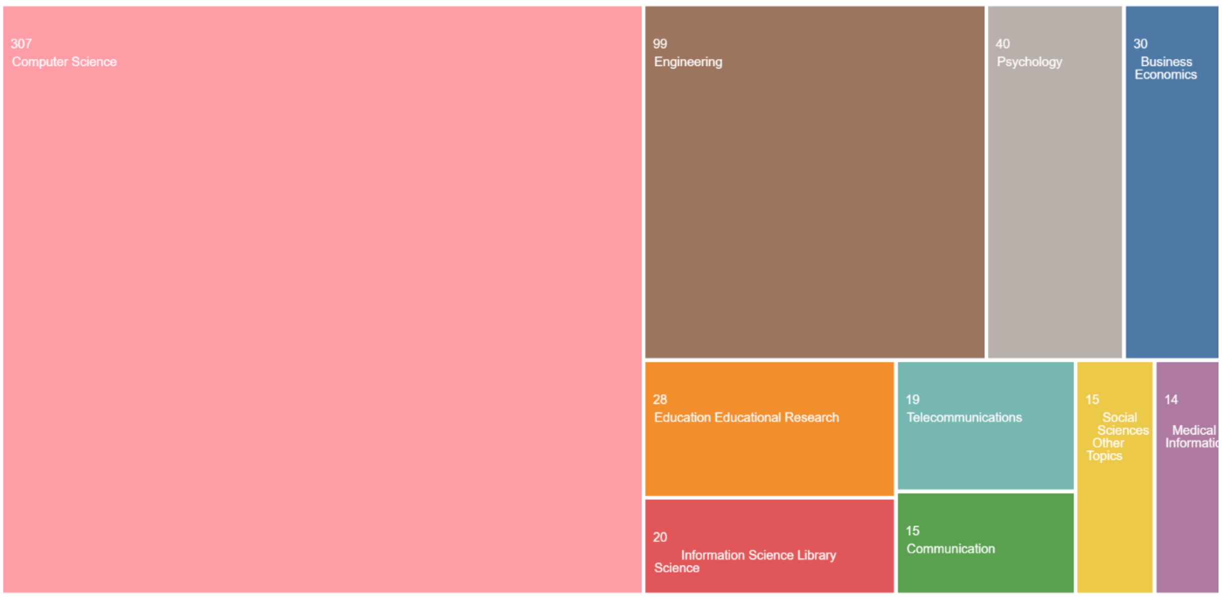


Figure 3: Classifying Human-AI Research by Research Area

Schuetz and Venkatesh (2020) explained the progression of intelligent machines' capabilities from having a mere Decision Support System (DSS) to building progressively more intelligent cognitive systems, namely Expert Systems (ES), Intelligent Agents (IA), and Cognitive Computing Systems (CSS), that are able to act and perceive like humans (see Figure 4).

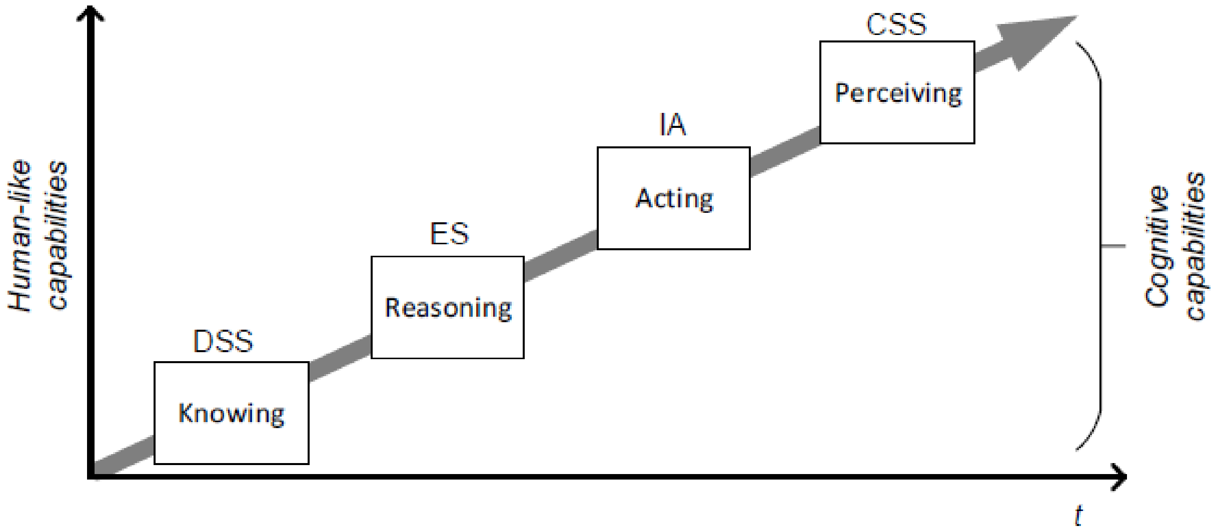


Figure 4: Progression Phases of Machine Capabilities (Schuetz and Venkatesh 2020)

Figure 5 illustrates that DSSs were the first step in the ladder of developing cognitive intelligent agents/machines. DSS are systems that operate through a set of decision rules and models and use a large database (Turban and Watkins 1986). DSSs enable users and decision-makers to generate aggregated reports or charts that can aid in the decision-making process. Still, it is the decision-maker's choice to pull out inferences from that generated information. Next comes ES that rely on a set of predetermined rules built into a knowledge base to reason like human experts and can provide some explanations for their recommendations (Turban and Watkins 1986). The third phase included the evolution of IA, that were empowered to operate autonomously without being reliant on humans to complete certain tasks (Turban and Watkins 1986). Such as utilizing IAs to detect malicious intentions of people who cross country borders (Nunamaker et al. 2011) or placing bids in auctions (Adomavicius et al. 2009). However, all of these levels of machine capabilities relied on a structured type of data to carry out tasks, making it difficult for humans to interact with it. Therefore, in the last "perceiving" phase, cognitive machines are introduced as machines that have the power to perceive like humans and make sense of the structured and unstructured environment surrounding them. They can eventually learn to become better over time and might improve themselves until reaching what's known as "the singularity" (i.e., a point when advancement in technology becomes uncontrollable) (Hutson 2023).

This evolution and progression of machine capabilities are attributed to three primary reasons. First, the growth of the big data era provided an opportunity to gain deeper and more informed insights about diverse issues/problems that previously were not possible. It is no longer limited to the structured type of data that is stored in a row-column format in databases. Now we also have semi-structured (e.g., XML) and unstructured data types (e.g., photos, videos) that are captured every second. Second, computing infrastructure and fast microprocessors have become more affordable and accessible. For example, cloud computing allows people to store, retrieve, and use data that is not stored on their personal devices to perform transactions or operations anytime and anywhere. Third, advances in ML algorithms and deep learning have made tremendous contributions to increasing operational efficiency, as well as creating machines and platforms that are more intelligent than ever (Daugherty and Wilson 2018). Hence, machines powered by AI can recognize and learn from complex patterns, draw conclusions, and predict future trends. Some AI have even exceeded human performance in identifying objects (Bughin et al. 2017).

AI can be classified into different evolutionary stages: 1) Artificial Narrow Intelligence (ANI), 2) Artificial General Intelligence (AGI), and 3) Artificial Super-Intelligence (ASI) (Joshi 2019; Kaplan and Haenlein 2019). ANI is also referred to as "Weak AI" because this type of AI is able to perform only one task. For example, DeepBlue created by IBM, was able to defeat human chess champions. However, we cannot ask DeepBlue to tell us the shortest route to get home or to predict the weather. ANI is the most common type that exists nowadays. Many of the AI that we use on a daily basis are also types of ANI, where it is intelligent and smart to do only one specific task. Hence, we can find ANI-powered applications in many domains and contexts. AGI, also referred to as “Human-Level AI”, or “Strong AI”, can be applied to different areas and is able to autonomously solve problems in different domains, such as IBM Watson. Finally, ASI, also called “Self-aware AI”, is a type of AI that can autonomously work and intelligently solve any problem in any domain or context and can outperform humans. Although self-aware AI does not yet exist (Kaplan and Haenlein 2019), great efforts are being devoted in this direction where we could see the influence that generative AI had in late 2022 when ChatGPT-3 was released to the public (Teubner et al. 2023).

Generative AI has emerged as a groundbreaking technology with the potential to transform various industries, including art, design, and content creation. The advancements in generative AI have led to the development of language models that can generate coherent and contextually relevant text (Hacker et al. 2023). OpenAI's GPT have garnered attention for their impressive text-generation capabilities. They are trained on vast amounts of text data, enabling them to capture syntactic and semantic patterns. These models can generate human-like text in a variety of styles, from news articles to poetry, by predicting the next word in a sequence given the preceding context. GPT-3, the subsequent iteration, boasts even more parameters and has demonstrated the ability to perform various tasks, including translation, summarization, and code generation (Hutson 2023). The widespread adoption and exploration of generative language models showcase their potential to automate content creation, streamline communication, and assist in a multitude of natural language processing tasks.

2.2. AI in Different Contexts

The idea of automating and augmenting humans' capabilities using intelligent or robotic machines is not new. In the manufacturing industry, several manual tasks were automated to aid humans and speed up the manufacturing process. For example, (Bright and Harry Asada 2017) designed a robotic arm that a human worker can wear to assist humans with completing heavy tasks and improve their productivity and safety. The robotic arm can lift and hold an item until the wearer secures it using a tool with both hands. Amazon is an excellent example of a company that automated numerous operations and departments for improved efficiency and effectiveness. For example, it uses a variety of robots that can assist in picking, moving, and packaging products (Correll et al. 2018; Laber et al. 2020).

Many applications and platforms powered by AI are also in the market and used by different organizations nowadays. For example, the Toronto Dominion (TD) Bank has acquired "layer6 AI" (a leading AI platform for prediction, personalization, and recommendation) to offer personalized banking and to anticipate customers' needs (Ligaya 2018). TELUS also introduced its AI virtual assistant for a 24/7 service to improve customer experience and to help navigate needs, offload call center demand, and create and implement frictionless solutions (Las 2018). In the recruitment domain, "Plum.io" is a hiring platform that matches human potential to the right job. Plum.io helps mid-size and enterprise companies solve tough talent acquisition, talent management, and workforce planning challenges while minimizing recruiter bias at the same time (Wiggers 2018). "Seeing AI" is another AI introduced by Microsoft. This free application is an ongoing research project designed to open up the visual world for the blind and visually-impaired community by harnessing the power of AI to describe nearby people, text, and objects to users (Coldewey 2019).

Another interesting example is the case of Humber River Hospital (HRH), nominated in 2016 as North America's first all-digital hospital (Kutscher 2016). In HRH, robots help employees by sorting medications, providing lunch trays using automated guided vehicles, and delivering test tubes that carry blood samples from patient floors to the laboratory (Kutscher 2016). The hospital also uses “da Vinci Robotic Surgery” which enables a surgeon to perform minimally invasive procedures by using high-end technology to guide precise and flexible surgical instruments. HRH also employs a robot called “Pepper” which is Canada’s first emotionally- sensitive robot for sick children, designed to reduce their anxiety by taking them through the hospital and explaining what to expect during their surgery (Adam 2018; Kiew 2018).

Intelligent chatbots can also be utilized to promote mental healthcare. Wysa and SERMO are two examples of intelligent chatbots used to assist people with mental health problems. Wysa is an emotionally-intelligent chatbot that is able to track a person’s mood and can detect whether their mood is negative or positive (Inkster et al. 2018). Based on the mood Wysa detects, it reacts in a personalized way and may suggest taking some tests (e.g., depression test, anxiety test…etc). Based on the test results, Wysa can provide further guidance or recommend seeking help from a professional. The chatbot also has meditation exercises that can help reduce anxious feelings or stress. Inkster et al. (2018) tested the chatbot with 129 participants. Some participants were frequent users of Wysa, and some used the app occasionally. Findings showed that frequent users of Wysa had a greater improvement in their mood than the group of occasional users, and two-thirds of the users indicated that conversing with Wysa was helpful for mental health (Inkster et al. 2018).

SERMO is another conversational agent for mentally ill people that is able to regulate feelings and thoughts (Denecke et al. 2021). It is based on the ABC theory developed by Albert

Ellis, where (A) represents an "Activating" event or situation that could be internal or external. (B) represents "Beliefs" consisting of behaviour and thoughts regarding the event. Finally, (C) represents "Consequences" reflecting feelings and attitudes. The theory suggests that when an individual perceives a stimulus, whether consciously or unconsciously, this stimulus is evaluated, which may lead to certain feelings and behaviours. SERMO asks its users about events and situations that occur to them on a daily basis and creates an emotion diary and a list of pleasant activities. Then, it employs natural language processing techniques to analyze responses and provide the appropriate treatment and necessary exercises for each user. For instance, if SERMO detects an anger emotion, it asks the user whether the anger is justified or not. If the user indicates that it is justified anger, SERMO asks the user if they would like to change the situation. If the user's anger is unjustified, SERMO follows a different approach and tries to focus the user's attention on something different by asking positive questions such as "What are you proud of?", "What do you like doing in your spare time?" (Denecke et al. 2021). Denecke et al. (2021) tested SERMO with 21 users and found that efficiency, perspicuity and attractiveness are good features that SERMO supports and that users enjoy. Expert psychologists and psychotherapists corroborated that SERMO is effective, especially for people who may not be willing to express their feelings and emotions in a face-to-face mode. Experts agree that a chatbot like SERMO is able to bridge this gap.

In examining the impact of introducing an AI in creative tasks, (Lysyakov and Viswanathan 2022) examined how the launch of an AI contestant for a logo design contest affected humans in a decentralized crowdsourcing platform. The authors found that AI could have different effects on humans with different capabilities. For example, humans with the lowest capabilities (i.e., in terms of the complexity and emotional content of their design submissions) are expected to avoid competing against an AI by exiting the platform. In contrast, designers with higher capabilities tend to accept the challenge and leverage their capabilities (e.g., imagination, creativity, emotional content…etc) after the launch of the AI.

During COVID-19 times, AI proved to be a great success when face-to-face communication between humans was advised to be avoided or controlled (Judson et al. 2020). For example, Laguarta et al. (2020) trained an MIT open voice model, where AI was used to detect and discriminate COVID-19 patients through audio cough recordings. The authors collected data from 5,320 subjects and created the largest audio COVID-19 cough-balanced dataset. Chatbots were also one of the most common AI that almost every organization has directed its attention to during the pandemic. It helped in communicating with customers during high volume requests anytime and anywhere to cope with the physical distancing requirement while ensuring smooth delivery of services in order not to lose customers. Hari et al. (2021) used the diffusion of innovation theory to study the factors that would influence customers' satisfaction and brand usage intention when interacting with banking chatbots. Interactivity with chatbots, compatibility with customers' needs, and trialability of chatbots did have an influence on customer brand engagement, which in turn affected satisfaction with the chatbot experience and the brand usage intention.

Han et al. (2022) investigated whether deploying emotional chatbots that can express positive feelings when interacting with customers would impact customer service evaluations. Varying three different levels of emotions through a laboratory experiment, the authors found that employing chatbots that show positive emotions to customers is not always beneficial to customers' expectations and may not always result in yielding enhanced customer evaluations. It was evidenced in their study that expressing positive emotions from a human agent had a better impact on customers than when the emotions were expressed by an AI. Interestingly, they also found that it is beneficial that an AI expresses positive emotions in a communal relationship with customers rather than an exchange relationship. This means that sometimes, it is better to account for the contextualization of the task when an AI is being introduced (Han et al. 2022).

2.3. AI Explainability

Organizations that have members with diverse backgrounds and expertise are expected to create a powerful synergy and perform better (Horwitz 2005; Rock and Grant 2016). However, in human-human groups, exchanging knowledge and explanations among members could be challenging. Human members might be hesitant to exchange their knowledge with others or to provide an explanation or justification for the decisions they make. This could be for several reasons, such as the fear of losing ownership or a position of power (Szulanski 1996) or fear that their job would be taken over by someone else. Members, therefore, make a tradeoff between the benefits they may gain from exchanging their knowledge and the associated costs. To encourage information exchange, some organizations may reward individuals who share their knowledge while punishing others who refrain from doing so (Bartol and Srivastava 2002). Firms also develop online platforms that enable them to reach subject matter experts from around the world to solve their problems (Dissanayake et al. 2015).

Likewise, in traditional and virtual teams, having members who can exchange their knowledge and explain their behaviour is critical to teams' success. Members who are knowledgeable about the task at hand can contribute better to team performance (Alsharo et al. 2017; Gardner et al. 2012). Therefore, organizations are now directing efforts to integrate AI as coworkers in the workplace who can complement employees' deficiencies and augment their knowledge-base and effectiveness. Since AI has a vast amount of extensible knowledge embedded within it and can be programmed to provide an explanation for their decisions and recommendations, this exchange of knowledge in a social context can be seen through the lens of the Social Exchange Theory.

Social Exchange Theory posits that people exchange favours in expectation of some future but unclear returns (Emerson 1976). This theoretical approach defines a group of social actors as two or more humans whose interactions affect their behaviours and actions. They usually behave in ways that maximize their benefits and minimize their costs (Alsharo et al. 2017). Posard and Gordon Rinderknecht (2015) have extended this definition by identifying a group as one that involves both humans and AI computers. The main benefit acquired from exchanging knowledge with AI is to achieve effective collaboration. However, human members might be uncertain about sharing their unique knowledge and paying the cost of losing ownership of it. On the other hand, an AI would not hesitate to provide an explanation for any decision or action it takes as it does not fear losing its job or its position of power. Organizations may view this as an attractive advantage of adopting AI as a collaborator in their workplace.

To develop such powerful AI, many ML algorithms are widely used by many organizations to foster competitiveness (Behl 2020). These AI can handle data of big sizes and may utilize different ML algorithms to provide explanations and make more informed decisions. Various ML algorithms such as decision trees, neural networks, logistic regression, and others can be used to serve different contexts and tasks (e.g., image classification, loan prediction, candidate selection…etc) (Burrell 2016). However, many ML algorithms embedded in AI remain non-transparent and inexplainable (Rai 2020). They are deemed as black boxes to decision-makers (Azodi et al. 2020). We have little understanding of how and why AI makes certain decisions or produces particular recommendations. In this context, we define a black box model as *"a model whose outcomes are hard to explain to decision-makers since its internal reasoning is unknown* *or hidden from its users"*. It could be due to the ML algorithm being proprietary, complex (Burrell 2016), or highly technical. The decision-maker may not know how the ML algorithm works or how representative the data is that was used to train the model.

Improperly trained AI might result in what is called a "Bias in, Bias out" problem (Mayson 2018), leading to favouring some groups over others or excluding some minorities (Ajunwa 2020; Rai 2020). Therefore, providing proper explanations to decision-makers and users is essential as a way to understand the reasoning behind a certain decision. Danks and London (2017) argued that non-transparent, inscrutable systems might have bias not only in the outcomes but also in the decision-making process itself. Although such systems may be perceived as impartial, they might employ biased algorithms that go unnoticed and uncorrected until it is perhaps too late (Danks and London 2017), which can significantly hamper users' trust in these systems and their willingness to rely on them. Guidotti et al. (2018) discussed that decisions generated by black box models without proper explanations might have counterintuitive effects and may create unconscious discrimination.

Furthermore, some regulatory agencies are emphasizing the need for transparency in ML algorithms to the extent that explanations are becoming more compulsory than voluntary (Krafft et al. 2022; Peukert et al. 2022). Explainability is essential for various reasons (Freitas 2014): first, it is argued that an explainable model is crucial for users to trust and accept its outcomes (Gilpin et al. 2018; Symeonidis et al. 2009). For example, Shin (2021) invited 350 subjects to study the impact of AI explainability on trusting the AI and the attitude towards it. Participants used a media lab equipped with computers and were asked to browse, view, and read automatically generated news on algorithm-based sites for about 1–2 hours. Subjects were presented with an explanation of why the AI recommended certain content to them and that it was based on the subject's preferences for some aspects. Results showed that explainability did have a significant effect on subjects' trust and attitude towards AI.

Second, users and decision-makers should be empowered by the ability to explain the rationale behind the recommendations that AI provides. Third, decision-makers would be able to elicit hidden insights about important relationships in the data and improve the ML algorithm used. For example, decision-makers would be able to identify which factors were critical when predicting an outcome or recommending a decision (Freitas 2014b).

Many researchers have introduced different methods in an attempt to open the black box models and replace them with more transparent and explainable ones (Gilpin et al. 2018; Handelman et al. 2018; Molnar 2019; Murdoch et al. 2019; Tamagnini et al. 2017). Before explaining some of the methods, it is important to discern between two types of model explainability: global and local. Global model explainability means that the method used can explain the internal logic of the *whole* ML algorithm at once and why it produced such outcomes (Buhrmester et al. 2019). This level of explainability is about understanding how the model makes decisions based on a *holistic* view of its variables (Molnar 2019). In contrast, local explainability means that the method employed can explain only *one* single instance (e.g., one recommendation) that a ML algorithm produced (Buhrmester et al. 2019).

The two most common methods of local explainability are called LIME (Local Interpretable Model-Agnostic Explanations), introduced by Ribeiro et al. (2016) and SHAP (Shapley Additive exPlanations), introduced by Lundberg and Lee (2017). LIME and SHAP are two methods that locally explain a black box ML classifier in a human-understandable way (Visani et al. 2020). Both methods estimate the variable attributions on individual instances of a dataset, which capture the contribution of each variable on the black box prediction. They are mostly used for image and text classification (Garreau and Luxburg, 2020).

Decision trees are another technique that shows some levels of explainability to ML algorithm and how it reached a certain decision or recommendation. Decision trees depict which features are of relative importance when making a decision. However, for a decision tree to be comprehensible, its hierarchical structure should not be complex and very long. The smaller the depth of an attribute, the more relevant the attribute is for classification. Therefore, small trees containing few attributes are preferred, but this might not be applicable to all contexts. Furthermore, an attribute might occur two or more times in the same path from the root to a leaf, and some paths may contain irrelevant attributes (Freitas 2014). Efforts were also made to make deep learning and convolutional neural network models in a human-interpretable way (Kim et al. 2020).

Some organizations are also devoting efforts toward making ML algorithms explainable and understandable. IBM offers an open-source toolkit that provides explainability methods and interactive demos called AI Explainability 360 (or AIX 360) (IBM Research Trusted AI n.d.). Arya et al. (2020) argue that there is no one best approach to explainability as the definition of an explanation varies by context and depends on its audience. For example, AIX 360 uses Feature Importance (FI) as a method of explanation for bank customers who want to understand why their loan application was rejected. While for a loan officer who wants to understand why a ML algorithm recommends a person's credit to be approved or denied and how it compares individuals with similar profiles, AIX 360 uses a different method of explainability (i.e., ProtoDash) (IBM Research Trusted AI n.d.). More can be found here http://aix360.mybluemix.net/consumer.

Another group of researchers at Microsoft claimed that a centralized platform for recruiting candidates that utilizes a unified tool to screen thousands of candidates across different professions is an inefficient and unfair hiring process. The researchers developed an experimental platform to examine ways to mitigate gender bias in hiring decisions (Peng et al. 2019). They defined bias as "*any observed difference in recommendation decisions such that one gender is favoured over another in a manner that does not correspond to the distribution input into the algorithm or human*" (Peng et al. 2019, p.127). The authors discovered that across different professions with varying gender distributions, balancing gender representation in candidate slates can help mitigate biases for some jobs. They also concluded that the gender of the decision-maker, the complexity of the decision-making task and the unbalanced representation of genders in the candidate's pool could greatly impact the final decision. This means that a more transparent process and explainable versions of ML algorithms can greatly help eliminate many of the problems associated with the decision-making process. Recent work suggests that explainability is an audience-dependant instead of a model-inherent property (Miller 2019a; Mohseni et al. 2021) that may differ from context to context and from field to field.

In an attempt to foster transparency when using recommender systems, Symeonidis (et al. 2009) proposed "MoviExplain" as a platform that offers a transparent way to justify and explain why certain movies are being recommended to users. The platform enables users to check the reasons for a recommendation and understand the strengths and limitations of the recommendation process. To do so, MoviExplain relies on two crucial factors: 1) the most important features for a rated movie and 2) users' history relating to these features. This approach proved to contribute significantly to users' satisfaction compared to previous methods that only recommended movies without proper explanation. The platform also provides a help section that explains how

MovieExplain works and how users can improve its performance (http://delab.csd.auth.gr/MoviExplain/moviexplain.htm#recommendations).

In the medical field, it is also critical to understand the decisions produced by an AI to build trust with it. Liu et al. (2022) raise the concern that AI is not common in hospitals, and he justifies that one contributing reason to this could be due to the lack of clear explanations from AI to the physicians. If physicians cannot acquire comprehensible explanations from an AI, they will lack confidence in it and may hesitate to incorporate it into the decision-making process (Liu et al. 2022; Markus et al. 2021).

2.4. AI Autonomy

Hoffman and Novak (1996) believe that human control is a key feature of interactive technologies. The emergence of cognitive systems and AI that can mimic human behaviour reshaped the degree of human involvement and control in completing tasks. Furthermore, user control leads consumers to believe they can influence their goal attainment process and thereby increases their confidence about the outcome (Bateson and Hui 1987). Meerbeek et al. (2008) also assert that human control is an important aspect in human-computer interactions and human-robot-interactions. They define control as "*one's ability to affect the outcome of the interaction*" (Meerbeek et al. 2008, p.208). Also, Skinner (1996) defined control as "*the extent to which an agent can intentionally produce desired outcomes and prevent undesired ones*" (Skinner, 1996, p.554).

Another critical concept in the field of intelligent systems and robotics that is closely associated with control is "autonomy" (Meerbeek et al. 2008). The term "Autonomy" is derived from the ancient Greek word "autos," meaning "self," and "nomos," which refers to "rule or law."

Indicating that people can make their own laws or rules (Dworkin 1988). Autonomy has been studied in various contexts and disciplines, such as psychology (Deci et al. 1989; Karasek 1979), organizational studies (Mazmanian et al. 2013; Trevelyan 2001), philosophy (Dworkin 1988), marketing (Wertenbroch et al. 2020), and information systems (Hua Ye and Kankanhalli 2018; Igbaria et al. 1991; Moore 2000); and can greatly impact teams (Kakar 2016) and individual outcomes (Hua Ye and Kankanhalli 2018).

Accordingly, manifold conceptualizations of "Autonomy" emerged. In psychology, autonomy is a fundamental concept for different theories, such as the job characteristics model (JCM) (Hackman and Oldham 1976) and the self-determination theory (SDT) (Deci et al. 1989). Hackman and Oldham defined autonomy as "*the degree to which the job provides substantial freedom, independence, and discretion to the employee in scheduling the work and in determining the procedures to be used in carrying it out*" (Hackman and Oldham 1976, p.258). Dodd & Ganster, (1996) conceptualize autonomy as " *the control the worker enjoys with respect to choosing among the operations, ordering the operations, and selecting a work pace*." Others defined autonomy as "*the ability to make and enact decisions on their own, free from external influences imposed by other agents*." (Wertenbroch et al. 2020, p. 430). More definitions of "Autonomy" can be found in (Beer et al. 2014).

While autonomy in previous studies was mostly attributed to humans (André et al. 2018; Wertenbroch et al. 2020), the evolution of mobile technologies and artificially intelligent agents gave rise to attributing autonomy levels to these technologies and intelligent machines (Zhang et al. 2022). The degree of intelligent systems' autonomy affects the degree of human control in a human-machine-interaction. For example, intelligent systems can free humans from some cognitive load when performing a task that involves processing large quantities of information.

However, this may come at the expense of lessened control for humans (Meerbeek et al. 2008). Some predict that intelligent systems could be completely autonomous in the future and that even programmers and system developers, at some point, may lose control over influencing the direction of autonomous AI or their underlying algorithms (Maitra 2020).

Parasuraman et al. (2000) developed a taxonomy for the levels of human control and machine autonomy in a human-machine interaction realm. The authors proposed ten different levels, where level one is complete human control, and the tenth level is full machine autonomy. In between these levels, a shared responsibility of accomplishing tasks or making a decision is encountered between humans and machines. They proposed that these levels of automation can be applied to any of four broad tasks: information acquisition, information analysis, decision and action selection, and action implementation. Since the word "autonomous" machines or intelligent agents could be vague as the extent of autonomy may not be defined, with these levels of automation, one can get an idea of the level of autonomy a machine or a human may possess.

PricewaterhouseCoopers (2017) defined four categories to classify AI in terms of autonomy and adaptability: i) assisted intelligence, ii) automation, iii) augmented intelligence, and iv) autonomous intelligence. *Assisted intelligence* is where AI interacts with and assists humans in carrying out tasks or making decisions. However, they do not learn from these interactions and perform the same thing in the same way again and again. Expert systems that are based on a predetermined set of rules used to make a decision or recommend an action fall into this class. Similarly, *automation* is when an AI accomplishes tasks without human assistance or adaptability to new situations, such as robots operating on a manufacturing assembly line.

On the other hand, *augmented intelligence* refers to the type of AI that works collaboratively with humans and assists (or augments) them in their decision-making process.

Conversational agents and chatbots fall under this category of AI as they also learn from interacting with humans and can continuously adapt to new environments and situations. Finally, *autonomous intelligence* is the type of AI that can work autonomously and adapt to different situations without human assistance. Although this type is still under development, some robots and AI software programs in this class are now emerging in the marketplace.

The most common applications and areas of research where autonomous machines or systems are being utilized are in the automobile industry (Garidis et al. 2020; Hein et al. 2018). SAE International (2014) has even introduced a classification for the levels of autonomy of vehicles and introduced five main levels. In the first level, the human driver takes control of the vehicle, and in level five, the vehicle is entirely autonomous and can operate the vehicle in all situations without the lack of human assistance. Ernst and Reinelt (2017) conducted survey-based research to study how people's perceptions of traffic safety and the enjoyment of driving a car would influence their acceptance of autonomous cars. Respondents indicated that personal driving enjoyment negatively influences their perceived enjoyment of autonomous cars, and that their perceived traffic safety positively influences the perceived usefulness and their perceived enjoyment. Findings confirmed that perceived usefulness and perceived enjoyment positively impact autonomous car acceptance. These findings suggest that while autonomous cars can entirely do the job and drive the car without human assistance, it asserts the notion that humans should, in turn, be given the option to be the drivers to maintain their enjoyment cues.

Although driverless cars attracted the greatest attention, autonomous trucks are also being studied. Due to their potential in the operations management and supply chain networks domains, Sternberg et al. (2020) interviewed experts and conducted a scenario analysis to explore the factors that may potentially impact the adoption of autonomous trucks. The authors concluded that technological maturity and regulations would be the two most significant factors in influencing the adoption of autonomous trucks.

However, despite the levels of autonomy introduced to autonomous machines or agents, they still lack the capabilities humans possess. We are still far away from completely replacing humans. Consequently, these intelligent agents could be very helpful in accomplishing tasks that may cause harm to humans (Hyder et al. 2018) or are considered routine laborious tasks (Wang et al. 2020). In other terms, delegating the negative aspects of work to an AI instead of humans (Seeber et al. 2018; Welfare et al. 2019). Therefore, humans can devote themselves to more complex tasks that would, for example, require creativity, judgment, special skills, or complex rational thinking (Seeber et al. 2018).

To tackle the debate of having humans lose control in the presence of autonomous AI, Akmeikina et al. (2022) suggest that we should not look at it as a controller-controlled interaction. But rather, look at it from a lens where all actors have a sufficient degree of control and capabilities to enter an interaction (Liu and Zawieska 2020). Akmeikina et al. (2022) stress that to account for autonomy when interacting with intelligent machines, the well-being of everyone must be a priority, and self-interest should always be secondary. Such that autonomy is where the autonomous agent is not only aware of the needs of others but should also be aware of what they can do for the good of the group.

Since autonomy is essential for human workers in most work contexts, Fernández-Macías et al. (2018) highlighted the importance of examining the role of AI autonomy and its impact on the workplace. The authors discussed how the nature of working environments might be altered by AI learning capabilities that autonomously acquire new skills or knowledge and can apply them in different tasks and contexts. Prior literature has also discussed the issue of autonomy. For example, Smithers (1997) suggested that intelligent agents need to be autonomous to an extent in case they have to effectively and intelligently work in environments where we live and work. Covrigaru and Lindsay (1991, p.111) stated, "*An entity must be autonomous to be truly intelligent; truly living; and, thus, truly humanoid*."

2.5. Human-AI Collaboration

In 1990, Blattberg and Hoch published a paper entitled "Database Models and Managerial Intuition: 50% Model + 50% Manager" (Blattberg and Hoch 1990). The authors were interested in studying whether decision outcomes would be better if database models alone were used without human judgment in two different forecasting circumstances. They found that the best results were always achieved when a manager's intuition was combined with a database model, compared to when either worked alone to solve the problem. This outcome was justified by proposing that each party has strengths and weaknesses, and when they are combined to work together, better results are always achieved. For example, experts may be influenced by organizational politics, suffer from ego or social pressures, get tired or bored, and not explain or provide evidence for their judgment. However, models do not suffer from these weaknesses.

On the other hand, experts also enjoy many strengths that models do not have. Experts can apply intuition and provide subjective assessments of variables that models are not able to evaluate objectively (Einhorn 1974). Experts can also be flexible and adapt to different circumstances, and unlike models, they may use cues that cannot be quantified or represented in a linear statistical model. This means that the statistical and mathematical power of models combined with experts who can apply and share their knowledge is the recipe for enhanced decisions. Sowa et al. (2021) also suggest that having human workers and AI collaborating together in performing managerial tasks increases productivity rather than having each working individually. These findings advocate the value of having humans and machines collaborate such that they utilize each's strengths and overcome each's weaknesses in order to boost the decision-making process and achieve superior outcomes. Therefore, instead of the conventional unidirectional system use, where humans issue a request to a system and then wait for the system to deliver a result. The newly emergent capabilities and agency of AI systems advocate the notion of bilateral interactions (Schuetz and Venkatesh 2020).

One of the first attempts that talked about having machines and humans in the loop is Fitts's MABA-MABA ('Men are better at; Machines are better at') concept (Fitts 1951). Fitt discussed whether certain tasks should be carried out by a machine or a human, in which Fitt suggested that humans and machines both have a list of different capabilities and limitations that we should be aware of when making decisions about task allocation and division of responsibility. For example, in the context of air navigation and traffic control systems in the early 1950s, Fitt mentioned that expediting traffic could be attributed to men. While monitoring collisions could be the responsibility of machines since humans are not very good at monitoring systems for a long time. Humans might get bored, sleepy, or inattentive (Fitts 1951; Mackworth 1950). However, such lists motivated one to compare human and machine capabilities who can replace each other in certain tasks rather than thinking of designing hybrid activities where humans and machines intertwine (Mackeprang et al. 2019)

Furthermore, Nass and his colleagues were pioneers in introducing the Computers as Social Actors (CASA) paradigm (Nass et al. 1994, 1996). The authors conducted five different experiments with experienced computer users to answer five different research questions in an attempt to understand which social rules people would apply to computers. Despite users' awareness that computers do not have feelings, gender, or a sense of motivation, in the five experiments, users perceived computers as social actors and applied social norms when interacting with the computers. It was evidenced that computer users apply gender stereotypes when interacting with computers, and they may even disclose some private information when computers converse in a human-like way (Nass and Moon 2000). This confirms what the media equation theory suggests, as introduced by Reeves and Nass in 1997, indicating that people tend to treat and respond to communication media and computers as they do to real people (Mou and Xu 2017).

Many researchers attempted to understand how people would perceive AI as a collaborator. Papachristos et al. (2021) developed an AI tool for a waste sorting task. The task entailed that once someone opens a physical waste bin and throws an item, the AI processes the image of the item and classifies it to either a waste or a recyclable item on a large screen attached to it. The AI also gives the user a score of how confident the AI is about its recommendation, and the final decision is up to the user whether to accept or reject the recommendation. The authors then interviewed 35 participants and asked them about their perceptions. The majority of participants reported a positive experience when collaborating with the AI, and the authors classified the roles of the AI according to the interviewees' responses into four roles: mirror (i.e., the AI mirrors or confirms participant's decisions), assistant (i.e., participants expect the AI would assist them only when unsure about their decision), guide (i.e., the AI provides a recommendation and a confidence score that participants happily accept it), and oracle (i.e., participants believe that the AI would always have better judgment than themselves). Additionally, Mou and Xu (2017) compared people's initial interactions with humans versus AI to see whether humans would disclose their personality traits and communicative attributes when interacting with an AI chatbot versus a human. Their findings suggested that when participants interacted with an AI, they exhibited different personality traits and communication attributes from interacting with humans. Participants were more open, agreeable, extroverted, more self-disclosing when interacting with humans than with an AI. When conversing with a chatbot, people also tend to like chatting with the chatbot through shorter messages for longer periods of time than they would do with a human (Hill et al. 2015).

Fan et al. (2022) wanted to test how synchronicity and explanations of AI collaborators would help User Experience (UX) evaluators when they evaluate usability test videos. The authors built an AI tool and experimented it quantitatively and qualitatively with 24 UX evaluators. Participants were sent to four different treatments (with explanation, without explanation, synchronous, and asynchronous). In the treatments with asynchronous AI, the AI highlighted to UX evaluators all the problems it identified at once early in the evaluators' analyses and was always available to the evaluators during the analysis. In contrast, the synchronous AI presented usability problems to evaluators gradually as they progressed with watching the video. In each of these treatments (i.e., synchronous and asynchronous), some participants had explanations about the AI suggestions, and some did not see any explanation for the AI recommendations. The authors found that regardless of whether being in the synchronous or the asynchronous treatment, having an AI that provided explanations to its suggestions helped evaluators to be more engaged and highly likely to accept the AI's recommendations. It helped participants to perform equally well overall. While in the case where the AI did not show explanations of their recommendations, evaluators' performance and engagement in the synchronous treatment were higher than those in the without explanation and asynchronous group. This implies that explainable AI plays a vital role in shaping users' acceptance and perception of collaborating with an AI.

Others have studied how different settings of Human-AI collaborations can shape people's perceptions and acceptance of collaborating and interacting with AI. Fogliato et al. (2022) investigated whether providing humans with an AI recommendation before versus after the human reviews a diagnostic task would affect humans' decisions and the decision-making process. The authors conducted an online experimental study where 19 veterinary radiologists were assigned to two treatments. In one treatment, radiologists were presented with 20 X-ray images of dogs along with the AI inferences of the images to aid the radiologist with their decision. In the other treatment, radiologists were shown only the 20 X-ray images and were asked about their diagnoses before they were presented with the AI inferences. Results showed that the one-step approach (i.e., showing participants the X-ray images and the AI advice at the same time) yielded better outcomes overall than the two-step approach. The authors also found that radiologists who were not shown the AI inferences except after they were asked about their own diagnoses (i.e., the two-step approach) were less likely to agree with the AI recommendation even if it was accurate, and they also rated the AI to be less useful. While participants from the one-step approach reported that the AI was useful in the decision-making process. This reveals that keeping AI in the loop with humans from the beginning may be a good approach for an improved decision-making process.

Even in systems development, researchers are exploring ways where humans and machines can be partners. Mackeprang et al. (2019) developed a tool in which they tested how AI could collaborate with humans in the information extraction process using collaboration ideation platforms (i.e., a platform that collects different ideas from the crowd or from a number of ideators in a distributed setting). The authors' ultimate goal was to find the right level of algorithmic support without compromising the quality of the extracted information and considering that the human effort should be low. Testing automating the various stages followed to extract information, the authors were successful in identifying and understanding the potential areas where Human-AI collaboration could be valuable and efficient.

AI could also be a time saver. In a study conducted by (Rzepka et al. 2020), it was found that users welcome the idea of collaborating with voice assistants when they purchase routine products online. Participants attributed the reason to the voice assistant being helpful as it saves them time by skipping regular steps to make a purchase, such as opening the app, searching through products, and choosing a product. They admit that the help they get from such AI made the process much easier and quicker (Rzepka et al. 2020).

In complex, uncertain decision-making domains such as military issues related to defence and security, (van den Bosch and Bronkhorst 2018) believe that the decision-making process has to be a joint endeavour of humans and intelligent agents collaborating together. To do so, the authors advocate that the AI has to be perceived by humans as a collaborator that can interact, adapt, communicate, and be aware of the goals and the situations. The authors also assert that having explainable AI that increases its transparency is essential for Human-AI collaboration.

Amershi et al. (2019) proposed 18 design guidelines for Human-AI collaboration to serve as a reference to researchers studying the interactions between humans and AI and to practitioners who design AI. The 18 guidelines were validated and used across ten different AI product categories. The authors grouped the 18 guidelines into four main stages of interaction: 1) initially, 2) during the interaction, 3) when the AI system is wrong, and 4) over time. The guidelines stress the importance of keeping human collaborators informed of what the AI capabilities are, giving humans the autonomy to correct or edit any AI mistakes, and designing AI collaborators that can explain its actions to humans.

Despite the fact that Human-AI collaboration emphasizes synergy, wherein AI complements human capabilities, filling gaps and enhancing overall performance, some may fear that AI will compete with humans and take over their jobs (de Cremer and Kasparov 2021). This perspective underscores AI's increasing capabilities and humans' potential to excel in specific domains or tasks. It also pushes human experts to continually enhance their skills and raises concerns about job displacement and exacerbating inequalities. Thus, in this work, the focus is on leveraging humans' abilities as well as AI's capabilities to better understand how both can collaborate and work together.

Summary

AI is used in various contexts and has great potential to collaborate with humans. While most prior research views AI as a mere technology that people use, there is a gap in the literature that empirically studies AI as a social actor with which humans collaborate. The classifications that categorize AI into different types always take into consideration whether the AI is autonomous or assistive, which necessitates a deeper understanding of how the two modes might affect individuals' beliefs. In addition, integrating AI as collaborators within organizations, where they may be perceived as a black box to humans, raises the need to investigate the role of AI explainability since AI might be perceived as a black-box to humans, and collaborating with an opaque member could have an influence on people's beliefs towards collaborating with it. These gaps are addressed in this research.

Chapter 3: Theoretical Development

3.1. The Information Systems Field and Epistemology

Each discipline has its unique characteristics and approach to conducting research. Myers in 2019 suggested that there are some phases or steps that any research design should follow: 1) identifying the philosophical stance of the researcher; 2) determining the research method; 3) defining data collection techniques; 4) specifying how the data will be analyzed; and finally, 5) the write-up of the research conducted (Myers 2019).

Before conducting a research study, researchers should first clarify their epistemological stance or appropriate philosophical assumptions and the discipline to which the study contributes (Myers 2019). Philosophical assumptions are assumptions about the nature of the world and how knowledge about the world can be obtained. These assumptions are used as a foundational guide for the research that will be conducted.

Since this research study contributes to the Information Systems discipline, there are three main philosophical assumptions discussed by (Orlikowski and Baroudi 1991): the positivist, the interpretive, and the critical. The research follows a positivist stance if its ultimate objective is to test a theory as an endeavour to boost the predictive understanding of a certain phenomenon. Therefore, positivist research includes a set of formal propositions or hypotheses, assumes quantifiable variable measures, aims to test hypotheses, describes the phenomena from inferences of the sample and generalizes it to a larger population (Myers 2019; Orlikowski and Baroudi 1991).

In contrast, the ultimate goal of an interpretive approach is to explore a phenomenon in its natural setting rather than testing it. This takes place by understanding and examining the meanings and interpretations that research participants have about a phenomenon within its natural contextual settings. Thus, the generalizability of findings from the sample of participants to a greater population is not required in an interpretive approach (Myers 2019). Finally, the critical approach, from its name, seeks to critique and challenge the taken-for-granted assumptions that we have about a certain phenomenon. It attempts to uncover the contradictory nature of current social practices in organizations and information systems (Myers 2019).

Of these three philosophical assumptions, this research study follows a positivist philosophical stance. The reason for taking a positivist stance is that this study attempts to test a theory using a set of predetermined constructs and a set of quantifiable measures for the constructs. Moreover, generalizability is one of the research objectives of this study.

Furthermore, in terms of the research method, this study follows a mixed-methods approach to increase the robustness of the research through triangulation. Triangulation is about following more than one approach or doing more than one thing in one study (Myers 2019). It could be by using more than one research method (e.g., case study and ethnography), using more than one data collection technique (e.g., survey and interviews), combining quantitative and qualitative methods in one study, or even including more than one researcher in the study where each has his/her own expertise (Myers 2019).

Mixed-methods research enables the development of novel theoretical perspectives by combining the strengths of the quantitative and qualitative methods and overcoming the drawbacks of each in order to derive stronger inferences about the phenomenon under study (Creswell 2009; Venkatesh et al. 2016). Venkatesh et al. (2016) provided six guidelines to be followed when adopting a mixed-methods approach. The first guideline requires that the researcher determines the appropriateness of utilizing a mixed-method in their study. The authors mentioned that

"*Researchers should employ a mixed-methods design only when they intend to holistically explain a phenomenon for which extant research is fragmented, inconclusive, and/or equivocal*" (Venkatesh et al. 2016, p.437). Extant literature lacks sufficient empirical work that studies people's perceptions of AI as collaborators in organizations. Furthermore, current research work is ambiguous and equivocal in terms of defining what AI is and how it influences employees in organizations. Thus, employing a mixed-methods approach would be appropriate in this study.

Moreover, the authors discussed that the second important step in employing a mixed-methods approach is to decide on the strategy for the research design. Two strategies to mixed-methods research design were proposed by (Venkatesh et al. 2013, 2016): sequential or concurrent. The sequential mixed-methods design entails conducting a quantitative study followed by a qualitative study or a qualitative study followed by a quantitative study in which results from one study inform the other. Concurrent mixed-methods design, on the other hand, involves conducting both methods at the same time (i.e., independent from each other where the result from one does not inform the other). The rest of the guidelines proposed by (Venkatesh et al. 2016) are related to developing strategies for collecting and analyzing the mixed-methods data, drawing meta-inferences from it, assessing the quality of the meta-inferences, and discussing potential threats and remedies.

In this study, a proposed model is developed with a set of predetermined constructs and hypotheses. However, perceiving AI as a collaborator with humans in an organizational context is a nascent phenomenon that is associated with both concerns and benefits. These concerns and benefits are not fully understood in the literature. Such concerns and benefits are likely to influence people's willingness to collaborate with AI. Consequently, two essential constructs (i.e., perceived benefits and perceived concerns) are examined through a qualitative study to grasp people's perceptions about these concerns and benefits that will then be used to deduct and quantitatively validate the proposed model. Therefore, in this study, a two-stage sequential mixed-methods approach is employed in which qualitative data was collected and analyzed, and then a quantitative study took place.

For the qualitative part, this research follows a positivist approach (Charmaz 2014; Myers 2019) for different reasons. Positivists view concepts as variables for which they provide operational definitions to be used in hypotheses testing. Moreover, a positivist stance identifies relationships between concepts, explains and predicts these relationships, verifies these relationships through hypothesis testing, and aims for context-free generalizations (Charmaz 2014). For the quantitative part, a positivist stance is also employed to test a set of predetermined hypotheses that are grounded in theory and tested using structured, validated instruments (Venkatesh et al. 2016).

3.2. Theoretical Premise

Theories and theoretical frameworks are important avenues that assist in explaining, predicting, and understanding a phenomenon under study. They are used across disciplines as a lens for researchers to guide them through their study and to help them ground their arguments.

Since the Information Systems (IS) field is multidisciplinary in nature, such that it deals with human-machine artifacts (Gregor 2006), theories in IS are different as they are required to connect the difference in nature between the artificial world of technology and the social and behavioural world of humans. To put this study on solid grounds, it will be carried out using a robust theoretical stance that is appropriate to the context and goal of this research.

3.3. Contextualization

Hong et al. (2014) discussed the value of contextualization when developing a theory in the information systems field. Notwithstanding that contextualization might necessitate forgoing parsimony and generalizability (Hong et al. 2014), it produces richer and more practical theories that are relevant in information systems research (Johns 2006; Orman 2002). Rousseau and Fried (2001, p.1) mentioned that "contextualization entails linking observations to a set of relevant facts, events, or points of view that make possible research and theory that form part of a larger whole." Johns (2006, p.386) defined context as "situational opportunities and constraints that affect the occurrence and meaning of organizational behaviour as well as functional relationships between variables." The author emphasized that context is crucial as if researchers did not understand the contextual situations, it would be challenging to understand person-situation interactions. Alvesson and Kärreman (2007) also stressed that no theory is right or wrong all the time, but theories can be more or less helpful and relevant depending on the situation (i.e., the context). In addition, context helps researchers communicate the practical applications of their research findings (Johns 2006) and can act as the starting point for generating new theories (Hong et al. 2014).

Orman (2002) discussed that sometimes it is hard to predict the influence of information technologies on organizations. The author explains that this is mainly because, in IS research, all information technologies are almost treated and viewed to be the same. This ends up having confounding results as different technologies in different contexts can have dramatically different effects on organizations (Orman 2002). Rousseau and Fried (2001, p.1) highlighted that "*contextualizing entails linking observations to a set of relevant facts, events, or points of view that make possible research and theory that form part of a larger whole*."

Moreover, Whetten (2009, p.31) defined “context effects” as “*a set of factors surrounding a phenomenon that exert some direct or indirect influence on it*.” Weber (2003) outlined that when studying and accounting for patterns of a certain phenomenon, it has to be contextually driven. The author stated that “ *grounding our theories in rich contextual tapestries will lead to important insights about phenomena associated with humans, information technology, information systems, and organizations*.” (Weber 2003, p.x)

Hong et al. (2014) suggested two ways for contextualization in information systems research: single-context theory contextualization and cross-context theory replication. The former involves contextualizing existing theories by adding or removing contextual constructs either as antecedents to dependent variables or core variables in the model, as moderators of the relationships, or by deconstructing core variables into context-related ones. The second approach includes replicating theoretical frameworks in different contexts, then performing theory-grounded meta-analyses to integrate the findings into a context-dependent theory.

In this research, the single-context theory is followed where existing well-established theories are contextualized, and two strategies highlighted in the single-context theory contextualization approach are used. First, high-level core constructs of perceived benefits and concerns are decomposed into constructs related to the context of AI. Second, two contextual technology characteristics (AI autonomy and AI explainability) were added as antecedents to the decomposed constructs of the perceived benefits and concerns where these decomposed variables mediate the relationship between the endogenous construct and the AI contextual characteristics.

Context embodies several aspects such as tasks, technology, individual, and organization. While these contextual characteristics are important (Brown et al. 2010), in this study, the focus is on the technology characteristics (i.e., AI autonomy and AI explainability) as they are nascent characteristics of AI that have not been studied enough yet in the information systems literature. These features can also be attributed to AI as a technology with a social actor role in a network of hybrid Human-AI members which is in line with the premise of this research. Weber et al. (2020) suggest that with the increase of mobile technology, researchers should examine “Autonomy” in a contextualized setting to get richer theoretical significance in IS research. Finally, focusing on the AI characteristics only sets the boundary for this study and minimizes potential confoundings. This is in line with Gregor (2006), who asserts that even in theory building, it is important to specify a boundary or a scope of the theory that defines its level of generalizability. Having said that, other contextual factors are indirectly considered in this research (e.g., participants have various individual traits and work for a variety of organizations and on a variety of tasks).

3.4. Theories Utilized in this Research

As described in Chapter 1, one core objective of this research is to leverage theory to determine the most prevalent contextualized factors in hybrid Human-AI collaborative settings. To do so, it is important to understand how individuals make a trade-off between the positives and negatives when collaborating with AI that are not deemed to be a mere assistive technology but social actors with whom they can collaborate (Sowa et al. 2021). Since humans follow a “net valence” approach when considering dealing or working with new technology (Breward et al. 2017; Cazier et al. 2008), this study is grounded in the Actor-Network Theory (ANT) and the Net-Valence Theory to understand the phenomena.

3.4.1. Actor-Network Theory (ANT)

According to the Actor-Network Theory (ANT), humans and non-human actors should be treated as inseparable. ANT is a socio-technical approach that aims to examine the motivations and actions of a heterogeneous network of human and nonhuman actors all together (Walsham 1997). The theory tries to trace and explain the processes through which relatively stable networks of aligned interests are created and maintained or why such networks might fail to create themselves. Kaartemo and Helkkula (2018) suggested using the Actor-Network Theory as a lens to understand the agency of technology and how AI and humans can collaborate to create value. AI should, therefore, not be viewed merely as an innovative technology but rather as vital social actors in a network with humans with whom it can interact, exchange knowledge, and make joint decisions.

Applying this lens, this study views AI as a social actor who participates and takes an active role as a real member of a network facilitating hybrid Human-AI collaboration. Accordingly, AI (i.e., non-human) will be treated as inseparable from human members in an organizational setting. This research is interested in studying how some of the functions normally attributed to human members would affect human beliefs if they were attributed to the AI as a collaborator.

3.4.2. Net-Valence Theory (NVT)

The "valence" concept was first introduced by Lewin (1943) in the context of consumer behaviour, where consumers make a trade-off between the positive utilities and negative utilities of products. Bilkey (1953) discussed that consumers make choices regarding products or services by comparing both the favourable (positive valence) and unfavourable characteristics (negative valence). Consumers then aim to maximize the "net valence" which is the arithmetic difference between the expected benefits and costs (Bilkey 1953; Lewin 1943). If the benefits (i.e., positive valences) outweighed the costs (i.e., negative valences), consumers would buy the product or service and vice versa. Thus, the Net-Valence Theory suggests that individuals engage in a cost-benefit analysis activity when making decisions, as they try to maximize the benefits and minimize costs.

Featherman (2001) suggested the adoption of the Net-Valence Theory in the information systems discipline, where the author proposed a model to extend the Technology Acceptance Model (TAM) and suggested that if the expected benefits (positive utility from usage) exceed the potential concerns (negative utility from usage), the information system will be adopted. Likewise, Pentina et al. (2016) adopted a Net-Valence approach in the context of information-sensitive mobile application use, where they referred to as "privacy calculus." The authors explored the role that the perceived benefits and perceived privacy concerns played in determining people's intentions to use mobile applications that collect sensitive personal information. In such context, the authors assumed that users would be willing to use the mobile application if the benefits equal or surpass the concerns.

The Net-Valence perspective is thus helpful in terms of modelling a more balanced evaluation of information systems adoption and use, as well as yielding profound practical and theoretical contributions (Featherman 2001). Breward et al. (2017) outlined that following a Net-Valence perspective would be helpful when studying the use of information technologies that people viewed with skepticism.

From the above, when humans contemplate integrating new technology into their work or life, they follow a "net valence" approach by weighing both the benefits and concerns of incorporating that particular technology (Breward et al. 2017; Cazier et al. 2008; Dinev and Hart 2006). Hence, NVT is used in this research as a lens when examining the benefits and costs of collaborating with AI. This study proposes that in hybrid Human-AI collaboration, human members would weigh both the benefits (positive valences) and concerns (negative valences) of collaborating with AI, such that if the benefits are equal or greater than the concerns (i.e., net-valence is positive), humans would be willing to collaborate with AI. Figure 5 highlights a general framework that will be used in this study to inform the empirical research model.

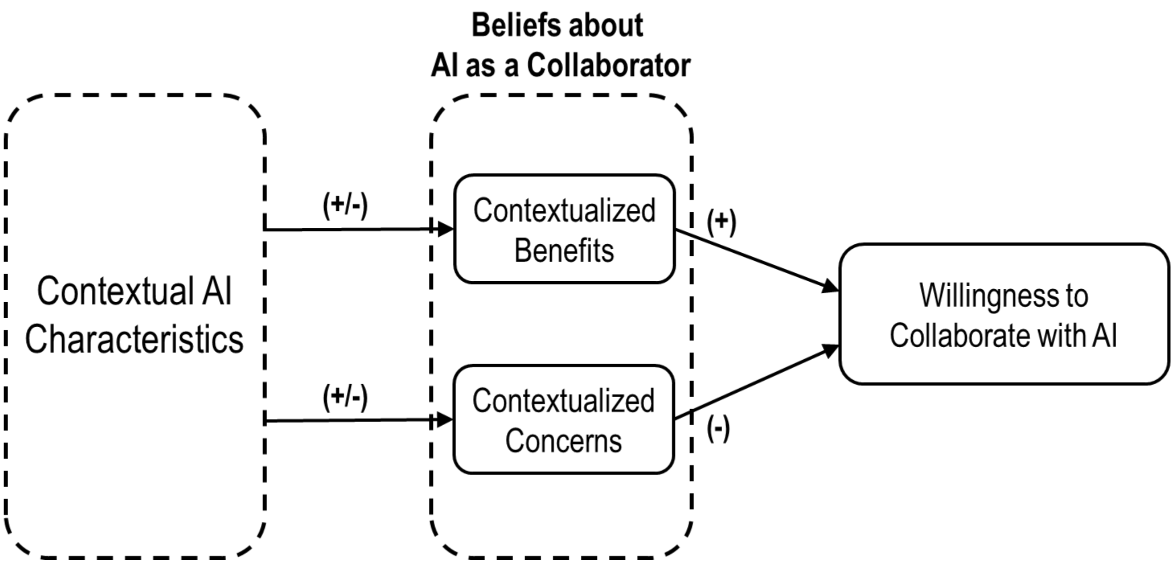


Figure 5: A General Framework to Understand the Willingness to Collaborate with an AI

Summary

The theories reviewed here are used to develop a contextualized research model.Although many contextual characteristics are important, the focus of this proposed study is on two contextual AI characteristics (i.e., Autonomy and Explainability). Further details of the selected AI contextual characteristics and why they were selected are discussed in the next chapter.

Chapter 4: The Qualitative Phase

AI is developed, implemented, and used by people in real-world settings. There is plenty of research on the technical aspects of AI. However, it appears that the literature is bereft in terms of empirically studying AI as a social actor rather than a mere assistive technology. Insufficient empirical research has been conducted to understand people's perceptions regarding collaborating with AI. Beliefs about the benefits and concerns of collaborating with AI are essential when designing and implementing such revolutionary technologies that would collaborate with employees in organizations. As such, a qualitative research design is best employed to understand human behaviours in the context of an AI's development, implementation, and use (Myers 2019). The objective of the qualitative phase of this study is to identify key benefits and concerns to inform the research model and test it empirically.

As suggested by Sarker et al. (2018), there are four key aspects to consider when designing a qualitative research study: theory, data, analysis, and the nature of findings. These four pillars are addressed in this research. First, with regard to theory, a conceptual framework is used to guide the data collection and analysis process. Second, data is collected through one-on-one interviews with participants to understand their perceptions regarding the benefits and concerns of collaborating with AI. Interviews are the most common and important qualitative research data collection technique (Myers 2019). They help researchers collect rich data from people in various situations. Third, in terms of the data analysis tactic used, a content analysis approach is followed to generate themes and concepts. Finally, concerning the nature of the findings, this study focuses on generating new insights (Walsham, 1995), as well as understanding the context of using new IS (i.e., AI) (Walsham, 1993).

4.1. Pilot Study

Before collecting the main interview data, a pilot study was carried out. It is claimed that before collecting all the interview data, pilot interviews should first take place to guide the rest of the interviewing process (Fox-Wolfgramm 1997), especially if it is the researcher's first journey to conducting interviews (Iyamu 2018). It alerts the interviewer to some issues in advance, such as the appropriateness of the approach used, the duration of the interview, how clear the questions are, and whether they are in the right direction as the researcher hoped (Iyamu 2018; van Teijlingen and Hundley 2002).

Thus, I conducted three online pilot interviews (see Appendix A for participants' details) to see if anything needed to be changed or modified in the interview structure. Prior to conducting the pilot or main studies, McMaster's research ethics clearance (MREB#5134) was obtained. Participants for the pilot and main studies were managers or employees with hiring experience of at least one year. All participants were recruited through social media, followed by a snowball sampling technique. They were also informed before the interview that interviews would be recorded for analysis reasons.

At the beginning of the pilot interviews, I welcomed participants and reviewed with them the information about the study. Since AI is a buzzword nowadays, it was important to put participants on common ground and provide them with a definition of AI that fits the context of this study. Therefore, I informed participants that our study defines AI as "*a technology that can imitate intelligent human behaviour. It can perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting, problem-solving, and decision-*

making" (Rai et al. 2019, p.iii). Participants were then asked if they had any questions. After ensuring participants were ready to move on, they were asked to imagine that their organization would adopt an AI to collaborate with them in hiring job candidates. Subsequently, participants were asked a set of questions.

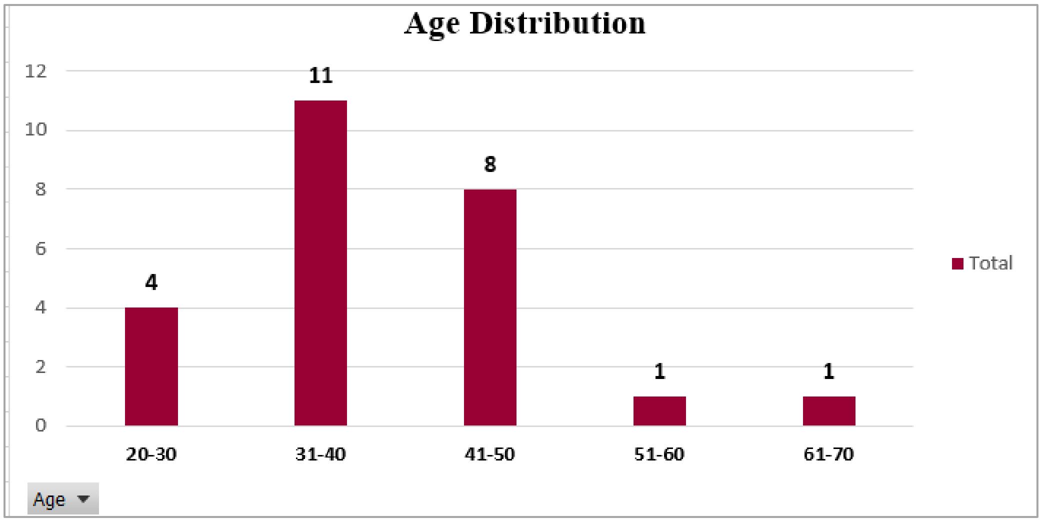
Myers (2019) discussed three possible formats for interview questions: structured, semi-structured, and unstructured. The structured format involves a set of predetermined questions a researcher has to ask and is not allowed to pose follow-up questions that might evolve from the interviewees' answers. In contrast, unstructured interview questions offer interviewees the freedom to speak, and follow-up questions can evolve from there. Finally, semi-structured questions are the most effective format as they combine the advantages of the other two types (Myers, 2019). They are pre-formulated to guide the interview and allow space for additional questions to emerge from the interviewees' answers. Babbie (1998) contends that semi-structured interviews give participants a chance to do most of the talking, which helps in gaining a better understanding of their views of the topic being studied. Accordingly, semi-structured interviews were used in this research that provided some structure to the interview using predetermined questions and also allowed participants to freely express their thoughts to obtain rich insights from the interviewees. Hence, participants were asked the following questions:

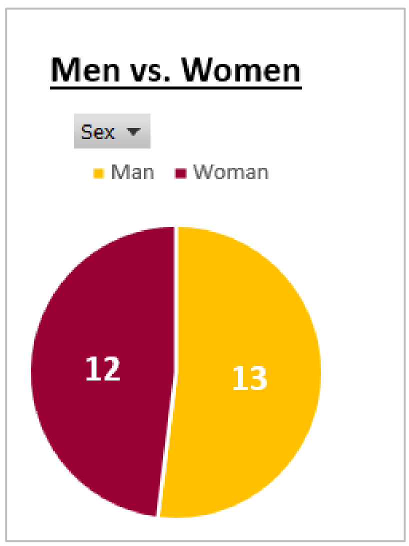
1. What would be the benefits/advantages of collaborating with AI to recruit candidates?
2. What would be the concerns/disadvantages of collaborating with AI to recruit candidates?
3. Please provide any other comments regarding collaborating with AI in the workplace.

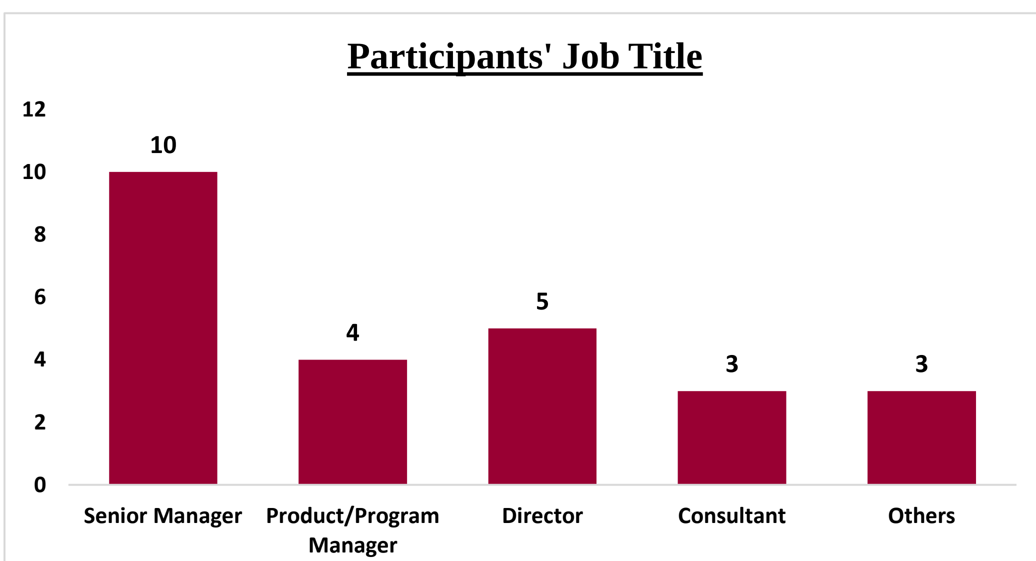
The pilot interviews were helpful in informing the interview structure. Interviews lasted between 25-30 minutes, and participants agreed that this duration was optimum for them. After conducting the pilot interviews, I noticed that participants had difficulty imagining what capabilities AI would offer and how AI would collaborate with them in a recruitment context. Therefore, in the main interviews, I chose to show interviewees a one-minute video that demonstrated how AI could be used in the recruitment process. Also, I found that participants during the pilot study sometimes mixed between the recruiter's perspective and the job candidate's when discussing the benefits and concerns. This was useful in the main study, where I had to ensure that participants were on track whenever they exceeded the boundaries of my investigation's objectives. Despite that, the three pilot interviewees provided enough answers to the interview questions asked, and their responses were accounted for in the main analysis.

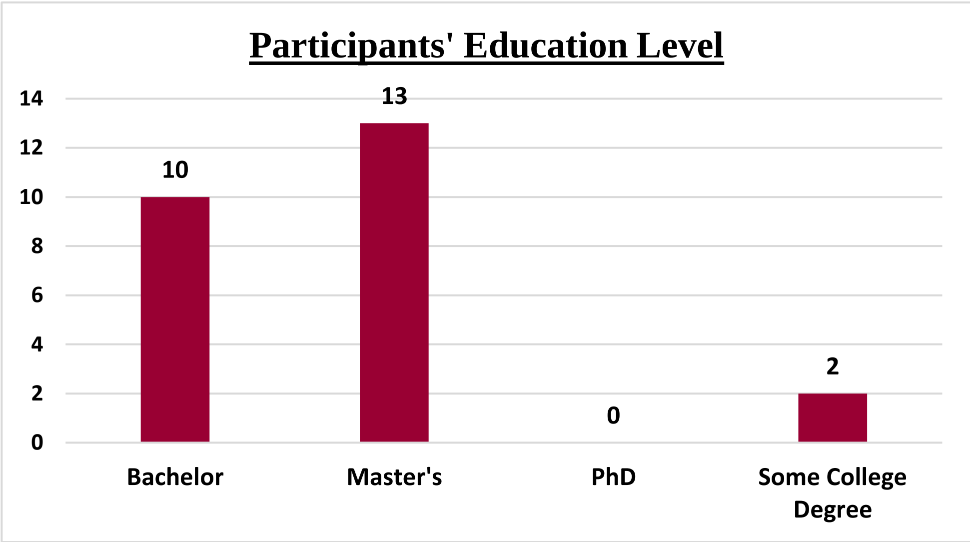
4.2. Main Study

After conducting the pilot interviews, the main study took place by interviewing 22 more participants during the period from July 2021 to August 2022. The sample consisted of employees and managers with hiring and recruitment experience. As mentioned earlier, participants were recruited through social media (specifically LinkedIn) and a snowball sampling technique, where interviewees recommended other prospective participants who could be interested in participating in my study. The sample of interviewees was as representative as possible in terms of age, gender, hiring experience, etc., as shown in the charts below.


Age Distribution
12
10
8
6
4
2
0
11
8
4
1
1
20-30
31-40
41-50
51-60
61-70
Total
Age


Men vs. Women
Sex
Man
Woman
12
13


Participants' Job Title
12
10
8
6
4
2
0
10
4
5
3
3
Senior Manager
Product/Program Manager
Director
Consultant
Others



Participants were employers of large organizations such as Deloitte, Google, Hatch, KPMG, and EY (see Appendix A for full details). All participants were compensated with a $20 Amazon gift card. The interviews took place during the COVID-19 pandemic and, thus, were conducted online over Zoom and lasted between 25-30 minutes. Prior to the interviews, participants were provided a consent form, a letter of information about the study, and the interview questions.

Like the pilot interviews, at the beginning of the main interviews, I welcomed participants and reviewed with them the information about the study. They were also provided with the AI definition as in the pilot study. Participants were then asked if they had any questions. After that, participants were shown the one-minute video that talked about a sample AI, which is now being used by many large organizations, such as Nike and Goldman Sachs to hire candidates. Showing participants the one-minute video impacted how they imagined the AI collaborating with them, as evidenced by their responses and the way the ensuing discussion was conducted. After watching the video, participants were asked to imagine that their organization wants to adopt a similar AI that would collaborate with them in hiring candidates. Then, participants were asked to answer the three interview questions discussed previously.

I stopped interviewing more people after reaching saturation (i.e., responses were repeated with no new insights). Hence, the total number of interviews conducted was 25 interviews. All interviews were recorded and transcribed for analysis purposes. The use of the proposed conceptual model helped guide the discussion during the interviews to focus on the positives and negatives of collaborating with AI. Participants were also sometimes asked to elaborate more on certain points that they raised. Either to clarify what they said or to let them provide more examples and freely express their opinions. An example is as follows:

Participant:*… I am a little nervous about the actual data points collected from the AI*.

Me: *can you please elaborate more on this?*

Participant: *I feel like it might exclude people from different cultures or people with disabilities that may be good candidates…. For example, if someone has autism, they may still be a good* *candidate for this role, but they don't have the same cues looking directly in the eye and that kind of stuff… I don't think that I can convince a candidate that doesn't even know me or the company yet to go through this, so I think I may lose candidates.*

To analyze the data collected, a content analysis approach was followed. Content analysis is a systematic and replicable approach for summarizing a body of text into fewer categories describing a phenomenon (Elo & Kyngäs, 2008; Stemler, 2001) as a means of providing knowledge and new insights (Krippendorff 2018). "*A category is a group of words with similar meaning or connotations*" (Weber, 1990, p.37). Content analysis is a suitable technique for this research as the main objective of the qualitative phase is to understand the main benefits and concerns of collaborating with AI and generate a set of potential constructs to be used in the research model proposed earlier.

Data was analyzed through phases. First, I analyzed the transcribed data from the first ten interviews (i.e., not including the pilot) iteratively by alternating between the interview scripts and the extant literature on AI and its related topics (e.g., chatbots). In this phase, I generated a list of general categories to come up with themes that reflected the interviewees' views. In the subsequent interviews, I succeeded in collecting twelve more perspectives that helped inform the general categories (see Table 1).

Table 1: List of General Categories

Benefits:
1. Saving time 2. Increased chance to reach a wider pool of candidates 3. Ability to screen more resumes 4. Ability to produce shortlisted candidates faster than humans

5. Less chance of human error when screening candidates 6. Increased ability to synthesize data 7. Removing much of the paperwork 8. Faster access to any resume or to recall any application's summary 9. Freeing recruiters to focus on other tasks 10. Can work anytime and anywhere 11. Ability to analyze data beyond humans' abilities (e.g., facial expressions) 12. Easier to give AI instructions than to a human (i.e., less chance of argument) 13. Increased objectivity 14. Increased Fairness 15. Less chance of evaluating resumes differently from each other 16. AI can justify why it selected certain candidates at anytime 17. Less chance of unconscious bias based on colour, race, ethnicity, gender, name…etc 18. Increased chance to give everyone equal opportunity
Concerns:
1) Inability to sense humans' feelings, moods, and emotion. 2) The quality of AI recommendations must be checked by a human first. 3) Loses the sense of getting into a conversation with a human, including the follow-up questions. 4) Humans may not be comfortable talking to machines. 5) What if the ethical policy and the culture of the company that developed AI in a certain country are different from the country where the AI will be used? 6) Not knowing how the AI makes decisions would make it difficult to know how the candidates would match the values and culture of the organization. 7) How would the AI measure work ethics? 8) Inability to predict how the AI would fit with the team spirit.

9) AI may learn from a very narrowed list, resulting in excluding good candidates. 10) How will it evaluate disabilities and people who are not comfortable with the technology? 11) Technological errors are inevitable. 12) Reliability of the technology in terms of system glitches, ongoing maintenance issues, system failure, etc. 13) Less control over how AI keeps my information, conversation…etc

After that, I downsized the list of general categories into fewer themes for the benefits and concerns (see Table 2). Example comments that pertain to the benefit and concern themes are provided in Appendix B. Next, another researcher familiar with qualitative research and the IS literature was given all the transcribed scripts and was asked to independently analyze the transcripts and produce fewer themes to double-check and enhance the robustness of results and strengthen the findings through triangulation, as suggested by (Benbasat et al. 1987). The list of generated themes is shown below:

Table 2: Themes Generated by the Two Researchers

	Principal Researcher	**Second Researcher**
Benefits:	1. Efficiency 2. Speed 3. Accuracy 4. Objectivity 5. Transparency	1. Efficiency 2. Complementarity 3. Integrity 4. Accuracy
Concerns	1. Lack of human touch/involvement/interaction 2. Compatibility 3. Privacy	1. Need for human interaction 2. Technical reliability

The two researchers then met to discuss their respective analyses. In this stage, differences in the coding were examined to reach a common understanding and satisfy the intercoder reliability. After discussions, the coders had a complete agreement in grouping the benefits and concerns categories into the themes listed in Table 3 (Cohen's kappa (k) = 1). For the benefits, the second researcher and I agreed that "efficiency" is the number one benefit. We also agreed that interviewees perceive AI to be more objective and able to make decisions with integrity. Given that "integrity" is a broader term that encompasses objectivity, fairness, and reduced bias (Przegalinska et al. 2019; Whang and Im 2018), we agreed to include "integrity" as the second major benefit. Speed, complementarity, and accuracy were seen as benefits related to being able to complement humans' abilities and accomplish things faster with minimized inaccuracies. Therefore, we both agreed to group them under the "efficiency" umbrella, which entails saving time, minimizing inaccuracy, and optimizing resources (Andrade and Tumelero 2022). Transparency was not included in the benefits since it is dependent on the type of AI utilized, and explainability as a way of making AI more transparent is a contextualized construct that will be studied in the quantitative phase of this study.

For the concerns, we both agreed that the absence of human touch when collaborating with AI is of great concern to interviewees. With all the AI advancements, participants still find it challenging to take humans out of the loop when accomplishing tasks. Hence, the "lack of human interaction" is listed as the first top concern. The second concern elicited from the interviews was the risk that the AI would produce outcomes or make decisions that do not fit the needs and values of the organization. Thus, after a discussion with the second researcher, "incompatibility" of AI decisions was agreed upon as the second salient concern since it was not listed in the second researcher's list of generated themes. The two other concerns listed in the final list of themes were "technical reliability", as some interviewees believe that AI is still a technology that would have its own glitches, and "privacy" as AI is able to record and learn from any information it comes across (Abdelhalim et al. 2019) which might put the privacy of individuals at risk.

Table 3: List of Final Themes

	Theme
Benefits:	1. Efficiency
	2. Integrity
Concerns	1. Lack of human interaction
	2. Incompatibility
	3. Technical reliability
	4. Privacy

After agreeing on the final themes, they were shared with two other information systems research experts for re-evaluation for consistency and to ensure that the most salient benefits and concerns from the interviewees' viewpoints were reported. As a result, the researchers agreed on the benefits; however, there was a debate about whether to include two of the reported concerns: technology reliability and privacy. On the one hand, the researchers agreed that technology reliability could be an issue with AI as well as with any other technology. Upon a further review of the transcripts, participants seem to be discussing this concern mostly for the candidates' sake (e.g., what if someone wears glasses or a mask). Lastly, only three participants raised this issue as a concern that would hinder their collaboration with AI.

On the other hand, privacy was the least concern among all respondents. Only one participant was concerned about their privacy and that the AI may share the participant's data and information without consent. Moreover, we do not expect that the AI would have a different level of access to employees' data or information compared to the human-human setting. Rather, we would expect that the AI might have more limited access since it could be programmed according to some restrictions. Furthermore, other participants clearly stated that privacy would not be a concern as long as the AI provides good recommendations and helps with the decision-making process, as exemplified by the following comments: "*Privacy is not a concern because if the AI would be better when learning from my data and my information, it is then a fair tradeoff*"; "*I guess AI would be using your interactions to improve itself and to help you make better decisions and things like that, so you have to let it do that*"; and "*As long as it solves a problem, my privacy would not be of a concern*." Therefore, privacy was not included as a distinct concern construct in our quantitative investigation.

After reaching a consensus on the final themes, findings were then shared with five participants. All agreed with the results of the study. The results of our qualitative study indicate that the two most salient perceived benefits are "efficiency" and "integrity." While the two most significant perceived concerns are "the lack of human interaction" and "incompatibility." These results inform the next component of our research investigation, where we develop and validate a model to understand humans' willingness to collaborate with AI.

4.3. Validation

Similar to quantitative research, it is important to validate qualitative research as it helps create a standard and common body of knowledge. Venkatesh et al. (2013) defined validity as "*how accurately the findings represent the truth in the objective world*" (p.32). Several aspects were considered throughout the qualitative journey to ensure a high-quality process. First, in the qualitative phase of my study, I was keen on interviewing a representative sample who work in different organizations to gauge the most salient benefits and concerns from different views. I stopped when saturation was reached and conducted multiple rounds of analyses to ensure the validity of the findings. Second, the validity of my qualitative part was assessed against the measures highlighted by (Venkatesh et al. 2013): a) design validity, b) analytical validity, and c) inferential validity, and the principles proposed by (Sarker et al. 2013). Even though these measures have interrelated components, I will discuss the elements that are relevant to my research design and how I addressed them in my qualitative phase.

"Design validity" in qualitative research resembles the internal and external validity in quantitative research. It refers to how the qualitative study is well designed and executed so that the outcomes are believable (or credible) and generalizable (or transferable). There are three main facets of design validity in qualitative research: 1) descriptive validity, 2) credibility and 3) transferability. Descriptive validity means how competent the qualitative researcher was in accurately describing and conveying events, contexts, objects, and so on. Introduced by (Maxwell 2002), descriptive validity stresses that qualitative researchers should report what participants exactly say, ensuring that any statement a participant makes is well-heard and not misunderstood or mistranscribed. Qualitative researchers should not omit any feature of a participant's speech that may reveal some emotions (i.e., stress) or aspects that convey the context of a discussion since this may threaten the descriptive validity of the research and take away some of the contextual factors of the issue being studied (Maxwell 2002). To ensure descriptive validity is achieved in my qualitative phase, all interviews were recorded and transcribed for future reference after getting participants' consent to ensure no important idea or thought was omitted or misheard. Second, in

analyzing and reporting what participants said, I tried not to remove any factor that touched upon context, emotion, or anything that would be of importance while analyzing and reporting results. Third, when a particular statement or discussion was not well-heard or well-understood, I either asked participants to repeat what they said or I repeated what they said so that they could correct any part that was misinterpreted.

The second pillar of design validity is credibility. Credibility in qualitative research reflects the notion of "internal validity" in quantitative analysis, which indicates the extent to which the observed findings explain the truth about a certain phenomenon in the targeted population and rule out any alternative explanation for it (Boudreau et al. 2001; Venkatesh et al. 2013). However, since truth in qualitative research is informant-oriented rather than researcher-defined, Guba and Lincoln (1982) suggested referring to internal validity in qualitative research as "credibility." Hence, credibility refers to the extent to which the outcomes of a qualitative study are believable or credible from the informants' perspective so that they rule out any alternative explanations. A qualitative study is deemed to be credible when it produces truthful explanations or interpretations of human experience such that the people having such experience would directly realize it from those explanations as their own (Sandelowski 1986). This was achieved in this study in two ways: first, the analysis of the qualitative findings was done through an iterative process. I tried to ensure that what participants elicited during an interview was consistent with the body of knowledge available in the literature. Second, I conducted member checking as a prominent validation technique in qualitative research (Creswell and Creswell 2018). Member checking is the process of following up with participants about the qualitative research findings to seek their feedback about the interpretation of those findings (Motulsky 2021). In this research, the outcomes of the 25 interviews were summarized and sent out to five different participants, who agreed that the findings of the qualitative study were representative and covered the main points they raised during the interview.

"Transferability" is the third element of design validity, which replicates "external validity" in quantitative research. It refers to the extent to which the findings of the qualitative study could be transferable (or generalizable) to other contexts or settings (Maxwell 2002). Although generalizability is more limited in qualitative research than quantitative studies, I tried not to be very specific about a certain AI type in this study. I emphasized viewing AI as a collaborator in the workplace that could collaborate with them to accomplish a predefined goal. However, to put this into context, we chose the recruitment domain for a couple of reasons. First, almost every organization has an HR department and goes through a recruitment process for some job vacancies. It is a domain that is not pertinent to a particular industry or requires specific conditions to exist. Whether it is a large, medium, or small organization, a recruitment process is necessary to some extent. Second, the HR domain replicates many organizational domains that require collaboration among colleagues to achieve a certain goal (i.e., hiring the best candidate). Third, the recruitment process itself is a well-defined process that does not differ much from one organization to another in order to hire a candidate for a certain job. This means that the findings of this research could be easily transferable to other domains that entail collaboration among workers with some minor adjustments to the context.

The second validity type emphasized by (Venkatesh et al. 2013) is "analytical validity, "*which refers to the way qualitative data was collected and analyzed such that results are dependable, consistent, and plausible*" (i.e., as opposed to measurement validity in quantitative research). In this study, I followed a systematic approach to collecting my qualitative data. The same exact steps, wording, and questions were followed in the same sequence across all interviews to ensure consistency. In order to affirm the plausibility of findings, analyzing the data was an iterative process that kept me going back and forth with the literature to confirm that results are plausible with a theoretical premise. Moreover, another researcher who is familiar with qualitative research and the IS literature was asked to analyze the data independently to compare and discuss the consistency and plausibility of findings.

Next, "inferential validity" was examined. It refers to *"the quality of interpretation that reflects how well the findings can be confirmed or corroborated by others*" (Venkatesh et al. 2013, p.34). As mentioned earlier, all interviews were transcribed and coded without any omission of any context or emotions. Moreover, whenever a sentence or opinion was misunderstood or misheard, I asked participants to restate what they shared by repeating what they said or by asking them to provide more explanations and examples to support their arguments. The recordings of the interviews also helped in revisiting any part that was not clear to ensure that the interpretation of findings reflected what the interviewees shared during the interview.

Besides the above, I assessed the validity of my qualitative phase through the principles provided by (Sarker et al. 2013) that were relevant to my study: 1) the principle of internal coherence as echoed in this study as I took a data-centric stance rather than being imaginative during the analysis phases, and I had to follow a neutral scientific presentation style to achieve coherence across the interviewing process and elements, 2) the role of IT was clear and significant in my study by focusing on AI as the IT artifact and engaging in discussion with practitioners to gain a practical and realistic understanding of the phenomena, which supports the principle of relevance, 3) in demonstrating how the qualitative phase of my study was carried out, I tried to be as transparent and detailed as possible through reporting <u>where, when, how, and from whom</u> data was collected, and how data was analyzed and inferences were made to support the principle of transparency.

Summary

This chapter demonstrated the qualitative phase undertaken as the first stage in the mixed method approach. The process of data collection and analysis was discussed, and the different ways of validating the qualitative findings were reviewed. In a nutshell, the two main benefits elicited from the qualitative study are *efficiency* and *integrity*, and the two main concerns are *lack of human interaction* and *incompatibility*.

Chapter 5: Research Model and Hypotheses

As discussed in Chapter 4, findings from the qualitative phase revealed that efficiency and integrity are the two most important benefits from the interviewees' perspectives and that lack of human interaction and incompatibility are the two most significant concerns when collaborating with AI in an organizational setting.

Integrating AI as collaborators in the workplace is a new phenomenon. It differs from the use of traditional information technologies that are merely facilitating tools. The literature lacks sufficient empirical research that examines how contextualized AI characteristics would influence people's perceptions about emerging technologies such as AI that can be collaborators in the workplace and how people's beliefs about the benefits and concerns impact their willingness to collaborate with AI. In this study, the definition of AI collaborator is adapted from (Rai et al. 2019), and is defined as "*an AI that can perform cognitive functions that we normally associate with human minds; can work autonomously; can interact with and learn from humans; can adapt to different situations; and can make proactive, predictive, or personalized decisions*." Willingness to collaborate refers to evaluating human partners' attitudes and intentions towards concrete collaboration situations (Rosas and Camarinha-Matos 2010). Thus, the willingness to collaborate in this study is defined as "*the readiness of a human to accept and jointly work with an AI as a collaborator to complete a task*" (see Table 4 for a full list of definitions).

To address this gap, the research model shown in Figure 6 is proposed to empirically test the influence of AI autonomy and AI explainability on peoples' beliefs. As well as how these beliefs can impact people's willingness to collaborate with AI. AI autonomy and AI explainability are two essential aspects to study in the field of artificial intelligence due to their significant implications on society and other critical domains. On the one hand, AI autonomy mainly refers to the ability of AI to make decisions on its own without human intervention (Nickerson and Reilly 2004). As AI advances, there is an increasing interest in developing AI systems that can operate autonomously in various applications, including self-driving cars, autonomous drones, medical diagnosis, and financial trading systems, among others. The decisions made by autonomous AI can have real-world consequences, and ensuring that these systems operate predictably and morally is vital. The deployment of autonomous AI in various industries may also raise legal and regulatory challenges, and it is important to examine how it would affect people's perceptions toward AI. In addition, understanding the balance between AI autonomy and human involvement is paramount to creating effective human-AI partnerships. Designing AI systems that can complement human skills and decision-making rather than replacing them is essential nowadays.

On the other hand, the ability of AI to provide understandable and transparent explanations for their decisions and actions is becoming a key feature. As AI is increasingly used in high-stakes domains, understanding the factors influencing AI decisions is crucial from an ethical and legal standpoint. For example, explainability is essential for ensuring that AI-generated medical recommendations can be justified and understood by healthcare professionals and patients. AI explainability can also help identify and mitigate biases in AI systems, ensuring fair treatment and reducing the risk of perpetuating discrimination. Thus, by studying AI autonomy and AI explainability, researchers, developers, policymakers, and users can work together to harness the full potential of AI collaborators while ensuring that it operates efficiently, fairly, and transparently in various domains.

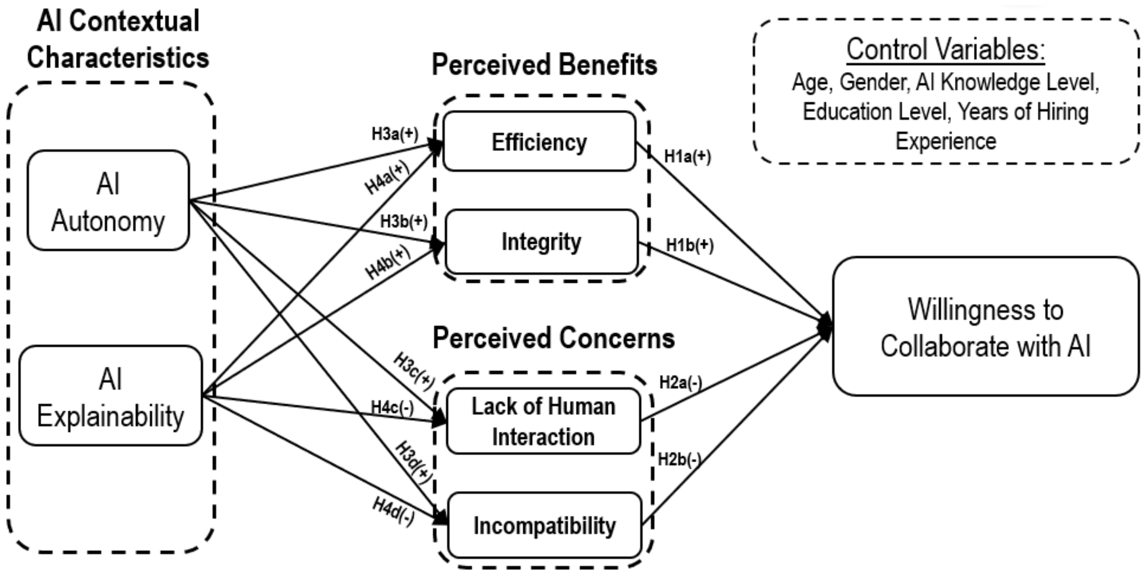


Figure 6: Proposed Research Model

Table 4: List of Definitions

Construct/Concept	Definition
AI Autonomy	The extent to which AI works independently and makes decisions on its own without human intervention
AI Explainability	The extent to which AI provides explanations about its recommendations/decisions
Efficiency	The ability to save time and speed up the decision-making process
Integrity	The extent to which an AI's recommendation is perceived as being objective, unbiased, and fair.
Lack of Human Interaction	The absence of interacting with a human when collaborating with AI

Incompatibility	The risk that an AI might be inconsistent with and is unable to provide recommendations that align with the organization's and/or humans' needs and values
Willingness to Collaborate	The readiness of a human to accept and jointly work with an AI as a social actor to complete a task

Based on ANT, NVT, the qualitative study findings, and the contextualized AI characteristics discussed above, we advance specific hypotheses for each contextualized benefit and concern related to collaborating with AI as determined from the qualitative analysis, namely, efficiency and integrity as benefits and the lack of human interaction and incompatibility as concerns.

5.1. Impact of Perceived Benefits and Concerns

In this section, the benefits and concerns of the qualitative study are discussed in more detail to formulate the research model hypotheses.

5.1.1. Efficiency

Andrade and Tumelero (2022) define efficiency in a service context as "*the highest performance in meeting human needs or the productivity of equipment, methods, actions or processes to achieve maximum gain with minimum inaccuracy, resource waste, strength or procedures*" (Andrade and Tumelero 2022, p.242). In this study, I define efficiency as "*the ability to save time and speed up the decision-making process.*" The non-refutable memory of AI, its capability to learn from and quickly process large amounts of information, work tirelessly, and discover hidden patterns beyond humans' abilities, enable AI to increase work efficiency and reduce human efforts and potential errors that might take place when working with humans. For example, in Norway, a key public service organization relies on chatbots to handle routine customer service requests. These chat-based AI can handle multiple chats and interact with many citizens simultaneously, which positively contributes to increased service efficiency (Vassilakopoulou et al. 2023). Rzepka et al. (2020) also found that individuals save a lot of time when using voice assistants than when performing their purchase activity by themselves. This is because customers are not required to open the app and select a product, so the buying process is expected to be much quicker.

This perception was evidenced through participants' comments in my qualitative phase, such as "*AI can downsize the number of applicants to go through being able to prioritize things in a more efficient manner*..", "*…avoidance of human error and ability to synthesize data to simplify things*" This added assurance of people's perception that AI can speed up completing tasks and save time should positively influence their willingness to collaborate with such technologies. Hence, I hypothesize that people would be more willing to collaborate with AI who can speed up the decision-making process and increase work efficiency:

***H1a**: Perceived task efficiency when collaborating with AI will have a positive association with one's willingness to collaborate with it*

5.1.2. Integrity

Integrity is a pillar in shaping an individual's trust (McKnight et al. 2002). In an organizational context, integrity is defined as "*the belief that an organization is fair and just*" (Przegalinska et al. 2019, p.788). In the context of recommender systems, Whang and Im (2018) define integrity as "*the extent to which the recommender system's advice is perceived to be* *unbiased*" (Whang and Im 2018, p.948). For decision support technologies, integrity is defined as "*the perception that these technologies adhere to a set of principles (e.g., honesty and keeping promises) generally accepted by consumers*" (Wang and Benbasat 2008, p.251). McKnight et al. (2002) define integrity in an e-commerce context as *"keeping commitments and not lying"*(McKnight et al. 2002, p.339). Based on these definitions, I define integrity in this study as "*the extent to which the AI's recommendation is perceived as being objective, unbiased, and fair*."

Integrity in different contexts contributes to one's positive perception regarding different experiences. Integrity is also critical as it reduces uncertainty and potential risks (Bhattacherjee 2002). Whang and Im (2018) claim that online shoppers would believe that recommender systems are unbiased (i.e., integrity) since they are able to provide personalized recommendations, which would increase shoppers' intention to adopt their recommendations. In a team setting, group members' perceptions of their leader's integrity positively impact their behaviour (White and Lean 2008). In the context of mobile banking, Lin (2011) argues that the integrity of mobile banking institutions plays a vital role in shaping one's attitude toward the use of mobile banking since integrity conveys a sense of objectivity, which encourages customers with high integrity to have a positive attitude toward using mobile banking. Ochmann and Laumer (2019) also propose, after interviewing 21 experts from various fields, that perceived fairness influences AI adoption in recruiting candidates. It is, therefore, reasonable to assume that people prefer to collaborate with those with integrity who are honest, unbiased, and fair. Particpants' comments from my qualitative study also support this argument, such as "*…AI would evaluate everyone in the same way, and this would help eliminate at least the unconscious bias in the initial stage of screening candidates*". This leads to the following hypothesis.

***H1b**: Perceived integrity when collaborating with AI will have a positive association with one's willingness to collaborate with it*

5.1.3. Lack of Human Interaction

The evolution of AI disrupted many of the professions that were known for years to be undertaken by humans to enhance users' experiences and improve business processes (Cath et al. 2018). This disruption sometimes comes at the expense of abandoning what users got used to when accomplishing a task or receiving a service. In customer service, for example, customers never imagined that they would receive a service from a non-human agent. Subsequently, this may put them in discomfort and influence their satisfaction (Ashfaq et al. 2020) and loyalty (Rajaobelina et al. 2021). Through surveying 500 random consumers, 87% disclosed that they still prefer humans to chatbots for quick interactions (Press 2019; Yin 2019). The lack of human interaction in computerized service delivery is defined as "*the need that some individuals feel for interacting with the service employee in a service encounter*" (Dabholkar 1992, p.564). In the context of self-service encounters, such as chatbots, the lack of human interaction refers to "*the importance of human interaction to the customer in service encounters of individuals*" (Dabholkar and Bagozzi 2002, p.188). Thus, in this study, we define the lack of human interaction as "*the concern of losing human interaction when collaborating with AI.*"

Nguyen (2019) discussed that some firms may still feel reluctant to implement chatbots since users lack the human touch when interacting with them. Luo et al. (2019) found that when customers know they are conversing with a non-human partner trying to make a sales call for them, they tend to be curt and purchase less because they perceive the chatbot to be less knowledgeable and less empathetic. Several studies also showed that there is an inverse relationship between the lack of human interaction and consumers' behaviour toward self-service technologies (Collier and Sherrell 2010; Lee et al. 2010; Reinders et al. 2008) and their intention to use them (Lee 2017).

Participants in the qualitative study in this research have also raised the concern that collaborating with an AI may eliminate the human touch from the interaction, which would hinder their willingness to collaborate with it. Some of the excerpts that affirm this argument are "*Humans are not comfortable talking to machines", "AI has to match the energy of the human co-worker that makes them more comfortable .. the AI would be very transactional. People want to work in a company that they feel comfortable …",* and *"the feeling of another person… the human element is very important.*" Thus, we hypothesize that:

***H2a**: Perceived lack of human interaction when collaborating with AI will have a negative association with one's willingness to collaborate with it*

5.1.4. Incompatibility

The qualitative study listed "incompatibility" as the second top concern when collaborating with AI. Some individuals perceive AI as a black-box. This means that the inner workings of AI and how its components come together may not be known. Some participants mentioned that: "*Compatibility with the values and morals as a company would be a concern...it can be detected through asking AI certain questions to make sure it is compatible with the company's values"*, "*Culture-fit and team-fit assessment is not easy …. someone who is great for one company might not be good in another*", and "*Compatibility with values and ethics is a concern*." Consequently, deploying an AI that is created and developed by another organization, and sometimes in another country, may influence individuals' perceptions about how it processes data and provides recommendations. This may raise the concern that the AI would unintentionally make recommendations that are incompatible with the organization's values and beliefs.

According to the Diffusion of Innovation (DOI) theory, compatibility is defined as " *how innovation fits into the user's experience and needs and is consistent with existing values* (Rogers 2003). In the context of providing online government services, compatibility is defined as "*the degree to which an innovation is seen to be compatible with existing values, beliefs, experiences and needs of adopters*" (Carter and Bélanger 2005, p.8). Compatibility was found to be a strong predictor of citizens' intention to use e-government services (Carter and Bélanger 2005). This means that citizens are more willing to use e-government services if the services are compatible with their values. For the adoption of solar energy systems, compatibility positively influenced the adoption of such systems (Labay and Kinnear 1981) and was defined as "*the degree to which the innovation is seen as consistent with the innovator's existing values, past experiences, and needs*" (Labay and Kinnear 1981, p.272). In mobile bookings, compatibility is defined as "*the degree to which the mobile hotel booking technology fits the lifestyle and experiences of individuals*" (Ozturk et al. 2016, p. 1352). In this research, I conceptualize incompatibility as *"the concern that the AI might be inconsistent with and is unable to provide recommendations that align with the organization's and/or humans' needs and values*."

Since compatibility represents one's existing values and beliefs, empirical research demonstrated that compatibility significantly affects the intention to use technology (Lee and Lyu 2019). Hari et al. (2021) discovered that compatibility positively influenced customer brand engagement when banking chatbots were used, thereby influencing their satisfaction with the brand experience and customer brand usage intention. Compatibility also has effects on behavioural intention to use mobile technology and is considered to be one of the most significant factors influencing behavioural intention in a mobile shopping context (Wu and Wang 2005). Ewe et al. (2015) highlighted that the willingness to adopt mobile banking could be increased by offering and promoting various complementary services that give the impression that mobile banking is easy to use and is compatible with users' lifestyles.

When an innovation shows higher compatibility, it has a higher probability of being adopted (Eeuwen 2017; Rodríguez Cardona et al. 2019). Once an innovation is compatible, it is highly likely to be used, and hence, it becomes an important feature of such innovation (Rogers, 2003). Therefore, if AI produces incompatible recommendations, individuals would evaluate AI collaborators negatively. Such concern would adversely impact humans' willingness to collaborate with it. This leads to hypothesize that:

H2b*: Perceived incompatibility when collaborating with AI will have a negative association with one's willingness to collaborate with it*

5.2. Impact of AI Autonomy

Beale and Wood in 1994 described autonomous agents as "*agents that are able to work on behalf of their users without the need for any interaction or input from the user. They act without your presence, tirelessly performing tasks*" (Beale and Wood 1994, p.240). Autonomy is often observed as freedom from human intervention or control (Brown et al. 1998). This definition was later modified by Barber and Martin in 1999, who argued that autonomous agents should also take into account their goal, be able to make decisions about how to attain the goal, and act on these decisions (Barber and Martin 1999). Subsequently, they redefined autonomy as "*an agent's active use of its capabilities to pursue some goal without intervention, oversight, or control by any other agent*." (Barber and Martin 1999, p.8). This means that AI autonomy refers to the ability of artificial intelligence systems to make decisions and take actions independently. It involves the development of algorithms and models that enable machines to analyze vast amounts of data, learn from it, and make intelligent choices based on predefined objectives.

Furthermore, the US National Institute of Standards and Technology (NIST) discussed that systems are characterized by being fully autonomous if they can achieve the goals they were designed to achieve within a definite scope with no human interventions and can adapt to operational and environmental conditions (Ezenkwu and Starkey 2019). Nickerson and Reilly also defined machine autonomy as "*the ability of the machine to make decisions on its own*" (Nickerson and Reilly 2004, p.2). Therefore, and in the context of this research, I define AI autonomy as: *"the extent to which the AI collaborator works independently and makes decisions on its own without human intervention."*

Ten levels of machine autonomy were proposed by Parasuraman et al. (2000); such that the highest level is one where the machine decides everything and ignores the human. In contrast, the lowest level is when the machine can offer assistance while the human makes all the decisions. Given these levels, Nickerson and Reilly (2004) proposed that the level of machine autonomy will have an influence on humans' beliefs in that machine.

The emergence of these new innovative technologies, such as AI, has led many organizations and humans to relinquish many tasks to be completed by these technologies. Green and Chen (2019) discuss that many decisions are now made through an "*algorithm-in-the-loop*" process where algorithms that are based on ML algorithms have some level of control and can inform people. Given the unique nature and capabilities of AI, AI is able to instantly make predictions, handle very complex calculations, and improve inefficiencies much faster than the human mind (Deyo 2020). This improves humans' work by completing routine, repetitive tasks or analyzing overwhelming amounts of data for more informed decisions (Brynjolfsson and Mitchell 2017). For example, firms can divide various tasks between humans and robots to benefit from each other's strengths. Robots can take care of duties that need a fair amount of physical activity and are repetitive, while human workers can focus on tasks requiring human capabilities and judgment (Libert et al. 2020). Hence, assigning AI some levels of autonomy to complete a task on behalf of humans would save them time (Hua Ye and Kankanhalli 2018) and increase people's perception of efficiency (André et al. 2018). Therefore, I hypothesize that:

***H3a**: Increased AI autonomy when collaborating with AI will have a positive association with perceived efficiency*

Highly autonomous AI can increase work efficiencies and execute decisions and actions consistently without being influenced by emotions, unconscious biases, or external pressures. Although, in some instances, the AI may produce biased or undesirable outcomes if not trained well or if the data used to train it is not well-representative (Ebrahimi and Hassanein 2021), the perception that AI is characterized by consistency in making decisions can contribute to a perception of integrity because the AI is not subject to the same human vulnerabilities that can lead to biased or inconsistent decision-making (Polli 2019). In addition, highly autonomous AI bases its decisions on objective data-driven patterns. This objectivity can lead to a perception that the AI is making decisions based on rational analysis rather than subjective factors, thus enhancing its perceived integrity.

Many participants from my qualitative study corroborate this idea that when AI is involved in the hiring process, the decisions made would be more objective and free of bias, at least initially, and thus, would have more integrity. As some interviewees mentioned, *"AI can create some objectivity for everybody when evaluating them as some hiring managers might be unconsciously* *biased.", and "AI does not care about skin colour or ethnicity….it skips all the biases the humans might have .. but the final decision should still be mine."* (see Appendix B for more quotes). Thus, in the context of this study, the following hypothesize is advanced:

***H3b**: Increased AI autonomy when collaborating with AI will have a positive association with perceived integrity*

Several studies claim that algorithms usually outperform humans even if they make mistakes or produce some errors (Grove et al. 2000). Despite this outperformance, people may prefer humans' forecasts to algorithms' forecasts (Diab et al. 2011; Eastwood et al. 2012). Furthermore, in some cases, people create a negative perception of algorithm-based decision aids when they are used to seek external advice since they are nonhuman tools (Shaffer et al. 2013). People may also attribute greater attention and weight to the advice given by a human than by an algorithm (Önkal et al. 2009). Some may also prefer to rely on their own or other people's judgment and exhibit less tolerance for errors made by algorithms than errors made by other people (Dietvorst et al. 2015; Green and Chen 2019). They may even reject fully automated decisions or AI systems in fear of replacement (Huisman et al. 2021) or other threat-faced emotions (Hornung and Smolnik 2022; Akmeikina et al. 2022).

This means that even if algorithms show great performance, humans may still feel they are in need of human touch and interaction when collaborating on a certain task.

Besides, Mnih et al. (2015) highlight that adopters of AI systems should be given additional control to foster the acceptance of the system. Furthermore, Eilers et al. (2020) emphasize that employee empowerment is a critical factor in improving employee work effectiveness. Akmeikina et al. (2022) discuss that in the era of technological intelligence, people continue to be empowered through their skills since AI systems are not yet mature enough to think exactly like humans. Therefore, we hypothesize that:

***H3c**: Increased AI autonomy when collaborating with AI will have a positive association with perceived lack of human interaction*

While AI autonomy holds great promise in various fields, such as healthcare, transportation, and finance, it also raises concerns about incompatibility with human values and needs (Muggleton et al. 2021). The qualitative data collected in this study showed that incompatibility of the AI recommendations would be of concern to humans who use it since it may not match the needs and values of the organization and/or the team. As AI systems become more sophisticated and capable of independent decision-making, they may prioritize objectives in ways that are misaligned with human values or societal norms. This misalignment can arise due to ignoring the contextual and subjective nature of values across cultures and individuals, and lack of transparency in algorithmic decision-making (Ajunwa 2020), or a misinterpretation of human preferences. Thus, I hypothesize that:

***H3d**: Increased AI autonomy when collaborating with AI will have a positive association with perceived incompatibility*

5.3. Impact of AI Explainability

As discussed earlier, the issue of using black-box models in building an AI presents the need to develop a more transparent and interpretable algorithm that can provide explanations for their outcomes (Rader et al. 2018). In non-mathematical terms, (Biran and Cotton 2017, p.1) discussed that "*systems are interpretable if their operations can be understood by a human, either through introspection or through a produced explanation*." Inspired by Biran and Cotton's definition, (Miller 2019, p.8) defined interpretable systems as "*the degree to which a human can understand the cause of a decision*." Adadi and Berrada (2018) also mentioned that explainable AI is a research field that attempts to create AI systems whose outcomes are understood by humans.

In 2016 the European Union imposed a new regulation as part of the General Data Protection Regulation (GDPR) that calls for the "*right to explanation*" (Casey 2018). This new type of regulation was imposed due to the accelerated development of intelligent machines and ML algorithms that are now used to make automated decisions. Therefore, the European Union urged that in order to ensure fairness and transparency in processing personal information, its citizens should have the right to know – in a meaningful way – the logic behind automated decisions that affect them. The U.S. Department of Defense is also directing efforts toward developing "explainable AI" systems that are able to translate decisions resulting from complex algorithms to a language humans can understand (Castellanos 2018). As a result, in this study, "explainability" is defined as:*" the extent to which the AI collaborator provides explanations about its recommendations/decisions."*

Gregor and Benbasat (1999) emphasized the importance of the ability of intelligent systems to justify and provide explanations for their actions. The authors suggested that this feature can enhance people's beliefs about such intelligent systems. For example, if intelligent agents are able to explain what they do and why they do it, people might trust them more. Moreover, the authors suggest that the use of explanations influences efficiency and effectiveness (Gregor and Benbasat 1999). This conclusion was justified by arguing that explanations improve systems' transparency and help to transfer knowledge more efficiently, resulting in greater system effectiveness.

When building recommender systems or ML systems, (Kulesza et al. 2015) argue that the ability of such systems to provide explanations about their recommendations and reasoning helps end users understand the internal mechanics of these systems and enables them to better customize these systems (Bostandjiev et al. 2012). Without explanations, end users would find it inefficient and challenging to accurately modify such systems (Lim et al. 2009). Kleinberg et al. (2018) also discuss that humans sometimes find it difficult to justify their decisions or may not be aware of their own unconscious biases. Hence, they may look for explanations to rationalize their choices. Moreover, Herlocker et al. (2000) advocate that explanations facilitate the process of handling any errors that might come with a recommendation from intelligent recommender systems, which leads to greater efficiency. Yang et al. (2021) also agree that some humans need AI explanations to improve the efficiency of their decision-making process when collaborating with AI. Thus, we hypothesize that:

> ***H4a****: AI explainability when collaborating with AI will have a positive association with perceived efficiency*

Wang and Benbasat (2007) found that the use of trade-off explanations generated by recommendation agents significantly and positively influences users' integrity beliefs, where trade-off explanations inform users about the benefits and potential costs of different product features. Such balanced information delivers an image of objectivity and fosters users' integrity perceptions toward recommendation agents. Herlocker et al. (2000) discussed that recommender systems are black boxes that, if a recommender system is able to provide explanations about the reasoning and rationale of its recommendations, it would be perceived to be more transparent. The authors also recommend that adding an explanation facility into a recommender system improves users' sense of involvement in the recommendation process, enabling them to add their knowledge and expertise to complete the decision-making process.

Shin (2021) conducted an experiment with 350 students to examine the effects of explainability on perceived fairness and transparency, among other variables. Participants were given an algorithmic-based news website where they could surf, view, and read automatically generated news. They were told that the news content and recommendations were generated through ML algorithms. They were also given explanations about why they were shown certain recommendations. The authors found that explainability had a positive influence on perceived fairness and transparency of the algorithmic-based website (Shin 2021).

From the qualitative study, objectivity in making decisions is among the benefits that humans expect from an AI collaborator. Humans need objective information or explanations that would influence their judgment and final decisions. Objectivity was also found to be among the issues that consumers care about when using recommendation agents in virtual shopping (Komiak et al. 2004). We argue that when AI provides explanations to justify its decisions and outcomes, humans perceive the AI to be objective, unbiased and fair, resulting in overall more integrity. Therefore:

H4b*: AI explainability when collaborating with AI will have a positive association with perceived integrity*

As discussed in Chapter 2, exchanging knowledge and explaining actions is crucial in human-human collaboration. To encourage such exchange, some organizations may compensate workers who share their knowledge while punishing others who refrain from doing so (Bartol and

Srivastava 2002). Firms also develop online platforms that enable them to reach subject matter experts from around the world to solve their problems (Dissanayake et al. 2015). Herlocker et al. (2000) advise that building recommender systems with explanatory interfaces improves users' involvement and their feeling that they are kept in the loop when making a decision. Researchers also agree that understanding how humans explain to each other can serve as a basis for building and designing explainable AI (Miller 2019; Wang et al. 2019). They suggest that explanations help people to learn by extracting the distilled knowledge in order to predict and influence future phenomena (Miller 2019; Yang et al. 2021).

Moreover, sometimes humans seek to consult other humans to learn and augment their knowledge or to verify a piece of information. The use of search engines, knowledge bases, and AI played a great role in eliminating the need for this human-human consultation. ChatGPT, as an example, is able to comprehend and generate any text to communicate fluently like humans (Teubner et al. 2023). Meta's Chief AI Scientist "Yann LeCun" also mentioned that there are ''*half a dozen startups that basically have very similar technology*'' (Ray 2023). Such technologies that can interact and explain different phenomena are reducing the need for human-human interaction and consultation in many situations. Decision-makers would utilize explanations from such systems to scrutinize and debug any errors or to make the necessary decisions (Wang et al. 2019).

Therefore, when humans collaborate with an AI that is willing to provide explanations about its decisions, this would reduce the need to consult another human to discuss the rationale of a certain AI recommendation. That said, humans' perception that they would need to interact with other humans would decrease if the AI were able to explain its actions. Therefore, we hypothesize that:

H4c*: AI explainability when collaborating with AI will have a negative association with perceived lack of human interaction*

Participants from the qualitative study were concerned about the compatibility of the AI outcome with the organization's values. Without explanation, this concern will be greater. However, when an AI is able to provide an explanation about the factors it considers to make a decision or produce an outcome, humans will be less concerned or at least will be able to investigate the elements that led to an incompatible outcome. Yang et al. (2021) highlight that compliance with regulation is a fundamental reason for seeking to improve the explainability of AI systems. The European Union regulated the notion that AI systems should grant individuals "*The right to explain*," where individuals subject to a decision made by an AI system should be given an explanation of why the AI made a specific decision (Goodman and Flaxman 2017). Providing explanations can also help humans understand what could be changed to obtain a better result (Yang et al. 2021) that can comply with organizations' policies and values. Therefore, the more the AI is able to provide explanations, the less people would be concerned about the incompatibility of the outcome since it would be transparent to humans the logic the AI considered when making a decision (Rossi 2019). This leads to the following hypothesis:

H4d*: AI explainability when collaborating with AI will have a negative association with perceived incompatibility*

Summary

This chapter outlined a research model and conceptualized the variables used in this study. The model consists of eleven main hypotheses deliberated from the general framework provided earlier in Chapter 1. Specific contextual AI characteristics are added, and beliefs about collaborating with AI (i.e., perceived benefits and concerns) are included. Details about the quantitative study are discussed in the next chapter.

Chapter 6: Methodology and Results of the Quantitative Study

This chapter describes the methodology used to validate the research model presented in Chapter 5. Details about The methodology employed a between-subjects experimental survey design. Results from this experiment are presented at the end of this chapter.

6.1. Task Overview

One of the most promising uses of AI in organizations is its utilization in the hiring and recruitment process. A wide range of applications are already in the market (e.g., Knockri, HireVue, Seedlink, Gecko), with an increasing trend towards using AI to screen millions of candidates and applications, and recommend a shortlist of those who present the best fit. It has been argued that recruiters ignore 65 % of resumes submitted for a job posting receiving a receiving a high volume of applications and that AI can play a vital role in solving this problem (Min 2016). Unlike typical, unintelligent HR software programs, AI that uses ML can also screen, rate, compare, and rank every resume instantly (Min 2016).

Figure 7 describes the major phases of a typical recruitment process in different organizations such as Unilever, Nike, Goldman Sachs etc. The process starts by posting a job opportunity for which organizations seek to find the best candidate. After that, prospective applicants fill out an application form, upload their resumes, and submit a self-recorded video where they answer a set of predetermined questions. Recruiters in organizations then examine each candidate's resume and application form and watch the videos that the applicant submitted. After this, recruiters create a shortlist of the best candidates and invite them for a personal interview.

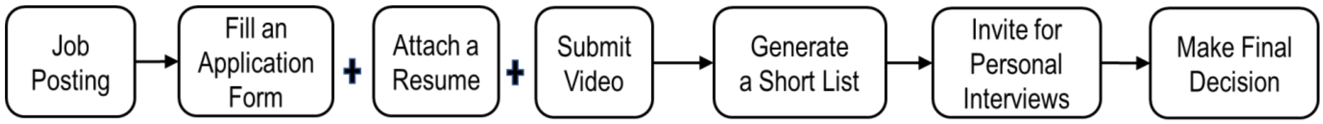


Figure 7: Steps of a Typical Hiring Process

After obtaining the research ethics clearance (MREB#: 5134), participants were invited through a market research firm to take part in a 2X2 factorial design scenario-based survey. The scenario of the study was designed in a way that imitates the real recruitment process described in Figure 7 but with the help of an AI called “AI-Assist”. The reason why I am applying a scenario-based study and not using a real AI is two-fold: 1) to eliminate any potential confounding effects (e.g., potential biases towards specific companies or AI products), and 2) to control the amount of time necessary for participants to complete the process. With this design, I tested the hypotheses through a between-subjects experimental approach.

At the beginning of the survey, a set of screening questions was used to filter out participants who were not eligible to complete the survey (see Appendix C). An ineligible participant is anyone with no hiring experience. After the screening phase, eligible participants were given a consent form that outlined general information about the study and what was expected from them during the study. Then, demographic information was collected from participants. After answering the demographic questions, all subjects were directed to watch a one-minute video about how AI could be used in the hiring process and the different ways it can process candidates’ profiles. Subsequently, participants were asked to imagine a scenario where their organization is contemplating the adoption of an AI called “AI-Assist” that can collect and analyze candidates' profiles and their recorded videos. They were asked to imagine that their organization is asking them to collaborate with this "AI-Assist" to recruit the best candidate for a customer service job.

A customer service job was selected in this scenario for different reasons. First, we assume that every participant is aware of what a customer service job is and the basic requirements that should be present in a candidate's profile to fill the job. Second, a good-fit candidate for a customer service job should possess a combination of qualities (e.g., body language, tonality, facial expressions, the language used, etc.) that AI can automatically process and analyze. Third, in a customer-facing job like a customer service representative, we argue that human judgment is still critical and that we are still far from relying only on an AI to hire candidates that will interact with customers. Last, employees always look to work with collaborative, friendly, and respectful people. Thus, having a human in the loop when recruiting candidates is still essential in organizations, and hiring a customer service candidate is an example of a job that would still require human input. Next, participants were randomly assigned to one of four treatment conditions where AI autonomy and AI explainability were manipulated (see Table 5).

Table 5: Treatment Groups

Treatment group	Autonomy Level	Explainability Level	No. of Participants
1	High	High	89
2	Low	High	89
3	High	Low	92
4	Low	Low	87

6.2. Treatment Conditions

Four treatment groups were used in this experimental study. Variations between groups were based on manipulating both the autonomy and explainability of "AI-Assist." Autonomy was manipulated as to who would make the final decision. A final decision in our context is defined in terms of who will be responsible for inviting shortlisted candidates to a personal interview. In the high autonomy condition (i.e., autonomous AI), AI-Assist screens candidates' applications, creates a shortlist of the top candidates and automatically decides and invites those shortlisted candidates for a personal interview. Also, since autonomy involves the freedom to choose from a variety of options unrestrictedly (Newman et al. 2022; Wertenbroch et al. 2020), autonomy was also manipulated in terms of the level at which the human recruiter has access to any other candidate's profile and not to be restricted only with what the AI provides. Therefore, in the high-autonomy condition, participants were only given access to the resumes and videos of the shortlisted candidates. They did not have the option to review other applicants' information who applied for the job.

In contrast, in the low autonomy condition (i.e., assistive or augmenting AI), AI-Assit screens candidates' applications and creates a shortlist of the top candidates. Then, the human recruiter takes a look at the shortlisted candidates so that he/she can decide whom to invite for a personal interview. Unlike the high-autonomy condition, participants in the low-autonomy condition were given access to the resumes and videos of the shortlisted candidates as well as access to any other applicant's information who applied for the job and was not shortlisted. So, in the low autonomy condition, the human recruiter makes the final decision, while in the high autonomy, "AI-Assist" makes the final decision.

For AI explainability, participants in the high explainability condition (i.e., explainable AI) were provided with an explanation for the AI-Assist's decision that justified why AI-Assist shortlisted certain candidates. In the low explainability condition (non-explainable AI), AI-Assist did not provide any explanation for its decision. A summary explanation of the four treatment groups can also be found in Table 6.

Participants in each condition watched a two-minute video that manipulated AI autonomy and AI explainability. In each video, participants watched and listened to a pre-recorded and written two-way conversation that took place between a human recruiter and AI-Assist. All conversations were written in a chat-dialogue format so that participants were also able to read the whole conversation. To showcase that the other conversing partner is an AI, a well-known icon representing an AI teammate was used, and the name "AI-Assist" was used and displayed beside the icon. In contrast, a human-like icon was used to represent the human collaborator.

In order to make sure that all participants watched the one-minute and two-minute videos, two approaches were used. First, prior to watching any of the videos, a message was displayed to all participants asking them to click on the "play" icon to listen to the video. The message also reminded them that they would not be able to proceed with the survey until they watched the video in full. Second, the experiment was designed so that participants could not proceed to the next page or see the "Next" button until the duration of the video being played had passed. After completing the experimental task, participants were directed to answer a set of survey questions that also included open-ended questions to justify and fortify the results of the quantitative study.

6.3. Sample Size

Participants were recruited and the survey study was carried out using Qualtrics. Qualtrics is a software company established in 2002 that offers user-friendly web-based tools and services to facilitate the development of surveys and the collection of survey data. Survey-based research is powerful in making inferences about the general population by collecting data from representative samples. The advantages of using web-based or online surveys are well documented in the literature as they (1) allow easy access to populations that would otherwise be difficult to reach, (2) make it easy to obtain a large sample, (3) are more convenient as data collection takes a shorter time, and (4) provide services at low administrative costs (Wright 2005). In this study, all participants were invited and compensated through Qualtrics.

For a statistical power of 0.8 with a medium effect size (f = 0.25), the G*Power 3.1.9.7 sample calculator suggested a sample size of 269 for five predictors (Faul et al. 2007, 2009). To allow for some spoiled surveys, a sample size of 300 was recommended. However, given that this study has four treatment groups and it is important to balance the number of participants in each treatment group as well as balancing the key demographics of age and gender, a total number of 380 participants were recruited in this study leading to a final number of 357 valid surveys, as will be discussed in section 6.6.

Table 6: Summary of the Four Experimental Conditions

Condition	Scenario
Condition 1 (LA / HE)	**AI**: Analyzes resumes, application forms, and videos and produces consolidated scores (ranks) about the shortlisted candidates. + **Explanations** are given regarding why these candidates were selected. **Human**: reviews and decides on the shortlist and gives permission to send invitations for personal interviews. Then, humans interview the candidates.
Condition 2 (HA / HE)	**AI**: Analyzes resumes, application forms, and videos and automatically sends invitations for personal interviews. + **Explanations** are given regarding why these candidates were selected. **Human:** is just informed about the invited candidates. Then, humans will interview the candidates.
Condition 3 (LA / LE)	**AI**: Analyzes resumes, application forms, and videos and produces consolidated scores (ranks) about the shortlisted candidates. **No Explanations** are provided. **Human**: reviews and decides on the shortlist and gives permission to send invitations for personal interviews. Then, humans interview the candidates.

Condition 4 (HA / LE)	**AI**: Analyzes resumes, application forms, and videos and sends invitations for personal interviews. **No Explanations** are provided. **Human**: is just informed about the invited candidates. Then, humans will interview the candidates.

Note: ***LA*** *= Low Autonomy,* ***HA****= High Autonomy,* ***LE****= Low Explainability,* ***HE****= High Explainability*

6.4. Pilot Study

Pilot studies are recommended to be employed before the main study, usually using a small, convenient sample to check the instrumentation before the main research design details are finalized (Boudreau et al. 2001). A pilot study took place before the main data was collected to refine the measurement instruments and the survey design, if necessary, and to detect any other possible issues with the research design. Therefore, 25 pilot surveys were collected through Qualtrics research firm, where a link to the study was sent to participants. All recruitment conventions were followed, and participants were required to electronically approve a consent form prior to sending them to the actual survey.

6.5. Main Study

For the main study, data was also collected through Qualtrics research firm. The use of a research firm received full ethics approval prior to the recruitment of any participants, including for the pilot study. A total of 380 participants were recruited, and all recruitment protocols were followed. Participants were required to electronically provide their consent prior to completing the survey. Since it was not possible to balance the representation of all the demographic variables in each treatment group, an effort was made to ensure that the participants' pool is representative in terms of the key demographics of age and gender.

6.6. Data Cleansing

Several procedures were followed to screen the data collected in this study and to ensure that all the responses collected were valid. The first step in cleaning survey-type data involved examining any missing values. When surveys are used, some participants may inadvertently or intentionally not answer one or more questions, resulting in incomplete data with missing values.

Second, other factors were considered to account for missing values and get rid of undesirable responses, such as the time participants spent completing the survey. From the pilot study, the average estimated time for participants to complete the survey was 10-15 minutes. The Qualtrics platform also has a tool that can predict the time participants can take to answer the survey questions based on the survey design and the number of questions included. Accordingly, any participant who finished the survey in less than 10 minutes was excluded from the pool. Another factor was to examine for straight lining, which is when a respondent selects the same response for most questions (Hair, Risher, et al. 2019). For example, in a 7-point Likert scale survey, the straight lining would be if a respondent selected 1 or 7 for most of the survey questions. This is of great importance, especially when incentives are offered to participants to complete the survey (e.g., providing participants with a financial reward for taking the survey).

In this study, missing values of a certain scale item were replaced with the mean values of that item since the number of missing values is less than 5% for that particular item, as suggested by (Hair, Risher, et al. 2019). In addition, to check for valid responses, one or more attention-check questions were inserted in the middle of the survey to ensure that participants were paying attention and taking the time to comprehend the question to decide on the right answer from their own perspective. Two attention questions were inserted into the survey to ensure participants' attentiveness while answering the questions[4]. Additionally, if a participant did not fully complete the survey or answered any of the attention check questions incorrectly, this participant's entire response was excluded. All of the aforementioned measures were considered in this study before deciding whether or not to remove a participant from our sample.

Next, an outlier analysis was performed. Outliers are instances or cases of a dataset with extreme or out-of-range values on a single variable (univariate) or a number of variables (multivariate) (Meyers et al. 2016). To detect univariate outliers, this study follows the boxplot method suggested by Cohen (2008) methods as described in (Meyers et al. 2016). Boxplots are based on median rather than mean values, where the upper and lower 'fences' of the boxplot are set at 1.5 times the Inter-Quartile Range (IQR). An outlier would be a value outside the upper or lower boundary of a boxplot (Meyers et al. 2016). The boxplot univariate outlier analysis was completed for each individual item of a construct to have a detailed inspection of any case causing a problem in the dataset. In Addition, the skewness and kurtosis of all the data points were checked to make sure all fall within the -2 and +2 ranges. Results revealed no serious departures of the skewness and kurtosis for all the individual items except for the first efficiency indicator (i.e., kurtosis = 2.42). Overall, a total of seven univariate outliers were detected, as shown in Figure 8 and Figure 9. After removing these outliers, the kurtosis fell within range (i.e., kurtosis= 1.82).

[4] In completing this task...

- I am not paying any attention to this survey

- I am answering questions in this survey without thinking

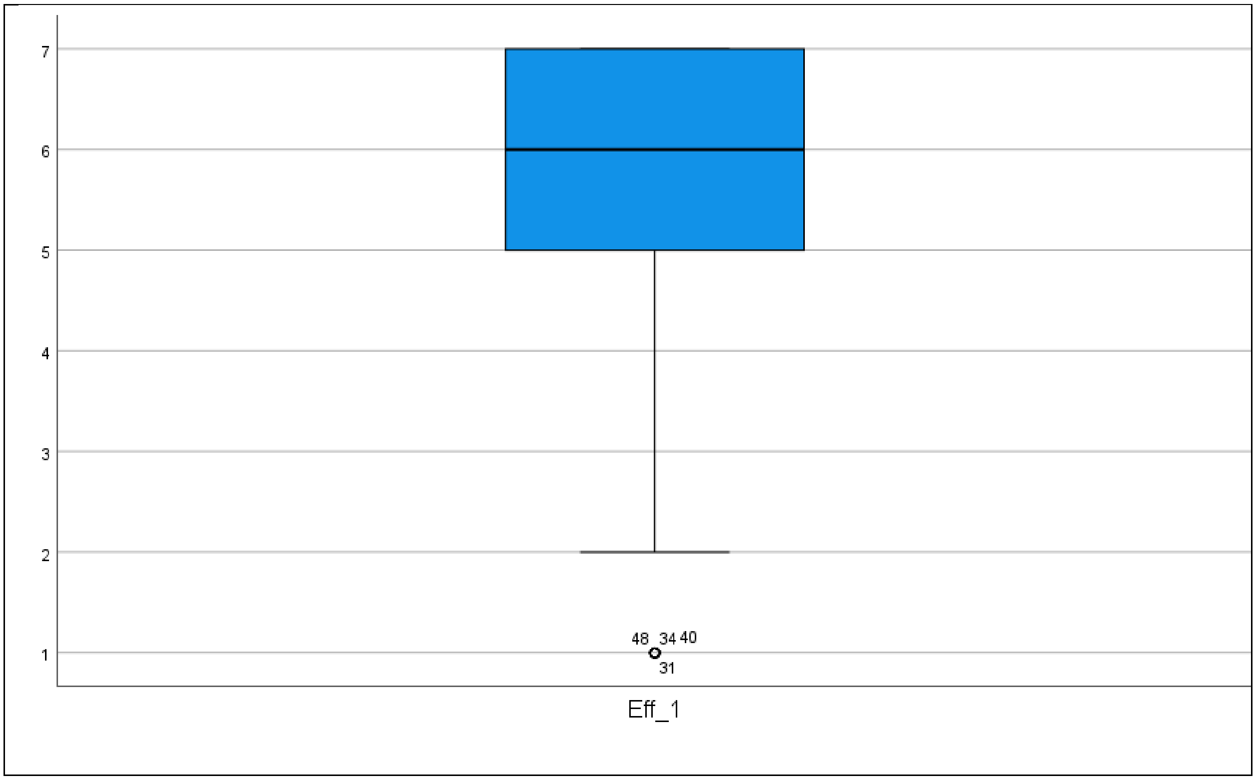


Figure 8: Univariate outliers detected from round one for the first efficiency indicator

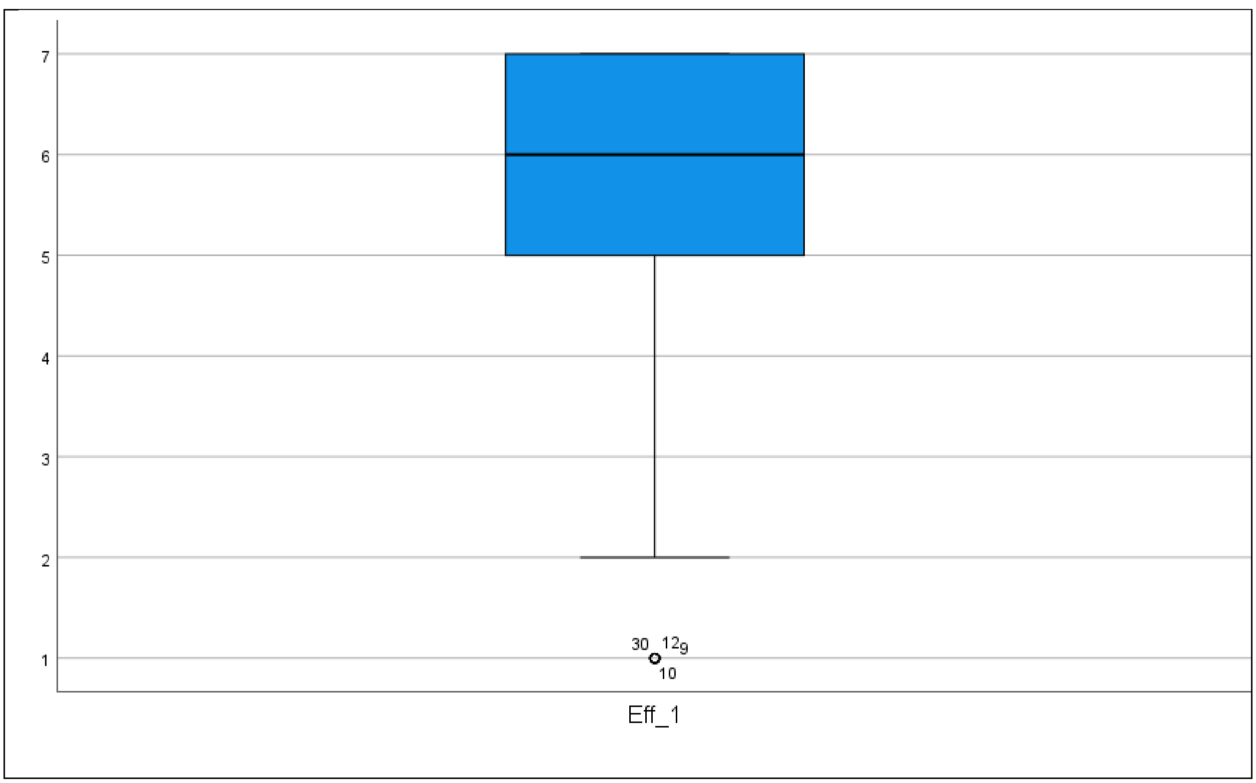


Figure 9: Univariate outliers detected from round two of univariate outlier analysis for the first efficiency indicator

Next, a Mahalanobis distance was performed using linear regression in SPSS Statistics for multivariate outlier analysis. Respondent identification (ID) numbers were used as the dependent variable, and efficiency, integrity, lack of human interaction, incompatibility, and willingness to collaborate were the independent variables. This analysis indicated that five cases are considered multivariate outliers. Such outliers were further examined, and it was found that removing them positively impacted the results. Therefore, they were removed from the dataset. From the 390 recruited subjects, a total of 357 were identified as the final valid responses after removing problematic cases (e.g., straight-lining, univariate, and multivariate outliers). This number also satisfies the PLS requirement since PLS requires that the sample size be ten times the number of items in the most complex (Gefen et al. 2000). From the above sample size requirements calculations, a sample of 357 participants was sufficient to run the analysis.

6.7. Demographics

As mentioned earlier, hiring managers or hiring employees who are responsible for hiring staff at their organizations were recruited to take the survey. Along with the main model constructs, a number of general demographic information was collected from respondents (see Appendix E) for a list of these questions. General demographic variables captured were age, gender, education level, years of hiring experience, and knowledge level of AI. A full breakdown of participants' demographics is provided in Table 7.

Table 7: Survey Participants' Demographics

Variable		Count	%
Gender	Man	208	58%
	Woman	148	41%
	Non-gender-binary, two-spirit,or similar	1	0%
	Others	0	0%
	Total	**357**	
Age	20-30	30	8%
	31-40	78	22%
	41-50	98	27%
	51-60	79	22%
	61-70	54	15%
	> 70	18	5%
	Total	**357**	
Education Level	High school diploma	33	9%
	Some college degree	88	25%
	Bachelors	140	39%
	Master's	76	21%
	Ph.D	11	3%
	Other	9	3%
	Total	357	
Years of Hiring Experience	1–5 years	79	22%
	6–10 years	93	26%
	11-20 years	99	28%
	21-30 years	50	14%
	>=31	36	10%
	Total	**357**	
Knowledge Level of AI	No knowledge	11	3%
	Very basic knowledge	91	25%
	Medium knowledge level	115	32%
	Very good knowledge level	96	27%
	Excellent	44	12%
	Total	**357**	

6.8. Measurement Scales

Participants recruited to take the survey were asked to answer a set of questions. Measurement scales used in the survey were borrowed from well-established, validated scales in the extant literature to ensure content validity. The scales were adapted to reflect the context of this study. A seven-point Likert scale, with "strongly disagree" and "strongly agree" as the endpoints, was followed. For the benefits, efficiency and integrity were measured using the scale adapted from (Wilkinson et al. 2021). For the concerns, the lack of human interaction scale was adapted from (Dabholkar and Bagozzi 2002), while incompatibility was adapted from (Hari et al. 2021). The explainability scale was adapted from (Jabagi et al. 2021), and autonomy was adapted from (Ahuja and Thatcher 2005). The willingness to Collaborate with AI was measured using a 3-item scale adapted from (Gursoy et al. 2019). See Appendix D for details about the scale items used in this study.

In addition to the aforementioned scales, three open-ended questions were also added at the end of the survey to solicit a richer understanding of each participant's point of view regarding the role of AI explainability and its level of control as opposed to just obtaining values from a scale. Responses gathered from the open-ended questions added richness to the findings and were used to confirm the quantitative analysis results.

6.9. Structural Equation Modeling (SEM)

Structural Equation Modeling (SEM) is used to validate the proposed research model. SEM assesses both the measurement model (i.e., uses factor analysis to determine the degree that the observed variables load on their latent constructs) and the structural model (i.e., estimates the assumed causal and covariance linear relationships among the exogenous and endogenous latent variables) in one analysis (Boudreau et al. 2001). SEM allows complicated variable relationships to be modelled through hierarchical or non-hierarchical structural relationships and is used to test the extent to which the statistical analysis of the IS research being studied is of high quality (Gefen et al. 2000). That is to test for statistical conclusion validity. SEM helps researchers answer interrelated research questions in one systematic and comprehensive analysis.

That said, Partial Least Squares (PLS) was used in this study as a SEM method to analyze the data and validate the proposed model. PLS is chosen for this study because a) it is more suitable for studies with an exploratory type which is the case in this study (Gefen et al. 2000), as the case in this study, b) it estimates the variance of dependent constructs and their associated latent variables (Chin et al. 2003), and c) it is best suited when: (i) hypotheses are derived from theory, and the relevant variables are not known, (ii) the relationships between theoretical variables and their indicators are unclear, and (iii) the relationships between latent variables are hypothetical (Falk and Miller 1992). To analyze the data and validate the model in this study, SmartPLS 4 was used, where the measurement model was assessed, and then the structural model. Specifically, consistent PLS (PLSc-SEM) is utilized, as it employs a correction of reflective constructs' correlations in order to make results consistent with a factor model (Dijkstra and Henseler 2015). The following sections present the results of both stages in detail.

6.9.1. Measurement Model

When evaluating the measurement model, the focus is placed on the reliability and validity of the measures used to represent the model’s constructs (Chin 2010).

6.9.1.1. Reliabaility Analyses

Since no observed data was used in this research, a reliability test of the scales used to measure study variables was conducted. Reliability reflects the measurement scale accuracy in representing a construct. It refers to "*the extent to which the respondent can answer the same questions or close approximations the same way each time*" (Straub et al. 2004, p. 400). Since all the study variables are composite variables consisting of multi-item scales, SPSS 26 was used to calculate Cronbach's alpha of each latent variable to check for its internal consistency. We did this test for each composite variable, and each of these measures exhibited high reliability (i.e. Cronbach alpha is greater than .70) as suggested by (Nunnally 1978) (see Table 8).

Composite reliability (CR) for each latent variable was also calculated using SmartPLS4 (see Table 8). All constructs exhibited good reliability (i.e., CR ≥ 0.6) (Bagozzi and Yi 1988). Thus, we can conclude that all the measurement scales used in this study meet the construct reliability requirements.

Table 8: Cronbach's Alpha, Composite Reliability and AVEs of the Study Variables

Variable	Cronpach's Alpha (α)	Composite Reliability	Average Variance Extracted (AVE)
Efficiency	**0.91**	**0.91**	**0.77**
Integrity	**0.89**	**0.81**	**0.72**
Lack of Human	**0.83**	**0.84**	**0.72**
Incompatibility	**0.81**	**0.81**	**0.60**
Autonomy	**0.74**	**0.86**	**0.77**
Explainability	**0.82**	**0.83**	**0.70**
Willingness to Collaborate	**0.94**	**0.94**	**0.85**

6.9.1.2. Convergent and Discriminant Validity Analyses

Construct validity refers to "*an issue of operationalization or measurement between constructs…. It raises the basic question of whether the measures chosen by the researcher 'fit' together in such as way so as to capture the essence of the construct*" (Straub et al. 2004, p.388). Since the constructs in our model are reflective, convergent validity and discriminant validity were examined. For reflective constructs, all measurement items are assumed to be unidirectional and highly correlated with one another to reflect the latent variable. To assess convergent validity, Average Variance Extracted (AVE) values are assessed, where AVE values should exceed 0.50 (Fornell and Larcker 1981). As Table 11 shows, all AVE values on the diagonal are >= 0.5, thus providing evidence of satisfactory convergent validity.

Discriminant validity means that "*the existence of a construct is that the measurement items posited to make up that construct differ from those that are not believed to make up the construct*" (Straub et al. 2004, p.389). To assess discriminant validity, two methods were followed: examining item loadings to construct correlations and examining the ratio of the square root of the AVE of each construct to the correlations of this construct to all the other constructs (Gefen and Straub 2005). For the first technique, scale items should be investigated to ensure they load more highly on their theoretically assigned latent variable than on any other latent construct (Gefen & Straub, 2005). Such that if each item loading is greater for its assigned variable by at least 0.1 and each of the variables loads highest with its corresponding items, it can be concluded that there is an adequate level of construct validity (Straub et al. 2004; Urbach and Ahlemann 2010).

To do this, SmartPLS4 was used to examine indicator/item loadings. As shown in Table 9, all indicators do load most highly on their own theoretically assigned construct at a minimum threshold of 0.50 as per (Hair, Babin, et al. 2019), except for the "Lack of Human Interaction" construct, where the first scale item loading was less than the minimum threshold (i.e., item loading = 0.268) and the difference between the item loadings and this item was greater than 0.1 (see Table 9). Therefore, this analysis took an iterative approach in examining the cross-loadings after removing the problematic item of the "Lack of Human Interaction" scale and rerunning the analysis. In the second iteration, all item loadings satisfied the construct validity requirements as shown in Table 10.

Table 9: First Iteration of Item Loadings and Cross Loadings of Measurement Scales

Construct Items	Efficiency	Integrity	Lackof H	Incomp	Autonomy	Explain	Willingness
Eff_1	**0.893**	0.601	-0.38	-0.367	0.021	0.464	0.658
Eff_2	**0.867**	0.635	-0.4	-0.37	0.04	0.439	0.643
Eff_3	**0.895**	0.661	-0.415	-0.415	-0.003	0.419	0.678
Integrity_1	0.604	**0.843**	-0.441	-0.326	0.014	0.46	0.599
Integrity_2	0.627	**0.836**	-0.417	-0.349	0.082	0.418	0.612
Integrity_3	0.608	**0.874**	-0.438	-0.359	0.042	0.379	0.63
LackofH_1	-0.12	-0.209	**0.268**	0.331	0.062	-0.288	-0.166
LackofH_2	-0.419	-0.461	**0.852**	0.531	0.072	-0.279	-0.514
LackofH_3	-0.422	-0.459	**0.976**	0.595	0.11	-0.288	-0.607
Incomp_1	-0.262	-0.239	0.483	**0.671**	0.125	-0.272	-0.335
Incomp_2	-0.35	-0.313	0.531	**0.797**	0.084	-0.305	-0.415
Incomp_3	-0.387	-0.372	0.459	**0.849**	0.027	-0.304	-0.454
Willing_1	0.689	0.672	-0.596	-0.486	-0.01	0.397	**0.944**
Willing_2	0.702	0.686	-0.555	-0.47	-0.032	0.435	**0.94**
Willing_3	0.68	0.63	-0.527	-0.495	0.016	0.462	**0.899**
Exp_1	0.417	0.413	-0.269	-0.338	0.018	**0.838**	0.394
Exp_2	0.421	0.416	-0.29	-0.304	-0.086	**0.836**	0.39
Aut_1	0.116	0.139	0.034	-0.014	**0.527**	0.047	0.056
Aut_2	-0.025	0.009	0.133	0.14	**0.998**	-0.077	-0.04

Table 10: Second Iteration of Item Loadings and Cross Loadings of Measurement Scales

	Efficiency	Integrity	LackofH	Incomp	Autonomy	Explain	Willingness
Eff_1	**0.863**	0.601	-0.42	-0.367	0.021	0.464	0.658
Eff_2	**0.867**	0.635	-0.437	-0.37	0.04	0.439	0.643
Eff_3	**0.909**	0.661	-0.451	-0.415	-0.003	0.419	0.678
Integrity_1	0.604	**0.845**	-0.469	-0.326	0.014	0.46	0.599
Integrity_2	0.627	**0.837**	-0.446	-0.349	0.082	0.418	0.612
Integrity_3	0.608	**0.860**	-0.471	-0.359	0.042	0.379	0.63

LackofH_2	-0.419	-0.461	**0.789**	0.531	0.062	-0.288	-0.514
LackofH_3	-0.422	-0.459	**0.903**	0.595	0.072	-0.279	-0.607
Incomp_1	-0.262	-0.239	0.508	**0.671**	0.11	-0.288	-0.335
Incomp_2	-0.35	-0.313	0.559	**0.797**	0.125	-0.272	-0.415
Incomp_3	-0.387	-0.372	0.481	**0.849**	0.084	-0.305	-0.454
Willing_1	0.689	0.672	-0.647	-0.486	0.027	-0.304	**0.935**
Willing_2	0.702	0.686	-0.607	-0.47	-0.01	0.397	**0.933**
Willing_3	0.68	0.63	-0.581	-0.495	-0.032	0.435	**0.891**
Exp_1	0.417	0.413	-0.269	-0.338	0.018	**0.838**	0.394
Exp_2	0.421	0.416	-0.29	-0.304	-0.086	**0.836**	0.39
Aut_1	0.116	0.139	0.034	-0.014	**0.527**	0.047	0.056
Aut_2	-0.025	0.009	0.133	0.14	**0.998**	-0.077	-0.04

Next, the square root of the AVE values for the latent variables used in this study was examined. Fornell and Larcker (1981) recommend that the square root of the AVE values for the latent variables should be larger than the correlations the construct has with any other construct. As demonstrated in Table 11, all square roots of the AVEs for each construct are greater than the correlation with any other construct, thus providing evidence of discriminant validity.

6.9.1.3. Multicollinearity Analysis

Multicollinearity occurs when variables are too highly correlated. As multicollinearity increases, it makes it difficult to correctly interpret the variable because it will be hard to ascertain the effect of any single variable owing to their interrelationships (Hair 1995). To assess multicollinearity in this study, different approaches were followed. First, inter-construct correlations were examined. Such that if bivariate correlations are greater than 0.80, it can be inferred that there could be multicollinearity issues that require either removing a variable (Meyers et al. 2016) or combining variables into a larger one (Stevens 2002). As presented in Table 11 shows, none of the inter-construct correlations indicate issues with multicollinearity.

Second, a more advanced multicollinearity analysis was completed by checking the Tolerance and Variance Inflation Factor (VIF) indices using SPSS 26. As suggested by Meyers et al. (2016), multicollinearity exists when predictor variables are strongly correlated but should not exist to correlations between predictor variables and dependent variables. Therefore, the Tolerance and VIF analysis were examined between predictor variables. Tolerance values should be less than 0.01, and the VIFs should be below ten as per (Esposito Vinzi et al. 2010; Stevens 2002). Using SPSS 26, all of the Tolerance values and VIFs for this study met the required threshold (i.e., Tolerance values < 0.01 and VIFs , 10). Consequently, multicollinearity is not deemed to be an issue in this research study.

Table 11: Descriptive statistics, correlation matrix, and Square Roots of AVEs

Variable	Mean	SD	Eff.	Integ.	Lack_H	Incomp	Willing	Autono	Expl	Blue_ Att	Pub_ Tr
Efficiency (1)	**5.68**	**1.12**	**0.89**								
Integrity (2)	**5.37**	**1.20**	.609**	**0.86**							
Lack of Human Interaction (3)	**4.13**	**1.56**	-.515**	-.525**	**0.85**						
Incompatibility (4)	**3.83**	**1.42**	-.46**	-.399**	.553**	**0.78**					
Willingness to Collaborate (5)	**5.50**	**1.25**	.742	.667**	-.625**	-.500**	**0.93**				
Autonomy (6)	**4.85**	**1.45**	.042	.072	.090	.074	.005	**0.77**			
Explainability (7)	**5.45**	**1.35**	.432**	.422**	-.279**	-.313**	.414**	.439**	**0.85**		
Blue Attitude (8)	**4.94**	**1.29**	.089	.101	-.022	-.057	.094	.028	.015	**0.73**	
Public_Transit (9)	**2.89**	**1.51**	.073	.090	-.055	-.040	.116*	.078	.012	.038	**0.70**
Notes: ** Correlation is significant at p<.01, * Correlation is significant at p<.05											

6.9.1.4. Common Method Bias

Since the study employs cross-sectional self-reported surveys, common method bias could be an issue. Common Method Bias (CMB) or Common Method Variance (CMV) is "*the systematic* *error variance shared among variables measured with and introduced as a function of the same method and/or source*" (Richardson et al. 2009, p.763). Common method bias can lead researchers to observe a significant effect when, in fact, the true effect is due to the method used (Whitman and Woszczynski 2003). This systematic error variance can bias the findings of empirical analyses and can also bias the estimated relationships among variables or measures (Jakobsen and Jensen 2015). Podsakoff et al. (2003) uncovered four main sources of CMB: a) having a common rater or the same person responding to the measures of the dependent and independent variables, b) having measurement items with some specific characteristics or properties (i.e., wording items with hidden cues as to how to respond, including complex and ambiguous items), c) the positioning of the survey items (i.e., items context) might influence respondents (e.g., the context and positioning of the items on a questionnaire), and d) the contextual influences in which the measures are obtained (i.e., location, time, media) (Podsakoff et al. 2003; Tehseen et al. 2017).

Many researchers have suggested the use of multiple methods or remedies to control for the effect of common method bias on the study results (Lindell and Whitney 2001; Podsakoff et al. 2003). Some remedies could be applied during the data collection phase, while others could be applied after the data collection phase (Tehseen et al. 2017). For the former, it was recommended that researchers try not to measure the dependent and independent variables from the same informants (Podsakoff et al. 2003; Williams et al. 2010). This means that researchers should collect survey data about the dependent variables from participants who are different from those informants invited to answer survey questions related to the independent variables since the use of the same respondent is one of the most common sources of common method bias. Although this may not be convenient and possible for all researchers, another method proposed by (Chin et al. 2013) is called "Measured Latent Marker Variable" (MLMV). This approach recommends collecting multiple theoretically unrelated measures, called "marker variables," at the same time as collecting data related to the primary constructs of the research model. Chin et al. (2013) state, "*A critical aspect of the MLMV approach is to select a set of measures that reflect underlying constructs that have no nomological relationship with the particular study in question while using the same survey format and scale to reflect the common method effects."* (p. 232). Williams et al. (2010) advise that adding a latent marker variable and its indicators to a research study will always yield more robust analyses and is likely to be feasible in most circumstances.

There are two important factors to consider when choosing a marker variable for the marker variable analysis proposed by (Simmering et al. 2015). First, the marker variable has to be theoretically unrelated to the variables in the model so that they do not share any meaningful variance with each other. Lindell and Whitney (2001) recommend that cross-sectional surveys should include at least one scale that is expected, based on prior theory, to be unrelated to the model variables. Likewise, Williams et al. (2010) suggested paying careful consideration to the role of theory to ensure that the marker variable is theoretically unrelated to the substantive variables of the study. The authors also recommended reviewing the flow and stages of the questionnaire that informants will go through to answer a measurement item and assess how the marker variable would fit with such flow and context. In addition, the marker variable should be susceptible to the same sources of biases as the items in the study and should use the same rating approach (Williams et al. 2010). For example, if the study variables are factual (e.g., age, demographics), the marker variable should not be behavioural (e.g., social desirability) or evaluative in nature (e.g., job performance) (Spector et al. 2019). Based on these criteria, two theoretically unrelated variables were used as marker variables in this study (i.e., blue attitude and public transit attitude) (see Appendix D). The questions pertaining to the two marker variables were included at the end of the survey to minimize the effects of respondents' fatigue on the pattern of responses relevant to the main study, as suggested by (Chin et al. 2013).

To evaluate the effectiveness of the marker variables, first, the correlation values between these two marker variables and the model constructs were tested as shown in Table 11. From the table, there is no significant correlation between the marker variables and the model constructs. Second, we followed the Construct Level Correction (CLC) approach introduced by (Chin et al. 2013) as in (Tehseen et al. 2017). In this method, the marker variables were modelled to have a path with each PLS model's construct. Then, path coefficients are estimated before and after introducing the marker variables on the model constructs. As Table 12 shows, it is observed that the changes between the original estimated path coefficient of the main model constructs before and after CLC are very small and not significant. Likewise, there were non-significant changes between the R^2 value in the original PLS model and the R^2 estimated by the CLC approach, as shown in Table 13. Therefore, we can conclude that the possibility of common method bias is very low in this study.

Table 12: Comparison of Path Coefficients by CLC Approach and Original PLS Models

Relationships	CLC Estimation (Path Coefficients)	Original PLS Estimates (Path Coefficient)
Efficiency → Willingness	0.463	0.464
Integrity → Willingness	0.236	0.239
Lack of Human Interaction → Willingness	(0.264)	(0.261)
Incompatibility → Willingness	(0.037)	(0.038)

Table 13: Comparison of R^2 Values by CLC Approach and Original PLS Models

Endogenous Construct	CLC Estimation (R^2)	Original PLS Estimate (R^2)
Willingness to collaborate	**0.751**	**0.751**

Another common and traditional approach could be followed after collecting the survey data to detect the common method bias problem after the data is collected. This method is called Harman's One Factor test. To assess for common method bias using this test, all scale items of the constructs in our research model were included in an exploratory principal components analysis (PCA).

Common method variance is present if either (a) a single factor will evolve from the factor analysis or (b) one general latent variable will explain more than 50% of the covariance among the measures" (Podsakoff et al. 2003). SPSS 26 was used to conduct a Principal Component Analysis, and all of the scale items of interest were used in the analysis. As shown in Table 14, a multi-factor solution emerged, with the first-factor accounting for only 42.496% of the variance. This provides sufficient evidence that the variables in the model do not load onto one factor and that the possibility of common method variance is low.

Table 14: Principal Component Analysis for Harman's Single Factor Test

Total Variance Explained						
Factor	**Initial Eigenvalues**			**Extraction Sums of Squared Loadings**		
	Total	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	8.116	45.091	45.091	7.649	**42.496**	42.496
2	1.836	10.198	55.290			
3	1.393	7.739	63.029			
4	1.264	7.024	70.053			
5	1.016	5.645	75.698			
.	.	.	.			

.	.	.	.			
19	0.097	0.537	100.000			

6.9.2. Structural Model

After the assessment of the measurement model, the structural relationships of the research model are examined. As mentioned earlier, to test the structural model and assess the hypotheses developed, PLSc-SEM is used in this study. The results of the structural model are shown below and detailed in the following section.

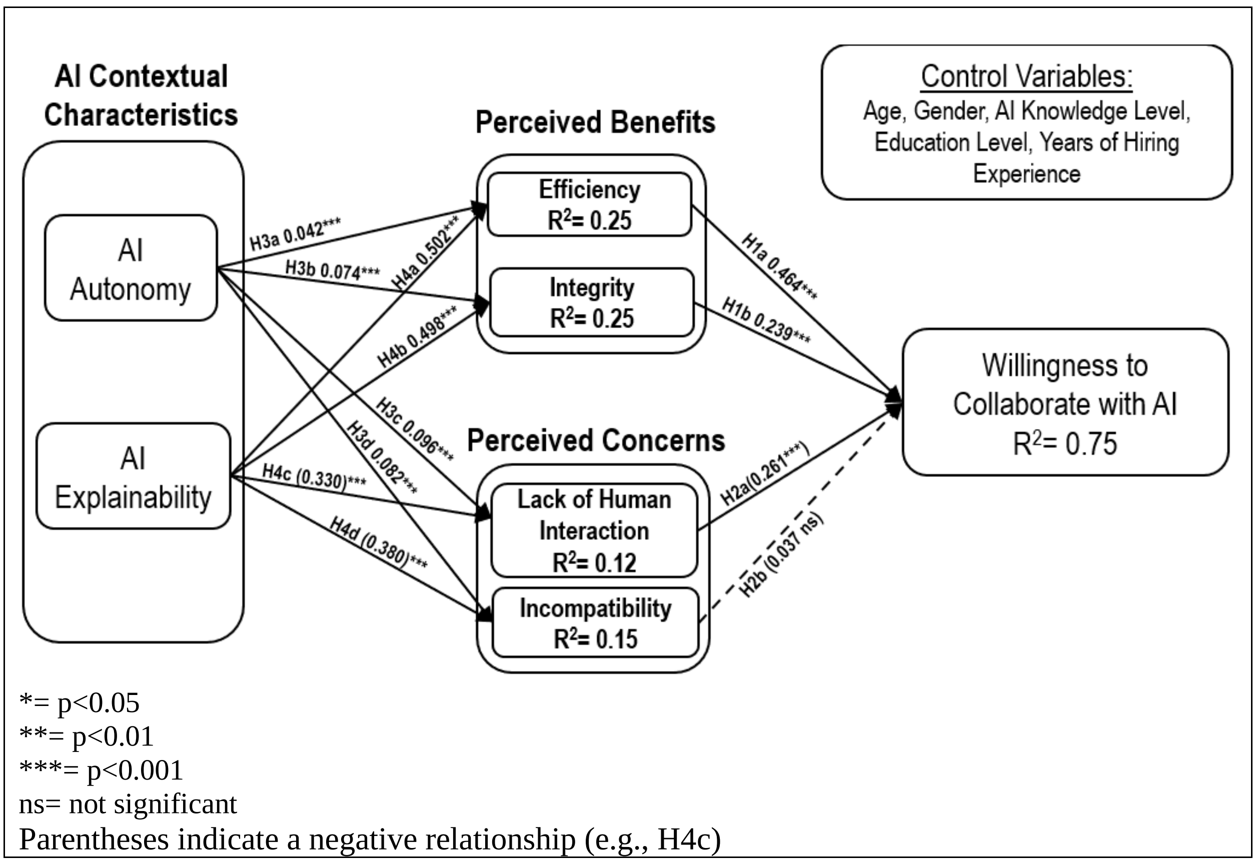


Figure 10: Final PLSc Model Results

6.9.2.1. Manipulation Validity

Manipulation checks are assessments of the degree to which the intended manipulation or treatment was correctly perceived by participants. This is a critical test performed to ensure the internal validity of experiments (Boudreau et al. 2001). Thus, this study manipulated two main variables in the experimental design: AI autonomy and AI explainability. Participants were divided into four groups (see Table 5), in which Groups 1 and 2 (n=178) were assigned to the high explainability condition, while Groups 3 and 4 (n=179) were assigned to the low explainability condition. Likewise, Groups 1 and 3 (n=181) were assigned to the high autonomy condition, while Groups 2 and 4 (n=176) were assigned to the low autonomy condition. To check whether the different conditions presented to participants were perceived as intended, the autonomy scale adapted from (Ahuja and Thatcher 2005), and the explainability scale adapted from (Jabagi et al. 2021) were used to check the validity of the manipulation of experimental treatments.

To assess the manipulation of autonomy and explainability, a one-way analysis of variance (ANOVA) was conducted to ascertain if there were significant differences between the responses to the manipulation check questions between the treatment groups. The results of the analysis are summarized in Table 15 and Table 16. It is shown that there are significant differences between treatment groups that were exposed to the high explainability condition versus low explainability (M=5.98 versus 4.91, $\rho<.001$). Similarly, there is a significant difference in the mean responses to the manipulation check questions that were examined for the high and low autonomy conditions (M=5.16 versus 4.54, $\rho<.001$). Appendix G also shows the results of a post hoc Tukey's HSD (Honestly Significant Differences) test, which confirmed significant differences between groups. Therefore, the manipulation check indicates the treatment was effective.

Table 15: One-Way ANOVA Analysis for Explainability Manipulation Check

Treatment Groups	**N**	**Mean**	**SD**	**95% Confidence Interval for Mean**	**ANOVA (Between Groups)**

				Lower Bound	Upper Bound	Sum of Squares	Mean Square	F	Sig.
High	178	5.98	.88	5.85	6.11	102.132	102.132	66.42	.000
Low	179	4.91	1.51	4.69	5.14				

Table 16: One-Way ANOVA Analysis for Autonomy Manipulation Check

Treatment Groups	N	Mean	SD	95% Confidence Interval for Mean		ANOVA (Between Groups)			
				Lower Bound	Upper Bound	Sum of Squares	Mean Square	F	Sig.
High	181	5.16	1.34	4.96	5.35	33.41	33.41	16.69	.000
Low	176	4.54	1.49	4.33	4.77				

6.9.2.2. PLSc Model Results

After assessing the validity of the measurement model and ensuring that participants did perceive the manipulation, the next step is to provide evidence for the proposed theoretical hypotheses by examining the structural model. Thus, Structural Equation Modeling using SmartPLS 4 was used to assess the proposed hypotheses and the significance of the path coefficients.

Results of the structural model are shown in Figure 10 and summarized in Table 17, where the original hypotheses are highlighted alongside the beta coefficients, t-statistic, the relationship significance, and its support. The beta coefficient provides information about the strength and direction of the relationship between variables in the model. They are typically reported alongside their associated t-values, which indicate the statistical significance of the relationship. The beta coefficient is standardized to allow for comparisons of the magnitudes of effects across different variables and scales within the SEM. Interpreting the beta coefficients involves considering their signs (positive or negative) and magnitudes. A positive beta coefficient indicates a positive relationship between the variables, meaning that an increase in the independent variable is associated with an increase in the dependent variable. Conversely, a negative beta coefficient indicates an inverse relationship.

As shown in Table 17, eleven out of the twelve hypotheses were supported. For all hypotheses, the algebraic sign (i.e., either positive or negative) of the path coefficient matched the hypothesized algebraic sign. As hypothesized, efficiency (β=0.46; ρ<0.001) and integrity (β=0.24; ρ<0.001) have a positive impact on the willingness to collaborate with AI, supporting H1a and H1b. Lack of Human Interaction (β= -0.26; ρ<0.001) negatively influences the willingness to collaborate with AI, supporting H2a. However, incompatibility (β= -0.37; ρ= 0.26) did not have a significant influence on the willingness to collaborate with AI. This indicates that all the benefits do have an influence on people's willingness to collaborate with AI, while only the Lack of Human Interaction concern impacts their willingness to collaborate but not the incompatibility.

Examining the effect of autonomy on the benefits and concerns, we can see that autonomy has a significant positive influence on the concerns and the benefits as hypothesized. Autonomy positively impacted efficiency (β= 0.042; ρ<.001) and integrity (β= 0.074; ρ<.001), as well as

Lack of Human Interaction (β= 0.096; ρ<0.01) and incompatibility (β= 0.082; ρ<0.001); supporting H3a, H3b, H3c and H3d. Likewise, explainability was found to influence the benefits and the concerns as hypothesized. It has a significant positive influence on the benefits, where the influence on efficiency (β=0.502; ρ<0.001) and integrity (β=0.074; ρ<0.001). Whereas the effect on the concerns was found to be negative on the Lack of Human Interaction (β= -0.330; ρ<0.001) or on incompatibility (β= -0.380; ρ<0.001), supporting H4a, H4b, H4c, and H4d.

Table 17: Validation of the Study Hypotheses

Hypothesis	Path Coefficient (β)	p-value	Supported?
H1a (+): Efficiency → Willingness	0.464***	0.000	Yes
H1b (+): Integrity → Willingness	0.239***	0.000	Yes
H2a (-): Lack_of_Human → Willingness	-0.261***	0.000	Yes
H2b (-): Incompatibility → Willingness	-0.038	0.518	No
H3a (+): Autonomy → Efficiency	0.184***	0.000	Yes
H3b (+): Autonomy → Integrity	0.074***	0.000	Yes
H3c (+): Autonomy → Lack_of_Human	0.096***	0.000	Yes
H3d (+): Autonomy → Incompatibility	0.082***	0.000	Yes
H4a (+): Explainability → Efficiency	0.502***	0.000	Yes
H4b (+): Explainability → Integrity	0.498***	0.000	Yes
H4c (-): Explainability → Lack_of_Human	-0.330***	0.000	Yes
H4d (-): Explainability → Incompatibility	-0.380***	0.000	Yes

*** Correlation is significant at the 0.001 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Effect Sizes

An effect size analysis was conducted to evaluate the impact of the independent antecedents on the dependent constructs (Cohen 2013). The analysis was performed as follows:

1. The R^2 value was calculated twice, where all the antecedents were included for all the dependent variables one time and then where one of the antecedents was excluded one at a time.
2. The change in the R^2 value was evaluated as follows: 0.02 corresponds to a small effect, 0.15 corresponds to a medium effect, and 0.35 corresponds to a large effect (Henseler and Sarstedt 2013; Roldán and Sánchez-Franco 2012).

Table 18: PLS Effect Size Analysis

Dependent Construct	Independents	R^2		ΔR^2	Effect Size
		Included	Excluded		
Willingness	**Efficiency**	0.751	0.650	0.101	Small
	Integrity		0.724	0.027	Small
	Need of Human		0.721	0.03	Small
	Incompatibility		0.750	0.001	ns
Efficiency	**Autonomy**	0.252	0.250	0.002	< Small
	Explainability		0.000	0.252	Medium
Integrity	**Autonomy**	0.253	0.25	0.003	< Small
	Explainability		0.003	0.25	Medium
Need of Human	**Autonomy**	0.12	0.109	0.011	< Small
	Explainability		0.011	0.109	Small
Incompatibility	**Autonomy**	0.154	0.145	0.009	< Small
	Explainability		0.009	0.145	Small

Note:
- *"ns" means that the path is not-significant*

From Table 18, results indicate that (efficiency → explainability) and (integrity → explainability) have a medium effect size. In comparison, the effect sizes of five significant paths in the original PLS model are considered to be small, except for (efficiency → autonomy), (integrity → autonomy), (lack_of_human → autonomy), and (incompatibility → autonomy), which have an effect size smaller than 0.02. These results are in line with prior literature that discusses that effect sizes in social science research are often small (Ferguson 2016; Rosnow and Rosenthal 2003). Therefore, the majority of effect sizes being small is not surprising.

Goodness of Fit Assessment

Unlike covariance-based structural equation modelling (CB-SEM), PLS path modelling does not involve multiple model fit indices, but a global goodness of fit index (i.e., GoF index) is used as suggested by (Tenenhaus et al. 2004). Such a global index is used to assess the performance of both the measurement model and the structural model, providing a single measure for the overall prediction performance of the research model (Esposito Vinzi et al. 2010). The global GoF index is calculated using the geometric mean of the average communality index and the average R^2 value model (Akter et al. 2011; Esposito Vinzi et al. 2010) as shown in the formula below:

$$GoF = \sqrt{(\text{Average Communality} * \text{Average } R^2)}$$

It produces a value between zero and one that is interpreted in the same way the value of the effect sizes is interpreted. According to (Wetzels et al. 2009), the thresholds for the effect sizes are considered small if between 0.10 and 0.25, medium if between 0.25 and 0.36, and large if greater than 0.36. Hence, based on the above formula and as shown in Table 19, the GoF value for the research model proposed in this study is 0.47, indicating a good model fit.

Table 19: GoF Index Calculation

Constructs	R-Square	AVE
Efficiency	0.252	0.77
Integrity	0.250	0.72
Incompatibility	0.154	0.60
Need of Human	0.120	0.72
Willingness	0.751	0.847
Average	0.305	0.7314
GoF	**0.47**	
SRMR	**0.040**	
NFI	**0.928**	

In addition to the global fit index, both the Standardized Root Mean Square Residual (SRMR) and the Normed Fit Index (NFI) or Bentler and Bonett Index provided by SmartPLS 4 were examined. SRMR is based on transforming both the sample covariance matrix and the predicted covariance matrix into correlation matrices. It is defined as the difference between the observed correlation and the model-implied correlation matrix. Thus, it allows assessing the average magnitude of the discrepancies between observed and expected correlations as an absolute measure of the (model) fit criterion. A value less than 0.10 is considered a good fit. (Hu and Bentler 1999). Henseler et al. (2014) argue that the SRMR as a goodness-of-fit measure for PLS-SEM that can be used to avoid model misspecification.

NFI, on the other hand, was one of the first fit measures proposed in the SEM literature is proposed by (Bentler and Bonett 1980). It computes the Chi-square value of the proposed model and compares it against a meaningful benchmark. NFI results in values between 0 and 1. The closer the NFI is to 1, the better the fit. NFI values above 0.9 usually represent an acceptable fit. Based on this, the SRMR fit index of the proposed research model is 0.040, and for NFI, the fit index is 0.928, indicating very good fit indices of the research model.

6.9.2.3. Post-hoc Analyses

For post-hoc analyses, the different combination effects of the autonomy and explainability interventions are examined, and a control variable analysis is performed.

Group Comparison of Autonomy and Explainability Interventions

Participants in this research received a different combination of autonomy and explainability treatments. Therefore an analysis of the potential treatment effects was completed. To do this, a one-way ANOVA analysis that examines the differences in means for the four groups was performed. Results of this analysis for autonomy are reported in Table 20, indicating that there are significant differences between the groups in the responses for the Lack of Human Interaction and incompatibility but not for efficiency and integrity. Results for explainability are reported in Table 21, indicating that there are significant differences between the groups in the responses for efficiency and integrity but not for the Lack of Human Interaction and incompatibility.

Table 20: One-Way ANOVA Group Comparisons for Autonomy

Variable	Sum of Square	df	Mean Square	F	Sig.
Efficiency	3.29	1	3.29	2.61	.107
Integrity	1.41	1	1.41	.978	.323
Lack of Human Interaction	15.20	1	15.20	6.30	**.012**
Incompatibility	7.62	1	7.62	3.80	**.052**

Table 21: One-Way ANOVA Group Comparisons for Explainability

Variable	Sum of Square	df	Mean Square	F	Sig.
Efficiency	6.91	1	6.91	5.53	**.019**
Integrity	6.11	1	6.11	4.26	**.040**
Lack of Human Interaction	6.68	1	6.68	2.75	.098
Incompatibility	5.55	1	5.55	2.76	.098

In order to assess which specific groups exhibited significant differences in the responses for efficiency, integrity, Lack of Human Interaction, and incompatibility, a Tukey's HSD was also completed as part of the one-way ANOVA.

The Tukey's HSD test is a post-hoc analysis that compares all possible pairs of means and identifies where there are significant differences. The results of the Tukey's HSD analysis are reported in Appendix G indicating that there are significant differences in the means between Group 1 (High Autonomy/High Explainability) and Group 4 (Low Autonomy/Low Explainability) related to efficiency, Lack of Human Interaction, and incompatibility. This indicates that giving

AI higher autonomy and the addition of explanations about its decisions do in fact, impact people's perceptions of efficiency, lack of human interaction, and incompatibility.

Furthermore, there are significant differences in means between Group 1 (High Autonomy/High Explainability) and Group 2 (Low Autonomy/High Explainability) for the lack of human interaction and the willingness to collaborate with AI. This means that giving AI more autonomy in making decisions does affect people's perceptions about the lack of human interaction and their willingness to collaborate with AI. All other mean differences between groups were found to be non-significant (see Appendix G).

Control Variables Analysis

This research follows the norm of other Information Systems (IS) research and has accounted for a number of variables to control for any result that may be due to extraneous factors (Archer and Cocosila 2011; Herath and Rao 2009). In this study, five control variables were included: Age, Gender, Education Level, Knowledge level of AI, and Years of Hiring Experience.

Bivariate correlations were calculated to see whether any of the control variables had a significant relationship with one or more of the endogenous variables in the model. Results are shown in Table 21, showing that **Table 22** only Age and the Knowledge Level of AI have significant relationships with all the endogenous constructs in the model. Age has a significant negative relationship with efficiency, integrity, and the willingness to collaborate, indicating that as people get older, they report lower scores for their perceptions of efficiency and integrity, and they will be less willing to collaborate with the AI. This could be due to the fact that as people age, they lose confidence in the technology and feel more comfortable following traditional, less advanced ways to complete tasks than dealing with new technologies (Barros Pena et al. 2021). In contrast, Age showed a significant positive relationship with the lack of human interaction and incompatibility, indicating that as people age, they exhibit stronger perceptions of the concerns of AI (i.e., lack of human interaction and incompatibility). Again, this could be due to the fact that as people age, they become skeptical about technology and how it complies with their needs and values, preferring collaborating with humans instead of AI.

Table 22: Control Variable and Endogenous Construct Bivariate Correlations

Control Variable		**Efficiency**	**Integrity**	**Lack_of_Human**	**Incompatibility**	**Willingness**
Age	Pearson Correlation (r)	**-.147****	**-.130***	**.108***	**.152****	**-.140****
	Sig. (2-tailed)	.005	.014	.041	.004	.008
Gender	Pearson Correlation (r)	.012	.048	.013	.003	-.008
	Sig. (2-tailed)	.821	.371	.800	.959	.883
Education Level	Pearson Correlation (r)	-.095	-.078	.028	.049	-.057
	Sig. (2-tailed)	.072	.141	.595	.359	.286
Knowledge Level of AI	Pearson Correlation (r)	**.285****	**.271****	**-.223****	**-.256****	**.336****
	Sig. (2-tailed)	.000	.000	.000	.000	.000
Years of Hiring Experience	Pearson Correlation (r)	-.067	-.069	.070	.094	-.077
	Sig. (2-tailed)	.209	.197	.188	.076	.150
** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).						

Knowledge Level of AI showed a significant negative relationship with incompatibility and lack of human interaction, indicating that the more knowledgeable people are about AI, the less they would be concerned about the incompatibility of the AI's recommendations as well as the lack of human interaction. Moreover, the Knowledge Level of AI also had a significant positive relationship with efficiency, integrity, and the willingness to collaborate. This indicates that the more knowledgeable people are about AI, the stronger their perceptions will be about efficiency and integrity, and they will be more willing to collaborate with an AI. This may be because as people become more aware of AI's capabilities, they become more confident about how AI would streamline work processes, and hence, increase perceptions of the benefits and decrease the perceptions of the concerns leading people to be more willing to collaborate with it. The remaining control variables did not show significant relationships, and therefore, their inclusion as control variables in the model may not be warranted as they would have little to no effect on the dependent variables.

In addition to testing the impact of these control variables collectively, each control variable was tested independently by running one-way ANOVAs for all the control variables with all the other endogenous variables to test for any significant mean differences between groups. The detailed results of the one-way ANOVA analysis, including Tukey's HSD, are reported in Appendix H. The results corroborate that only the Age and Knowledge Level of AI play a vital role in influencing all the benefits and concerns, as well as the willingness to collaborate with AI.

Since, from the above control variable analysis, Age and Knowledge Level of AI had significant impacts on the endogenous variables in the model, their impact on each endogenous variable in the structural model was analyzed by adding each of these two control variables one at a time to the model and linking each control variable to each endogenous variable. The significance of the paths for Age and Knowledge Level of AI were analyzed and reported in Tables 23 and 24.

Table 23: Impact of Age on the Endogenous Variables in the Model Using PLS

Path	Path Coeffifienct	P-value
Age → Efficiency	**-0.127****	**0.008**
Age → Incompatibility	**0.165****	**0.002**
Age → Integrity	**-0.106***	**0.021**
Age → Lack-of-Human	**0.114***	**0.017**
Age → Willingness	-0.001	0.488
** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).		

Table 24: Impact of Knowledge Level of AI on the Endogenous Variables in the Model Using PLS

Path	Path Coefficient	P-value
AI_Knowledge → Efficiency	**0.184*****	**0.000**
AI_Knowledge → Incompatibility	**-0.198*****	**0.001**
AI_Knowledge → Integrity	**0.175*****	**0.001**
AI_Knowledge → Lack-of-Human	**-0.169****	**0.003**
AI_Knowledge → Willingness	**0.072***	**0.022**
*** Correlation is significant at the 0.001 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).		

Results corroborate previous findings where Age had a significant negative impact on the benefits (i.e., efficiency and integrity), and had a positive relationship with the concerns (incompatibility and lack of human interaction). However, Age did not seem to have an effect on people’s willingness to collaborate with AI. On the other hand, the Knowledge Level of AI had a significant positive relationship with efficiency, integrity, and willingness to collaborate but had a negative impact on incompatibility and lack of human interaction. These findings encourage future research to investigate the role of these control variables in Human-AI collaboration. Finally, adding either of these two control variables did not change any of the hypothesized relationships in the original model (see Appendix I).

Since Age and Knowledge Level of AI had a significant impact on more than one endogenous variable in the model from the tables above, the effect size of the significant paths for the two control variables was analyzed, and the results are shown below. As mentioned before, the change in the R^2 value was evaluated as follows: 0.02 corresponds to a small effect, 0.15 corresponds to a medium effect, and 0.35 corresponds to a large effect (Henseler and Sarstedt 2013; Roldán and Sánchez-Franco 2012).

Table 25: PLS Effect Size Analysis for Significant Control Variable Paths

Control Variable	Endogenous Variables	R^2		ΔR^2	Effect Size
		Included	Excluded		
Age	Efficiency	0.266	0.251	0.015	Small
	Integrity	0.258	0.243	0.015	Small
	Lack of Human	0.136	0.123	0.013	Small
	Incompatibility	0.186	0.159	0.017	Small
AI Knowledge	Efficiency	0.282	0.251	0.031	Small
	Integrity	0.275	0.243	0.032	Small
	Lack of Human	0.149	0.123	0.026	Small
	Incompatibility	0.196	0.159	0.037	Small
	Willingness	0.756	0.751	0.005	< Small

As we can see from the above table, the effect size of adding the two control variables is small. Therefore, it could be concluded that the control variables do not change the conclusions derived from the original hypotheses of the study.

Summary

This chapter discussed the methodology that was followed to study the willingness of individuals to collaborate with AI. The quantitative method was outlined in detail, where the measurement model and the structural model were validated. Post-hoc analyses were then performed and discussed.

Chapter 7: Discussion and Conclusion

The area of Human-AI collaboration presents numerous opportunities for research. This research is only a step toward understanding the nature of collaboration in a specific context (i.e., HR context). The overarching objective of this work was to answer "what are the factors that influence humans' willingness to collaborate with AI in the workplace?" by viewing AI as a social actor in a network of humans and non-humans. In doing so, it was important to understand how people make a trade-off between the benefits and concerns when an AI is potentially introduced in their working environment and how these beliefs impact their willingness to collaborate with AI. To do so, a mixed-method approach was followed to investigate this phenomenon.

7.1. Discussion of Key Findings

Building upon ANT, this study views AI as a social actor who participates and takes an active role as a real member of a network of hybrid Human-AI collaborators. Interviewees of the qualitative study confirmed this notion by easily imagining that AI (i.e., non-human) can now be treated as inseparable from human members in an organizational setting. The study also revealed that there are pros and cons to collaborating with AI, as promoted by the NVT.

Results from the structural model uncovered that AI autonomy and AI explainability do play a vital role in influencing their perceptions when collaborating with AI. AI autonomy does have a significant influence on the lack of human interaction (β= -0.292; ρ<0.01) and incompatibility (β= -0.23; ρ= 0.05). This denotes that when workers inside organizations collaborate with AI, management should carefully consider the levels of autonomy attributed to the AI, as the more autonomous the AI is in making decisions, the more people would be concerned about the compatibility of the AI outcomes and recommendations and the more people would believe that they need human-human interaction. This is also supported by responses to open-ended questions provided by respondents who confirmed that it is important that the AI consults them before it makes a decision when they collaborate with AI-Assist. Respondents replied with statements such as:

- *"Yes, I want it to consult me first because there could always be an error overlooked. Computers make mistakes too."*
- *"The decision is ultimately mine, and my input is important in the hiring process."*
- *"Absolutely. Never give a machine authority to make any decisions."*
- *"Yes. Humans offer so much more than AI. I don't think AI could possibly be the best way to pick a qualified candidate for jobs."*
- *"Yes, I expected to be consulted before AI-Assist makes a recommendation because I need to monitor it. After all, it is a machine, and my confidence levels vary when it comes to trusting in a machine or software."*
- *"Yes, There are certain things an AI can't comprehend in my particular field."*
- *"Yes, I like how in this scenario, the AI-Assist allowed the user to review the other applications."*
- *"Yes of course. Every company has a personality, so that must also go into the equation."*
- *"The AI-Assist should allow me to choose which candidates to interview."*
- *"Yes, it would be helpful if I could help AI by shaping the AI considerations."*

- *"I would like to ultimately be in charge of the recommendation as I would be interacting with candidate/hiree, and AI doesn't take into account the human intangibles. As much as it is programmed to, it is still A, and a human is writing the program."*

AI autonomy was found to have a significant impact on people's perception of efficiency (β= 0.042; ρ<0.001) and integrity (β= 0.074; ρ<0.001). This means that people advocate that when collaborating with AI, giving AI more autonomy to complete a task increases their perception of efficiency since highly autonomous AI can make decisions quickly and process vast amounts of data in real-time. AI can also handle tasks that would require significant human effort and time. This speed and reduction in manual labour can be perceived as being more efficient, as it allows human workers to focus on more strategic or creative tasks. In addition, although AI may be, in some cases, trained on non-representative or biased data, they are perceived as being consistent in the way they make decisions as they rely on data-driven analysis rather than subjective factors. This consistency and objectivity contribute to the perception of integrity.

When examining the effect of AI explainability, it was found that it has a significant impact on people's perceptions. As many participants reported in the open-ended question, "When collaborating with AI-Assist, do you think it is important that AI provides an explanation for its recommendation?". For example, respondents highlighted the following:

- *"Absolutely. It is essential for AI to provide key points for its decision-making process."*
- *"Yes -- either before or after the determination, I would want to know the parameters used for the selection."*

- *"Yes, I think it's very important. I want to know exactly what AI thought."*
- *"Absolutely, as I would never blindly accept ANY recommendation without an explanation as to why an entity chose the way they did."*
- *"I think it is helpful. Ideally, I would consider the recommendation and the reasoning behind it and potentially use that to inform my own decision."*
- *"Yes, I don't feel confident considering their recommendation without an explanation."*
- *"Yes, otherwise, it is just picking out a candidate based solely on statistics."*
- *"I really like the concept, and it would benefit our company and will not interfere with our time as much as we need to get our jobs done."*
- *"Yes, without justification, it is less helpful and does not help identify qualities that are important in the hiring process and skill sets of the candidates."*
- *"Absolutely - I would have no confidence in AI's recommendations without that information."*

AI explainability, thus, was found to have a significant positive influence on efficiency (β=0.502; ρ<0.001) and integrity (β=0.4983; ρ<0.001). This means that the ability to understand and interpret the decision-making process of AI systems can significantly affect the overall efficiency in completing tasks. Explainability allows human users to easily verify decisions made by AI systems. When users understand the rationale behind AI decisions, they can have confidence in its output and rely on it to make informed judgments or take appropriate actions. Furthermore, explainability facilitates the identification and resolution of issues or errors in AI systems and gives humans visibility into the decision-making process and the ability to identify potential limitations in the AI system's operation. This streamlines the decision-making process, reduces unnecessary human intervention, and ultimately improves efficiency.

On the contrary, explainability was found to have a significant negative effect on the lack of human interaction (β= -0.330; ρ<0.001). This means that when the AI is able to explain its decisions, humans can better understand the rationale behind those decisions and will reduce the need for human-human consultations.

Likewise, explainability significantly impacts incompatibility (β= -0.380; ρ<0.001). This indicates that when AI is able to explain how it arrives at its decisions, humans will be able to understand the factors, data, and reasoning that contribute to AI outputs. It will also make it easier to identify and address any inconsistencies or incompatibilities found in the AI decision. Moreover, explainability allows for verifying whether AI decisions align with predefined requirements, guidelines, or regulations. If decisions are not in line with these criteria, corrective actions can be taken to adjust the AI system's behaviour and prevent incompatibility issues. Furthermore, explainability provides context around AI decisions, helping humans understand the circumstances under which certain decisions are made, which can mitigate misunderstandings that contribute to incompatibility concerns.

Based on the above discussion, we can conclude that if organizations are ready for their employees to collaborate with AI, organizations should pay careful consideration to the role of AI autonomy and explainability to mitigate people's concerns and maximize AI's perceived benefits.

In addition to the aforementioned findings, the two benefits and the two concerns were also studied in relation to their impact on the willingness to collaborate with AI. Results uncovered that all the perceived benefits do positively influence people's willingness to collaborate with AI (i.e., β=0.46; ρ<0.001 for efficiency and β=0.24; ρ<0.001 for integrity), while only the lack of human interaction negatively affects it (i.e., β= -0.26; ρ<0.001) but not the incompatibility (β= -0.24; ρ=

0.52). This delineates that when integrating AI as collaborators in the workplace, organizations have to ensure that AI improves efficiency and integrity to encourage a collaborative environment with humans. Furthermore, organizations should be wary of people's need to have a human-in-the-loop when making decisions since the lack of human interaction negatively influences people's willingness to collaborate with AI.

7.2. Contributions to Theory

This study contributes to the theory and literature in numerous ways. First, it is a pioneering study in conducting a mixed-method approach to understanding an emerging phenomenon such as Human-AI collaboration. The study contributes to theory through its insightful qualitative findings that distilled managers' and hiring employees' perceptions about AI collaborators. Interviewing 25 candidates resulted in highlighting the main benefits and concerns that people believe would influence their collaboration with AI in the workplace.

The qualitative conclusions were then utilized to empirically examine how they may affect humans' willingness to collaborate with AI in the workplace. It is an initial step towards identifying employees' perceptions of AI not as a facilitating assistive technology but as collaborators inside organizations. Findings from the qualitative and quantitative parts of my study corroborated the ideas raised in the literature about the potential benefits of Human-AI symbiosis. When both collaborate, people would expect higher levels of efficiency and integrity in completing tasks. However, it is important to balance responsibilities as people still need the human element and are yet skeptical about AI decisions' compatibility. Therefore, this study addressed the impact of efficiency, integrity, lack of human interaction and incompatibility on people's willingness to collaborate with AI. In doing so, I showcased that the lack of human interaction along with the two benefits are important factors to consider when asking employees inside organizations to collaborate with a newly hired AI.

Second, grounded in ANT and NVT, this study contributes to the IS literature by leveraging well-established theories to examine how contextualized AI characteristics, such as AI autonomy and AI explainability, would influence humans' beliefs and, in turn, their willingness to collaborate with AI in organizational settings. ANT is a sociological framework that explores the interactions and relationships between human and non-human actors within a network. Whereas, NVT examines the positive or negative valence associated with different elements within a network. AI autonomy can contribute to ANT by introducing autonomous AI systems as actors in the network that can independently make decisions, take actions, and interact with others. This is an important element to examine since the increased advancements in AI are predicted to replace many humans in several jobs. Hence, the autonomy given to an AI collaborator would shape people's perceptions about it.

AI explainability also allows human actors to understand the reasoning and processes behind AI decisions and facilitates negotiation, collaboration, and trust-building between human and non-human actors within the network. Through explainability, human actors can integrate AI systems into their decision-making processes, evaluate their impact, and make informed choices about their interactions with these autonomous actors. In addition, several efforts are now directed toward promoting transparent and fair AI systems that can take out many of the implicit biases inherent in humans when making decisions.

Finally, this study contributes to the IS literature by conceptualizing AI autonomy, AI explainability, AI collaborators, and the willingness to collaborate with an AI actor. Weber (2003) argued that one of the ways to make theoretical contributions is by defining constructs more precisely or by conceptualizing them in different ways.

7.3. Contributions to Practice

Besides the theoretical influences elaborated above, this research also has important implications for practitioners. The empirical findings of this work provide developers and designers of AI with guidelines that should be taken into consideration when designing AI that will collaborate with humans in the workplace.

Understanding the influence of AI autonomy and explainability on efficiency helps organizations optimize their processes and workflows. AI autonomy can enhance efficiency by automating repetitive tasks, accelerating decision-making, and streamlining operations. However, it is essential to strike a balance, as excessive autonomy without human oversight may introduce incompatible outcomes. Additionally, explainability aids in identifying bottlenecks, refining AI, and improving overall efficiency and incompatibility by providing insights into AI decision-making and enabling performance evaluation. Besides, organizations need to ensure that autonomous AI systems maintain integrity by addressing biases, ensuring compliance with regulations, and adhering to ethical guidelines. Explainability plays a crucial role in maintaining integrity by enabling organizations to understand and verify AI decisions.

While AI can improve efficiency, excessive automation that leads to eliminating human interaction may hinder collaboration. Human interaction is vital for creative problem-solving, critical thinking, and empathy. Organizations should carefully balance AI autonomy with human involvement to ensure meaningful collaboration. By leveraging explainability, organizations can also facilitate Human-AI interaction by providing insights into AI reasoning and enabling effective communication. This fosters a collaborative environment where AI systems and humans can complement each other's strengths, leading to more successful outcomes.

7.4. Limitations and Future Work

This work represents early research into the domain of Human-AI collaboration and, therefore, has several limitations that might point to areas for future research. First, the context of this study focuses on the HR domain. There are many other sectors and domains that can benefit from this research (e.g., healthcare, banking). As such, the results of this study can only be applied to other contexts with appropriate refinements. Furthermore, I believe that the concerns and benefits elicited from the qualitative study in the HR context could be different if the scenario and interview discussion targeted different people from different domains. Also, this work reflected only the perception of people who are managers or employees with hiring experience. Thus, conclusions about the benefits and concerns may be different for different stakeholders. This means that contextualization is important to understand the Human-AI symbiosis better.

Second, while the scenario-based experimental study in this research mimics a real-world hiring process and every provision was taken to ensure that participants perceived the portrayed scenarios as realistic as possible, still, participants had to imagine their collaboration with an AI on a fictitious HR task. Thus, the findings from this study are based on a simulation as opposed to a real collaboration. Future research can investigate the phenomenon in real settings where a real AI is used. This also opens new avenues of research to study how the different types of AI (e.g., virtual vs. physical) and its anthropomorphic features (e.g., gender, skin colour, tone), if any, would influence people's perceptions about it and their readiness to collaborate with it in organizations. Another avenue of research could also be directed to examine how an AI collaborator would incorporate the idea of equity, diversity, and inclusion (EDI) in a workplace setting as a key element that organizations have to consider when hiring new employees nowadays (Bernstein et al. 2020; Ferraro et al. 2023).

Third, I believe that the type of task participants were asked to imagine plays a vital role in shaping their responses. In this research, task type and complexity were not manipulated or examined. That said, it is interesting to study how different tasks with different levels of complexity can influence people's perceptions about the benefits and concerns, as well as their willingness to collaborate with AI. In addition, I presumed that the recruitment task chosen for this study resembles a task that has routine phases as well as requires human judgment. This means that the findings of this study can differ for other types of tasks that may contain only repetitive tasks that are easy to be automated and done entirely by an AI.

Fourth, this work focuses only on the willingness of humans to collaborate with AI in the pre-collaboration phase since this is the phase organizations are interested in right now. However, as humans interact and collaborate with AI in a workplace setting, their perceptions might change, and the impacts of different variables implemented in this study may not be the same in a post-collaboration stage. Thus, future research may direct some efforts toward studying how human collaborators' perceptions would change before and after collaboration. It is also worth studying their perceptions of the benefits and concerns as some new benefits and concerns may arise, and some old ones may need to be turned down.

Fifth, this work only examines humans' willingness to collaborate with AI and not the other way around (i.e., an AI's willingness to collaborate with humans). This could be interesting for future research which can study AI's perceptions when collaborating with humans.

Sixth, the control variables collected from this study showed significant influence on many of the hypothesized relationships. This could also be of interest to researchers who wish to investigate them more or wish to see their effect on the Human-AI phenomenon in a different context.

Seventh, this study focuses on two AI characteristics (i.e., AI autonomy and explainability). However, other factors, such as situational or task characteristics, may result in different findings and may report other kinds of individual beliefs about collaborating with AI. Though important, they were beyond the scope of this study and should be explored in future research.

Eighth, this study was limited to organizational settings. Future research could be conducted in other areas, such as humans collaborating with AI for education, learning, entertainment, hobbies, travel, personal finance, personal banking, etc.

Finally, future research should focus on developing frameworks and guidelines for effective collaboration between humans and AI systems. This involves understanding the roles, responsibilities, and decision-making processes of both humans and AI and identifying the optimal division of labour. Research should explore how to integrate AI seamlessly into existing workflows and develop interfaces and interaction mechanisms that facilitate smooth collaboration. Besides, future research should focus on developing AI systems that can adapt to different contexts, user preferences, and collaborative dynamics. AI systems need to be able to understand and respond to human behaviour, preferences, and feedback in real-time, enabling personalized and contextually appropriate interactions. Investigating techniques for adaptive AI that can adjust its behaviour and decision-making based on the evolving needs of collaboration is also essential.

7.5. Conclusion

Artificial Intelligence (AI) has emerged as a transformative technology that has permeated various domains of society, revolutionizing the way we live and work. With its ability to mimic human cognitive processes and perform complex tasks, AI has become a driving force behind advancements in fields such as healthcare, finance, transportation, and many others. The rapid progress in AI research and development has led to an increasing interest in understanding its capabilities, limitations, and societal implications. AI refers to the creation of intelligent machines that can perform tasks requiring human-like intelligence, such as visual perception, speech recognition, decision-making, and problem-solving. These intelligent machines are built using algorithms that enable them to learn from data, adapt to new information, and make predictions or take actions. The field of AI encompasses various sub-disciplines, including ML, natural language processing, computer vision, and robotics, each contributing to the broader goal of creating intelligent systems.

The advent of AI has introduced the concepts of autonomy and explainability, which have significant implications for collaboration with AI systems. AI autonomy, in this context, refers to the ability of AI systems to make decisions on their own, while AI explainability relates to the ability of the AI to provide explanations about how and why it made certain decisions or took specific actions. Such a nascent phenomenon drives people to make a trade-off between the benefits and concerns that can arise from collaborating with them on a certain task. To understand this trade-off, this study employs a mixed-methods approach in order to instill rich insights about people's perceptions. A qualitative study was conducted first with 25 interviewees who are managers or have hiring experience to better understand people's beliefs about the potential benefits and concerns of collaborating with an AI in an HR recruitment context. The qualitative interview findings showed that efficiency and integrity are the two most important benefits an AI collaborator can offer its human counterparts.

AI has the potential to greatly impact efficiency in a multitude of industries. By automating repetitive and time-consuming tasks, AI systems can enhance productivity, reduce errors, and improve overall operational efficiency (Brynjolfsson and McAfee 2014; Vassilakopoulou et al. 2023). Moreover, AI's ability to analyze vast amounts of data and extract valuable insights enables organizations to make data-driven decisions and gain a competitive edge in dynamic market environments (Davenport and Ronanki 2018).

Another salient benefit of AI is its ability to mitigate bias in decision-making. Research by Caliskan et al. (2017) highlights the potential of AI systems to reduce bias by relying on algorithms and data-driven approaches. Unlike human decision-makers who may be influenced by personal beliefs, prejudices, or cognitive biases, AI systems can operate objectively and consistently. By analyzing vast amounts of data and applying statistical models, AI systems have the potential to identify and mitigate biases in decision-making processes, promoting fairness and equality (Shin 2021b). Furthermore, AI's ability to learn from diverse data sources and consider multiple perspectives can contribute to a more inclusive decision-making process, ultimately reducing the impact of human biases. Explainability also allows human collaborators to understand the reasoning behind AI decisions and detect potential biases or errors. This transparency fosters trust and confidence in AI systems, ultimately enhancing their integrity within collaborative environments.

One notable application of AI where it can eliminate bias is in the recruitment and hiring process. Research by (Bertrand and Mullainathan 2004) demonstrates that human decision-makers often exhibit biases based on factors such as gender, race, or socioeconomic background, leading to discriminatory outcomes. AI recruitment systems, on the other hand, can be designed to focus on relevant qualifications and skills while disregarding personal characteristics that may introduce bias. By analyzing candidate profiles and historical data, AI systems can help identify the most suitable candidates based on objective criteria, thereby reducing bias and promoting a fairer selection process. These AI-driven approaches hold the potential to create more diverse and inclusive workplaces.

However, alongside the promises of AI, there are also concerns regarding the potential lack of human interaction in various domains and incompatibility with an organization's needs and values are the two most reported concerns from interviewees. While AI systems can handle tasks efficiently, Dabbish and Kraut (2004) have highlighted the significance of human-human interaction for effective collaboration. While AI systems can automate certain tasks, human interaction remains crucial for creative and complex problem-solving, brainstorming, social interaction, and building relationships. Therefore, striking a balance between AI autonomy and human collaboration is essential to ensure successful and productive partnerships to harness the full potential of collaboration. Furthermore, the integration of AI into existing infrastructures and practices may present challenges of incompatibility (Lee and Lyu 2019). Developing AI that is able to adapt to organizational structures, processes, workflows, and values may require significant investments and changes, potentially leading to resistance and reluctance to collaborate with such technologies. Rogers (2003) recommends that once an innovation is compatible, it is highly likely to be used, and hence, it becomes an important feature of such innovation. Therefore, individuals would negatively evaluate AI collaborators if AI produces recommendations that are incompatible with their values and needs.

References

Abdelhalim, E., Hassanein, K., and Head, M. 2019. "The Role of Knowledge Control and Knowledge Asymmetry in Trusting and Collaborating with AI-Teammates," *SIGHCI 2019 Proceedings*. (https://aisel.aisnet.org/sighci2019/7).

Adadi, A., and Berrada, M. 2018. "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access* (6), pp. 52138–52160. (https://doi.org/10.1109/ACCESS.2018.2870052).

Adam, A. 2018. "Meet Pepper – Canada's First Emotionally Sensitive Robot for Sick Kids," *Global News*. (https://globalnews.ca/news/4180025/pepper-canada-robot/, accessed November 22, 2019).

Adomavicius, G., Gupta, A., and Zhdanov, D. 2009. "Designing Intelligent Software Agents for Auctions with Limited Information Feedback," *Information Systems Research* (20:4), Linthicum: INFORMS, pp. 507–526. (https://doi.org/10.1287/isre.1080.0172).

Ahuja, M. K., and Thatcher, J. B. 2005. "Moving beyond Intentions and toward the Theory of Trying: Effects of Work Environment and Gender on Post-Adoption Information Technology Use," *MIS Quarterly* (29:3), Management Information Systems Research Center, University of Minnesota, pp. 427–459. (https://doi.org/10.2307/25148691).

Ajunwa, I. 2020. "The 'Black Box' at Work," *Big Data & Society* (7:2), SAGE Publications Ltd, p. 2053951720966181. (https://doi.org/10.1177/2053951720938093).

Akmeikina, E., Eilers, K., Li, M. M., and Peters, C. 2022. *Empowerment Effects in Human-Machine Collaboration - A Systematic Literature Review and Directions on Hybrid Intelligence Behavior Patterns*.

Akter, S., D'Ambra, J., and Ray, P. 2011. "AN EVALUATION OF PLS BASED COMPLEX MODELS: THE ROLES OF POWER ANALYSIS, PREDICTIVE RELEVANCE AND GOF INDEX," *AMCIS 2011 Proceedings - All Submissions*. (https://aisel.aisnet.org/amcis2011_submissions/151).

Alsharo, M., Gregg, D., and Ramirez, R. 2017. "Virtual Team Effectiveness: The Role of Knowledge Sharing and Trust," *Information & Management* (54:4), pp. 479–490. (https://doi.org/10.1016/j.im.2016.10.005).

Alvesson, M., and Kärreman, D. 2007. "Constructing Mystery: Empirical Matters in Theory Development," *The Academy of Management Review* (32:4), pp. 1265–1281. (https://doi.org/10.2307/20159366).

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., and Horvitz, E. 2019. "Guidelines for Human-AI Interaction," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow Scotland Uk: ACM, May 2, pp. 1–13. (https://doi.org/10.1145/3290605.3300233).

Andrade, I. M. D., and Tumelero, C. 2022. "Increasing Customer Service Efficiency through Artificial Intelligence Chatbot," *Revista de Gestão* (29:3), Emerald Publishing Limited, pp. 238–251. (https://doi.org/10.1108/REGE-07-2021-0120).

André, Q., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., Huber, J., van Boven, L., Weber, B., and Yang, H. 2018. "Consumer Choice and Autonomy in the Age of Artificial Intelligence and Big Data," *Customer Needs and Solutions* (5:1–2), New York: Springer US, pp. 28–37. (https://doi.org/10.1007/s40547-017-0085-8).

Archer, N., and Cocosila, M. 2011. "A Comparison of Physician Pre-Adoption and Adoption Views on Electronic Health Records in Canadian Medical Practices," *Journal of Medical Internet Research* (13:3), p. e1726. (https://doi.org/10.2196/jmir.1726).

Arya, V., Bellamy, R. K. E., Chen, P.-Y., Dhurandhar, A., Hind, M., Hoffman, S. C., Houde, S., Liao, Q. V., Luss, R., Mojsilovic, A., Mourad, S., Pedemonte, P., Raghavendra, R., Richards, J. T., Sattigeri, P., Shanmugam, K., Singh, M., Varshney, K. R., Wei, D., and Zhang, Y. 2020. "AI Explainability 360: An Extensible Toolkit for Understanding Data and Machine Learning Models," *Journal of Machine Learning Research 21(130)*, p. 6.

Ashfaq, M., Yun, J., Yu, S., and Loureiro, S. M. C. 2020. "I, Chatbot: Modeling the Determinants of Users' Satisfaction and Continuance Intention of AI-Powered Service Agents," *Telematics and Informatics* (54), p. 101473. (https://doi.org/10.1016/j.tele.2020.101473).

Azodi, C. B., Tang, J., and Shiu, S.-H. 2020. "Opening the Black Box: Interpretable Machine Learning for Geneticists," *Trends in Genetics* (36:6), pp. 442–455. (https://doi.org/10.1016/j.tig.2020.03.005).

Babbie, E. R. 1998. *The Practice of Social Research*, (8th ed.), Belmont, CA: Wadsworth Pub. Co.

Bagozzi, R. P., and Yi, Y. 1988. "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science* (16:1), pp. 74–94. (https://doi.org/10.1007/BF02723327).

Baird, A., and Maruping, L. 2021. "The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts," *Management Information Systems Quarterly* (45:1), pp. 315–341.

Barber, K. S., and Martin, C. E. 1999. "Agent Autonomy: Specification, Measurement, and Dynamic Adjustment," in *In Proceedings of the Autonomy Control Software Workshop, Agents '99*, pp. 8–15.

Barros Pena, B., Clarke, R. E., Holmquist, L. E., and Vines, J. 2021. "Circumspect Users: Older Adults as Critical Adopters and Resistors of Technology," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA: Association for Computing Machinery, May 7, pp. 1–14. (https://doi.org/10.1145/3411764.3445128).

Bartol, K. M., and Srivastava, A. 2002. "Encouraging Knowledge Sharing: The Role of Organizational Reward Systems," *Journal of Leadership & Organizational Studies* (9:1), pp. 64–76. (https://doi.org/10.1177/107179190200900105).

Beale, R., and Wood, A. 1994. “Agent-Based Interaction,” in *In Proceedings of People and Computers IX: Proceedings of HCI’94*, Cambridge: Cambridge University Press. (https://doi.org/10.1017/CBO9780511600821).

Bedwell, W. L., Wildman, J. L., DiazGranados, D., Salazar, M., Kramer, W. S., and Salas, E. 2012. “Collaboration at Work: An Integrative Multilevel Conceptualization,” *Human Resource Management Review* (22:2), Construct Clarity in Human Resource Management Research, pp. 128–145. (https://doi.org/10.1016/j.hrmr.2011.11.007).

Beer, J. M., Fisk, A. D., and Rogers, W. A. 2014. “Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction,” *Journal of Human-Robot Interaction* (3:2), p. 74. (https://doi.org/10.5898/JHRI.3.2.Beer).

Behl, A. 2020. “Antecedents to Firm Performance and Competitiveness Using the Lens of Big Data Analytics: A Cross-Cultural Study,” *Management Decision* (60:2), Emerald Publishing Limited, pp. 368–398. (https://doi.org/10.1108/MD-01-2020-0121).

Benbasat, I., Goldstein, D. K., and Mead, M. 1987. “The Case Research Strategy in Studies of Information Systems,” *MIS Quarterly* (11:3), pp. 369–386. (https://doi.org/10.2307/248684).

Benbya, H., Deakin University, Melbourne, Australia, Pachidi, S., Cambridge JudgeBusiness School, University of Cambridge, United Kingdom, Jarvenpaa, S. L., and McCombs School of Business, University of Texas at Austin, U.S.A. 2021. “Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research,” *Journal of the Association for Information Systems* (22:2), pp. 281–303. (https://doi.org/10.17705/1jais.00662).

Bentler, P. M., and Bonett, D. G. 1980. “Significance Tests and Goodness of Fit in the Analysis of Covariance Structures,” *Psychological Bulletin* (88:3), US: American Psychological Association, pp. 588–606. (https://doi.org/10.1037/0033-2909.88.3.588).

Bernstein, R. S., Bulger, M., Salipante, P., and Weisinger, J. Y. 2020. “From Diversity to Inclusion to Equity: A Theory of Generative Interactions,” *Journal of Business Ethics* (167:3), pp. 395–410. (https://doi.org/10.1007/s10551-019-04180-1).

Bertrand, M., and Mullainathan, S. 2004. “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination,” *American Economic Review* (94:4), pp. 991–1013. (https://doi.org/10.1257/0002828042002561).

Bhattacherjee, A. 2002. “Individual Trust in Online Firms: Scale Development and Initial Test,” *Journal of Management Information Systems* (19:1), Taylor & Francis, Ltd., pp. 211–241.

Bilkey, W. J. 1953. “A Psychological Approach to Consumer Behavior Analysis,” *Journal of Marketing* (18:1), pp. 18–25. (https://doi.org/10.1177/002224295301800103).

Biran, O., and Cotton, C. 2017. *Explanation and Justification in Machine Learning: A Survey*, p. 6.

Blattberg, R. C., and Hoch, S. J. 1990. “Database Models and Managerial Intuition: 50${\tt\%}$ Model + 50${\tt\%}$ Manager,” *Management Science* (36:8), pp. 887–899.

van den Bosch, K., and Bronkhorst, A. 2018. "Human-AI Cooperation to Benefit Military Decision Making," *Proceedings of the Special Meeting on Big Data, Artificial Intelligence, and Military Decision Making*, p. 13.

Bostandjiev, S., O'Donovan, J., and Höllerer, T. 2012. "TasteWeights: A Visual Interactive Hybrid Recommender System," in *Proceedings of the Sixth ACM Conference on Recommender Systems*, RecSys '12, New York, NY, USA: Association for Computing Machinery, September 9, pp. 35–42. (https://doi.org/10.1145/2365952.2365964).

Boudreau, M.-C., Gefen, D., and Straub, D. W. 2001. "Validation in Information Systems Research: A State-of-the-Art Assessment," *MIS Quarterly* (25:1), pp. 1–16.

Breward, M., Hassanein, K., and Head, M. 2017. "Understanding Consumers' Attitudes Toward Controversial Information Technologies: A Contextualization Approach," *Information Systems Research* (28:4), pp. 760–774. (https://doi.org/10.1287/isre.2017.0706).

Bright, Z., and Harry Asada, H. 2017. "Supernumerary Robotic Limbs for Human Augmentation in Overhead Assembly Tasks," in *Robotics: Science and Systems XIII*, Robotics: Science and Systems Foundation, July 12. (https://doi.org/10.15607/RSS.2017.XIII.062).

Brown, S. A., Dennis, A. R., and Venkatesh, V. 2010. "Predicting Collaboration Technology Use: Integrating Technology Adoption and Collaboration Research," *Journal of Management Information Systems* (27:2), pp. 9–54. (https://doi.org/10.2753/MIS0742-1222270201).

Brown, S. M., Santos, E., Banks, S. B., and Oxley, M. E. 1998. "Using Explicit Requirements and Metrics for Interface Agent User Model Correction," in *Proceedings of the Second International Conference on Autonomous Agents*, AGENTS '98, Minneapolis, Minnesota, USA: Association for Computing Machinery, May 1, pp. 1–7. (https://doi.org/10.1145/280765.280767).

Brynjolfsson, E., and McAfee, A. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W. W. Norton & Company.

Brynjolfsson, E., and Mitchell, T. 2017. "What Can Machine Learning Do? Workforce Implications," *Science* (358:6370), American Association for the Advancement of Science, pp. 1530–1534. (https://doi.org/10.1126/science.aap8062).

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Henke, N., and Trench, M. 2017. "How Artificial Intelligence Can Deliver Real Value to Companies | McKinsey & Company." (https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-artificial-intelligence-can-deliver-real-value-to-companies, accessed August 13, 2018).

Buhrmester, V., Münch, D., and Arens, M. 2019. "Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey," *arXiv:1911.12116 [Cs]*. (http://arxiv.org/abs/1911.12116).

Burrell, J. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society* (3:1), p. 205395171562251. (https://doi.org/10.1177/2053951715622512).

Caliskan, A., Bryson, J. J., and Narayanan, A. 2017. “Semantics Derived Automatically from Language Corpora Contain Human-like Biases,” *Science* (356:6334), American Association for the Advancement of Science, pp. 183–186. (https://doi.org/10.1126/science.aal4230).

Carter, L., and Bélanger, F. 2005. “The Utilization of E-Government Services: Citizen Trust, Innovation and Acceptance Factors*,” *Information Systems Journal* (15:1), pp. 5–25. (https://doi.org/10.1111/j.1365-2575.2005.00183.x).

Casey, B. 2018. “Rethinking Explainable Machines: The Next Chapter in the GDPR’s ‘Right to Explanation’ Debate,” *Oxford Law Faculty*, , May 15. (https://www.law.ox.ac.uk/business-law-blog/blog/2018/05/rethinking-explainable-machines-next-chapter-gdprs-right-explanation, accessed January 14, 2020).

Castellanos, S. 2018. “Companies Grapple With AI’s Opaque Decision-Making Process,” *WSJ*, , May 2. (https://blogs.wsj.com/cio/2018/05/02/companies-grapple-with-ais-opaque-decision-making-process/, accessed January 28, 2020).

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., and Floridi, L. 2018. “Artificial Intelligence and the ‘Good Society’: The US, EU, and UK Approach,” *Science and Engineering Ethics* (24:2), pp. 505–528. (https://doi.org/10.1007/s11948-017-9901-7).

Cazier, J. A., Jensen, A. S., and Dave, D. S. 2008. “The Impact of Consumer Perceptions of Information Privacy and Security Risks on the Adoption of Residual RFID Technologies,” *Communications of the Association for Information Systems* (23). (https://doi.org/10.17705/1CAIS.02314).

Charmaz, K. 2014. *Constructing Grounded Theory*, SAGE.

Chin, W. W., Marcolin, B. L., and Newsted, P. R. 2003. “A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study,” *Information Systems Research* (14:2), INFORMS, pp. 189–217. (https://doi.org/10.1287/isre.14.2.189.16018).

Chin, W. W., Thatcher, J. B., Wright, R. T., and Steel, D. 2013. *Controlling for Common Method Variance in PLS Analysis: The Measured Latent Marker Variable Approach*, Springer Proceedings in Mathematics & Statistics, New York, NY: Springer New York, pp. 231–239. (https://doi.org/10.1007/978-1-4614-8283-3_16).

Claudé, M., and Combe, D. 2018. *THE ROLES OF ARTIFICIAL INTELLIGENCE AND HUMANS IN DECISION MAKING: TOWARDS AUGMENTED HUMANS? : A Focus on Knowledge-Intensive Firms.* (http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-150022).

Cohen, B. H. 2008. *Explaining Psychological Statistics*, John Wiley & Sons.

Cohen, J. 2013. *Statistical Power Analysis for the Behavioral Sciences*, Academic Press.

Coldewey, D. 2019. “Blind Users Can Now Explore Photos by Touch with Microsoft’s Seeing AI,” *TechCrunch*. (http://social.techcrunch.com/2019/03/12/blind-users-can-now-explore-photos-by-touch-with-microsofts-seeing-ai/, accessed November 22, 2019).

Collier, J., and Sherrell, D. 2010. "Examining the Influence of Control and Convenience in a Self-Service Setting," *Journal of the Academy of Marketing Science* (38), pp. 490–509. (https://doi.org/10.1007/s11747-009-0179-4).

Correll, N., Bekris, K. E., Berenson, D., Brock, O., Causo, A., Hauser, K., Okada, K., Rodriguez, A., Romano, J. M., and Wurman, P. R. 2018. "Analysis and Observations From the First Amazon Picking Challenge," *IEEE Transactions on Automation Science and Engineering* (15:1), pp. 172–188. (https://doi.org/10.1109/TASE.2016.2600527).

Covrigaru, A. A., and Lindsay, R. K. 1991. "Deterministic Autonomous Systems," *AI Magazine* (12:3), pp. 110–110. (https://doi.org/10.1609/aimag.v12i3.907).

de Cremer, D., and Kasparov, G. 2021. "AI Should Augment Human Intelligence, Not Replace It," *Harvard Business Review*. (https://hbr.org/2021/03/ai-should-augment-human-intelligence-not-replace-it, accessed August 15, 2023).

Creswell, J. W. 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, (3rd ed.), Thousand Oaks, Calif: Sage Publications.

Creswell, J. W., and Creswell, J. D. 2018. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, (Fifth edition.), Thousand Oaks, California: SAGE Publications, Inc.

Dabbish, L., and Kraut, R. 2004. *Controlling Interruptions: Awareness Displays and Social Motivation for Coordination*, CSCW '04, ACM, pp. 182–191. (https://doi.org/10.1145/1031607.1031638).

Dabholkar, P. A. 1992. "Role of Affect and Need For Interaction in On-Site Service Encounters," *ACR North American Advances* (NA-19). (https://www.acrwebsite.org/volumes/7357/volumes/v19/NA-19/full).

Dabholkar, P. A., and Bagozzi, R. P. 2002. "An Attitudinal Model of Technology-Based Self-Service: Moderating Effects of Consumer Traits and Situational Factors," *Journal of the Academy of Marketing Science* (30:3), SAGE Publications Inc, pp. 184–201. (https://doi.org/10.1177/0092070302303001).

Danks, D., and London, A. J. 2017. "Algorithmic Bias in Autonomous Systems," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, Australia: International Joint Conferences on Artificial Intelligence Organization, August, pp. 4691–4697. (https://doi.org/10.24963/ijcai.2017/654).

Daugherty, P. R., and Wilson, H. J. 2018. *Human + Machine: Reimagining Work in the Age of AI*, Harvard Business Press.

Davenport, T. H., and Ronanki, R. 2018. "Don't Start with Moon Shots.," *Harvard Business Review* (96(1)), pp. 108–116.

Deci, E. L., Connell, J. P., and Ryan, R. M. 1989. "Self-Determination in a Work Organization," *Journal of Applied Psychology* (74:4), Washington, DC: American Psychological Association, pp. 580–590. (https://doi.org/10.1037/0021-9010.74.4.580).

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., and Ebel, P. 2019. *The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems*. (https://doi.org/10.24251/HICSS.2019.034).

Denecke, K., Vaaheesan, S., and Arulnathan, A. 2021. “A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability Test,” *IEEE Transactions on Emerging Topics in Computing* (9:3), pp. 1170–1182. (https://doi.org/10.1109/TETC.2020.2974478).

Deyo, K. 2020. “Artificial Intelligence: Empowering People, Not The Rise Of The Machines.” (https://www.orchestracms.com/Artificial-Intelligence-Empowering-People-Not-The-Rise-Of-The-Machines, accessed June 23, 2023).

Diab, D. L., Pui, S.-Y., Yankelevich, M., and Highhouse, S. 2011. “Lay Perceptions of Selection Decision Aids in US and Non-US Samples,” *International Journal of Selection and Assessment* (19:2), pp. 209–216. (https://doi.org/10.1111/j.1468-2389.2011.00548.x).

Dietvorst, B. J., Simmons, J. P., and Massey, C. 2015. “Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err,” *Journal of Experimental Psychology: General* (144:1), US: American Psychological Association, pp. 114–126. (https://doi.org/10.1037/xge0000033).

Dijkstra, T. K., and Henseler, J. 2015. “Consistent Partial Least Squares Path Modeling,” *MIS Quarterly* (39:2), Management Information Systems Research Center, University of Minnesota, pp. 297–316.

Dinev, T., and Hart, P. 2006. “An Extended Privacy Calculus Model for E-Commerce Transactions,” *Information Systems Research* (17:1), pp. 61–80. (https://doi.org/10.1287/isre.1060.0080).

Dissanayake, I., Zhang, J., and Gu, B. 2015. “Task Division for Team Success in Crowdsourcing Contests: Resource Allocation and Alignment Effects,” *Journal of Management Information Systems* (32:2), pp. 8–39. (https://doi.org/10.1080/07421222.2015.1068604).

Dodd, N. G., and Ganster, D. C. 1996. “The Interactive Effects of Variety, Autonomy, and Feedback on Attitudes and Performance,” *Journal of Organizational Behavior* (17:4), pp. 329–347. (https://doi.org/10.1002/(SICI)1099-1379(199607)17:4<329::AID-JOB754>3.0.CO;2-B).

Dworkin, G. 1988. “The Theory and Practice of Autonomy,” *Cambridge University Press*.

Eastwood, J., Snook, B., and Luther, K. 2012. “What People Want From Their Professionals: Attitudes Toward Decision-Making Strategies,” *Journal of Behavioral Decision Making* (25:5), pp. 458–468. (https://doi.org/10.1002/bdm.741).

Ebrahimi, S., and Hassanein, K. 2021. “Decisional Guidance for Detecting Discriminatory Data Analytics Recommendations,” *Information & Management* (58:7), p. 103520. (https://doi.org/10.1016/j.im.2021.103520).

Eeuwen, M. van. 2017. “Mobile Conversational Commerce: Messenger Chatbots as the next Interface between Businesses and Consumers,” Info:eu-repo/semantics/masterThesis, University of Twente, February 8. (http://essay.utwente.nl/71706/, accessed June 14, 2023).

Eilers, K., Simmert, B., and Peters, C. 2020. *Doing Agile vs. Being Agile - Understanding Their Effects to Improve Agile Work*.

Einhorn, H. J. 1974. “Cue Definition and Residual Judgment,” *Organizational Behavior and Human Performance* (12:1), pp. 30–49. (https://doi.org/10.1016/0030-5073(74)90035-X).

Elo, S., and Kyngäs, H. 2008. “The Qualitative Content Analysis Process,” *Journal of Advanced Nursing* (62:1), pp. 107–115. (https://doi.org/10.1111/j.1365-2648.2007.04569.x).

Emerson, R. M. 1976. “Social Exchange Theory,” *Annual Review of Sociology* (2:1), pp. 335–362. (https://doi.org/10.1146/annurev.so.02.080176.002003).

Ernst, C.-P., and Reinelt, P. 2017. “Autonomous Car Acceptance: Safety vs. Personal Driving Enjoyment,” *AMCIS 2017 Proceedings*. (https://aisel.aisnet.org/amcis2017/AdvancesIS/Presentations/7).

Esposito Vinzi, V., Chin, W. W., Henseler, J., and Wang, H. 2010. *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer Handbooks of Computational Statistics, Berlin, Heidelberg: Springer Berlin / Heidelberg.

Ewe, S. Y., Yap, S. F., and Lee, C. K. C. 2015. “Network Externalities and the Perception of Innovation Characteristics: Mobile Banking,” *Marketing Intelligence & Planning* (33:4), Emerald Group Publishing Limited, pp. 592–611. (https://doi.org/10.1108/MIP-01-2014-0006).

Ezenkwu, C. P., and Starkey, A. 2019. “Machine Autonomy: Definition, Approaches, Challenges and Research Gaps,” in *Intelligent Computing*, Advances in Intelligent Systems and Computing, K. Arai, R. Bhatia, and S. Kapoor (eds.), Cham: Springer International Publishing, pp. 335–358. (https://doi.org/10.1007/978-3-030-22871-2_24).

Falk, R. F., and Miller, N. B. 1992. *A Primer for Soft Modeling*, A Primer for Soft Modeling, Akron, OH, US: University of Akron Press, pp. xiv, 103.

Fan, M., Yang, X., Yu, T., Liao, Q. V., and Zhao, J. 2022. “Human-AI Collaboration for UX Evaluation: Effects of Explanation and Synchronization,” *Proceedings of the ACM on Human-Computer Interaction* (6:CSCW1), 96:1-96:32. (https://doi.org/10.1145/3512943).

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. 2009. “Statistical Power Analyses Using G*Power 3.1: Tests for Correlation and Regression Analyses,” *Behavior Research Methods* (41:4), pp. 1149–1160. (https://doi.org/10.3758/BRM.41.4.1149).

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. 2007. “G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences,” *Behavior Research Methods* (39:2), pp. 175–191. (https://doi.org/10.3758/BF03193146).

Featherman, M. 2001. *Extending the Technology Acceptance Model by Inclusion of Perceived Risk*, p. 4.

Ferguson, C. J. 2016. *An Effect Size Primer: A Guide for Clinicians and Researchers*, Methodological Issues and Strategies in Clinical Research, 4th Ed, Washington, DC, US: American Psychological Association, p. 310. (https://doi.org/10.1037/14805-020).

Fernández-Macías, E., Gómez, E., Hernández-Orallo, J., Loe, B. S., Martens, B., Martínez-Plumed, F., and Tolan, S. 2018. "A Multidisciplinary Task-Based Perspective for Evaluating the Impact of AI Autonomy and Generality on the Future of Work," *arXiv:1807.02416 [Cs]*. (http://arxiv.org/abs/1807.02416).

Ferraro, C., Hemsley, A., and Sands, S. 2023. "Embracing Diversity, Equity, and Inclusion (DEI): Considerations and Opportunities for Brand Managers," *Business Horizons* (66:4), pp. 463–479. (https://doi.org/10.1016/j.bushor.2022.09.005).

Fitts, P. M. 1951. "Human Engineering for an Effective Air-Navigation and Traffic-Control System," Oxford, England: National Research Council, Div. of, pp. xxii, 84.

Fogliato, R., Chappidi, S., Lungren, M., Fisher, P., Wilson, D., Fitzke, M., Parkinson, M., Horvitz, E., Inkpen, K., and Nushi, B. 2022. "Who Goes First? Influences of Human-AI Workflow on Decision Making in Clinical Imaging," in *2022 ACM Conference on Fairness, Accountability, and Transparency*, Seoul Republic of Korea: ACM, June 21, pp. 1362–1374. (https://doi.org/10.1145/3531146.3533193).

Fornell, C., and Larcker, D. F. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), SAGE Publications Inc, pp. 39–50. (https://doi.org/10.1177/002224378101800104).

Fox-Wolfgramm, S. J. 1997. "Towards Developing a Methodology for Doing Qualitative Research: The Dynamic-Comparative Case Study Method," *Scandinavian Journal of Management* (13:4), Scandinavian Journal of Management, Elsevier Ltd, pp. 439–455. (https://doi.org/10.1016/S0956-5221(97)00028-6).

Freitas, A. A. 2014a. "Comprehensible Classification Models: A Position Paper," *SIGKDD Explor. Newsl.* (15:1), pp. 1–10. (https://doi.org/10.1145/2594473.2594475).

Freitas, A. A. 2014b. "Comprehensible Classification Models: A Position Paper," *ACM SIGKDD Explorations Newsletter* (15:1), pp. 1–10. (https://doi.org/10.1145/2594473.2594475).

Gambino, A., Fox, J., and Ratan, R. A. 2020. "Building a Stronger CASA: Extending the Computers Are Social Actors Paradigm," *Human-Machine Communication* (1), Communication and Social Robotics Labs, pp. 71–85. (https://doi.org/10.3316/INFORMIT.097034846749023).

Garidis, K., Ulbricht, L., Rossmann, A., and Schmäh, M. 2020. "Toward a User Acceptance Model of Autonomous Driving," *Hawaii International Conference on System Sciences 2020 (HICSS-53)*. (https://aisel.aisnet.org/hicss-53/da/iot/2).

Garreau, D., and Luxburg, U. 2020. "Explaining the Explainer: A First Theoretical Analysis of LIME," in *International Conference on Artificial Intelligence and Statistics*, PMLR, June 3, pp. 1287–1296. (http://proceedings.mlr.press/v108/garreau20a.html).

Gefen, D., and Straub, D. 2005. "A Practical Guide To Factorial Validity Using PLS-Graph: Tutorial And Annotated Example," *Communications of the Association for Information Systems* (16), Atlanta: Association for Information Systems, pp. 5-. (https://doi.org/10.17705/1CAIS.01605).

Gefen, D., Straub, D., and Boudreau, M.-C. 2000. “Structural Equation Modeling and Regression: Guidelines for Research Practice,” *Communications of the Association for Information Systems* (4:1). (https://doi.org/10.17705/1CAIS.00407).

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., and Kagal, L. 2018. “Explaining Explanations: An Overview of Interpretability of Machine Learning,” *arXiv:1806.00069 [Cs, Stat]*. (http://arxiv.org/abs/1806.00069).

Goodman, B., and Flaxman, S. 2017. “European Union Regulations on Algorithmic Decision-Making and a ‘Right to Explanation,’” *AI Magazine* (38:3), pp. 50–57. (https://doi.org/10.1609/aimag.v38i3.2741).

Green, B., and Chen, Y. 2019. “The Principles and Limits of Algorithm-in-the-Loop Decision Making,” *Proceedings of the ACM on Human-Computer Interaction* (3:CSCW), 50:1-50:24. (https://doi.org/10.1145/3359152).

Gregor, S. 2006. “The Nature of Theory in Information Systems,” *MIS Quarterly* (30:3), pp. 611–642. (https://doi.org/10.2307/25148742).

Gregor, S., and Benbasat, I. 1999. “Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice,” *MIS Quarterly* (23:4), pp. 497–530. (https://doi.org/10.2307/249487).

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., and Nelson, C. 2000. “Clinical versus Mechanical Prediction: A Meta-Analysis,” *Psychological Assessment* (12:1), US: American Psychological Association, pp. 19–30. (https://doi.org/10.1037/1040-3590.12.1.19).

Guba, E. G., and Lincoln, Y. S. 1982. *Effective Evaluation*, (1st ed.), Jossey-Bass Higher Education Series, San Francisco: Jossey-Bass Publishers.

Gugerty, L. 2006. “Newell and Simon’s Logic Theorist: Historical Background and Impact on Cognitive Modeling,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (50:9), pp. 880–884. (https://doi.org/10.1177/154193120605000904).

Gunkel, D. 2012. “Communication and Artificial Intelligence: Opportunities and Challenges for the 21st Century,” *Communication 1* (1:1), pp. 1–25. (https://doi.org/10.7275/R5QJ7F7R).

Gursoy, D., Chi, O. H., Lu, L., and Nunkoo, R. 2019. “Consumers Acceptance of Artificially Intelligent (AI) Device Use in Service Delivery,” *International Journal of Information Management* (49), pp. 157–169. (https://doi.org/10.1016/j.ijinfomgt.2019.03.008).

Hacker, P., Engel, A., and Mauer, M. 2023. “Regulating ChatGPT and Other Large Generative AI Models,” in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, FAccT ’23, New York, NY, USA: Association for Computing Machinery, June 12, pp. 1112–1123. (https://doi.org/10.1145/3593013.3594067).

Hackman, Jr., and Oldham, G. R. 1976. “Motivation through the Design of Work: Test of a Theory,” *Organizational Behavior and Human Performance* (16:2), Elsevier B.V, pp. 250–279. (https://doi.org/10.1016/0030-5073(76)90016-7).

Hair, J. F. 1995. *Multivariate Data Analysis with Readings*, (4th ed.), Englewood Cliffs, N.J: Prentice Hall.

Hair, J. F., Babin, B. J., and Anderson, R. E. 2019. *Multivariate Data Analysis*, (8th ed.), Cengage.

Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. 2019. "When to Use and How to Report the Results of PLS-SEM," *European Business Review* (31:1), Emerald Publishing Limited, pp. 2–24. (https://doi.org/10.1108/EBR-11-2018-0203).

Han, E., Yin, D., and Zhang, H. 2022. "Bots with Feelings: Should AI Agents Express Positive Emotion in Customer Service?," *Information Systems Research*, INFORMS. (https://doi.org/10.1287/isre.2022.1179).

Handelman, G. S., Kok, H. K., Chandra, R. V., Razavi, A. H., Huang, S., Brooks, M., Lee, M. J., and Asadi, H. 2018. "Peering Into the Black Box of Artificial Intelligence: Evaluation Metrics of Machine Learning Methods," *American Journal of Roentgenology* (212:1), pp. 38–43. (https://doi.org/10.2214/AJR.18.20224).

Hari, H., Iyer, R., and Sampat, B. 2021. "Customer Brand Engagement through Chatbots on Bank Websites– Examining the Antecedents and Consequences," *International Journal of Human–Computer Interaction* (0:0), Taylor & Francis, pp. 1–16. (https://doi.org/10.1080/10447318.2021.1988487).

Hein, D., Rauschnabel, P., He, J., Richter, L., and Ivens, B. 2018. "What Drives the Adoption of Autonomous Cars?," *ICIS 2018 Proceedings*. (https://aisel.aisnet.org/icis2018/practice/Presentations/4).

Hendler, J. 2008. "Avoiding Another AI Winter," *IEEE Intelligent Systems* (23:2), pp. 2–4. (https://doi.org/10.1109/MIS.2008.20).

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Hair, J. F., Hult, G. T. M., and Calantone, R. J. 2014. "Common Beliefs and Reality About PLS: Comments on Rönkkö and Evermann (2013)," *Organizational Research Methods* (17:2), pp. 182–209. (https://doi.org/10.1177/1094428114526928).

Henseler, J., and Sarstedt, M. 2013. "Goodness-of-Fit Indices for Partial Least Squares Path Modeling," *Computational Statistics* (28:2), pp. 565–580. (https://doi.org/10.1007/s00180-012-0317-1).

Herath, T., and Rao, H. R. 2009. "Encouraging Information Security Behaviors in Organizations: Role of Penalties, Pressures and Perceived Effectiveness," *Decision Support Systems* (47:2), pp. 154–165. (https://doi.org/10.1016/j.dss.2009.02.005).

Herlocker, J. L., Konstan, J. A., and Riedl, J. 2000. "Explaining Collaborative Filtering Recommendations," in *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, CSCW '00, New York, NY, USA: Association for Computing Machinery, December 1, pp. 241–250. (https://doi.org/10.1145/358916.358995).

Hill, J., Randolph Ford, W., and Farreras, I. G. 2015. "Real Conversations with Artificial Intelligence: A Comparison between Human–Human Online Conversations and Human–Chatbot

Conversations," *Computers in Human Behavior* (49), pp. 245–250. (https://doi.org/10.1016/j.chb.2015.02.026).

Hinton, G. E. 1991. *Connectionist Symbol Processing*, (1st MIT Press ed.), Special Issues of Artificial Intelligence, an International Journal, Cambridge, Mass: MIT Press.

Hoffman, D. L., and Novak, T. P. 1996. "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," *Journal of Marketing* (60:3), SAGE Publications Inc, pp. 50–68. (https://doi.org/10.1177/002224299606000304).

Hong, W., Chan, F. K. Y., Thong, J. Y. L., Chasalow, L. C., and Dhillon, G. 2014. "A Framework and Guidelines for Context-Specific Theorizing in Information Systems Research," *Information Systems Research* (25:1), pp. 111–136. (https://doi.org/10.1287/isre.2013.0501).

Hornung, O., and Smolnik, S. 2022. "AI Invading the Workplace: Negative Emotions towards the Organizational Use of Personal Virtual Assistants," *Electronic Markets* (32:1), pp. 123–138. (https://doi.org/10.1007/s12525-021-00493-0).

Horwitz, S. K. 2005. "The Compositional Impact of Team Diversity on Performance: Theoretical Considerations," *Human Resource Development Review* (4:2), pp. 219–245. (https://doi.org/10.1177/1534484305275847).

Hu, L., and Bentler, P. M. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives," *Structural Equation Modeling: A Multidisciplinary Journal* (6:1), Routledge, pp. 1–55. (https://doi.org/10.1080/10705519909540118).

Hua Ye, J., and Kankanhalli, A. 2018. "User Service Innovation on Mobile Phone Platforms: Investigating Impacts of Lead Userness, Toolkit Support, and Design Autonomy," *Management Information Systems Quarterly* (42:1), pp. 165–187.

Hughes, A. 2023. "ChatGPT: Everything You Need to Know about OpenAI's GPT-4 Tool," *BBC Science Focus Magazine*. (https://www.sciencefocus.com/future-technology/gpt-3/, accessed June 6, 2023).

Hutson, M. 2023. "Can We Stop Runaway A.I.?," *The New Yorker*. (https://www.newyorker.com/science/annals-of-artificial-intelligence/can-we-stop-the-singularity).

Hyder, Z., Siau, K., and Nah, F. 2018. "Use of Artificial Intelligence, Machine Learning, and Autonomous Technologies in the Mining Industry," *MWAIS 2018 Proceedings*. (https://aisel.aisnet.org/mwais2018/43).

Hymas, C. 2019. "AI Used for First Time in Job Interviews in UK to Find Best Applicants," *The Telegraph*. (https://www.telegraph.co.uk/news/2019/09/27/ai-facial-recognition-used-first-time-job-interviews-uk-find/).

IBM. 2017. "2016 Corporate Responsibility Report," *2016 IBM Responsibility Report – Our Approach*, , July 13. (https://www.ibm.com/ibm/responsibility/2016/at-a-glance/our-approach.html, accessed October 13, 2022).

IBM Research Trusted AI. (n.d.). "AI Explainability 360 - Demo." (aix360.mybluemix.net/explanation_cust, accessed April 8, 2021).

Igbaria, M., Greenhaus, J. H., and Parasuraman, S. 1991. "Career Orientations of MIS Employees: An Empirical Analysis," *MIS Quarterly* (15:2), Minneapolis: The Society for Information Management and The Management Information Systems Research Center of the University of Minnesota, pp. 151–169. (https://doi.org/10.2307/249376).

Inkster, B., Sarda, S., and Subramanian, V. 2018. "An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study," *JMIR mHealth and uHealth* (6:11), p. e12106. (https://doi.org/10.2196/12106).

Iyamu, T. 2018. "Collecting Qualitative Data for Information Systems Studies: The Reality in Practice," *Education and Information Technologies* (23:5), New York: Springer US, pp. 2249–2264. (https://doi.org/10.1007/s10639-018-9718-2).

Jabagi, N., Croteau, A.-M., Audebrand, L., and Marsan, J. 2021. "Unlocking Perceived Algorithmic Autonomy-Support: Scale Development and Validation," *Hawaii International Conference on System Sciences 2021 (HICSS-54)*. (https://aisel.aisnet.org/hicss-54/os/socio-technical_issues_in_it/9).

Jakobsen, M., and Jensen, R. 2015. "Common Method Bias in Public Management Studies," *International Public Management Journal* (18:1), Routledge, pp. 3–30. (https://doi.org/10.1080/10967494.2014.997906).

Johns, G. 2006. "The Essential Impact of Context on Organizational Behavior," *The Academy of Management Review* (31:2), pp. 386–408. (https://doi.org/10.2307/20159208).

Joshi, N. 2019. "7 Types Of Artificial Intelligence," *Forbes*. (https://www.forbes.com/sites/cognitiveworld/2019/06/19/7-types-of-artificial-intelligence/, accessed January 29, 2020).

Judson, T. J., Odisho, A. Y., Young, J. J., Bigazzi, O., Steuer, D., Gonzales, R., and Neinstein, A. B. 2020. "Implementation of a Digital Chatbot to Screen Health System Employees during the COVID-19 Pandemic," *Journal of the American Medical Informatics Association* (27:9), pp. 1450–1455. (https://doi.org/10.1093/jamia/ocaa130).

Kaartemo, V., and Helkkula, A. 2018. "A Systematic Review of Artificial Intelligence and Robots in Value Co-Creation: Current Status and Future Research Avenues," *Journal of Creating Value* (4:2), pp. 211–228. (https://doi.org/10.1177/2394964318805625).

Kakar, A. 2016. "Enhancing Reflexivity in Software Development Teams: Should We Focus on Autonomy or Interdependence?," *Journal of Information Technology Theory and Application (JITTA)* (17:3). (https://aisel.aisnet.org/jitta/vol17/iss3/2).

Kaplan, A., and Haenlein, M. 2019. "Siri, Siri, in My Hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence," *Business Horizons* (62:1), pp. 15–25. (https://doi.org/10.1016/j.bushor.2018.08.004).

Karasek, R. A. 1979. "Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign," *Administrative Science Quarterly* (24:2), Ithaca, N.Y: Cornell University Graduate School of Business and Public Administration, pp. 285–308. (https://doi.org/10.2307/2392498).

Kiew, J. 2018. "The Digital Surgery: Humber River Hospital Reinvents Itself with AI & Robotics," *The Change Leadership*, , June 16. (https://thechangeleadership.com/humber-river-hospital-reinvents-itself-with-ai-robotics/, accessed November 22, 2019).

Kim, B., Park, J., and Suh, J. 2020. "Transparency and Accountability in AI Decision Support: Explaining and Visualizing Convolutional Neural Networks for Text Information," *Decision Support Systems* (134), p. 113302. (https://doi.org/10.1016/j.dss.2020.113302).

Kleinberg, J., Ludwig, J., Mullainathan, S., and Sunstein, C. R. 2018. "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* (10), pp. 113–174. (https://doi.org/10.1093/jla/laz001).

Komiak, S. Y. X., Wang, W., and Benbasat, I. 2004. "Trust Building in Virtual Salespersons Versus in Human Salespersons: Similarities and Differences," *E-Service Journal* (3:3), Indiana University Press, pp. 49–64. (https://doi.org/10.2979/esj.2004.3.3.49).

Krafft, T. D., Zweig, K. A., and König, P. D. 2022. "How to Regulate Algorithmic Decision-Making: A Framework of Regulatory Requirements for Different Applications," *Regulation & Governance* (16:1), pp. 119–136. (https://doi.org/10.1111/rego.12369).

Krippendorff, K. 2018. *Content Analysis: An Introduction to Its Methodology*, SAGE Publications.

Kulesza, T., Burnett, M., Wong, W.-K., and Stumpf, S. 2015. "Principles of Explanatory Debugging to Personalize Interactive Machine Learning," in *Proceedings of the 20th International Conference on Intelligent User Interfaces*, IUI '15, New York, NY, USA: Association for Computing Machinery, March 18, pp. 126–137. (https://doi.org/10.1145/2678025.2701399).

Kutscher, B. 2016. "Inside North America's First All-Digital Hospital," *Modern Healthcare*. (https://www.modernhealthcare.com/article/20160430/MAGAZINE/304309981/inside-north-america-s-first-all-digital-hospital, accessed November 22, 2019).

Labay, D. G., and Kinnear, T. C. 1981. "Exploring the Consumer Decision Process in the Adoption of Solar Energy Systems," *Journal of Consumer Research* (8:3), Oxford: Journal of Consumer Research, pp. 271–278. (https://doi.org/10.1086/208865).

Laber, J., Thamma, R., and Kirby, E. D. 2020. "The Impact of Warehouse Automation in Amazon's Success," *International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 8, August 2020* (7:8).

Laguarta, J., Hueto, F., and Subirana, B. 2020. "COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings," *IEEE Open Journal of Engineering in Medicine and Biology* (1), pp. 275–281. (https://doi.org/10.1109/OJEMB.2020.3026928).

Lamb, R., and Kling, R. 2003. “Reconceptualizing Users as Social Actors in Information Systems Research,” *MIS Quarterly* (27:2), Management Information Systems Research Center, University of Minnesota, pp. 197–236. (https://doi.org/10.2307/30036529).

Las. 2018. “The TELUS Virtual Assistant: 24/7 Service for a 24/7 World,” , August 22. (https://forum.telus.com/t5/Neighbourhood-Blog/The-TELUS-Virtual-Assistant-24-7-Service-for-a-24-7-World/ba-p/88688, accessed November 22, 2019).

Lee, H., Jeong Cho, H., Xu, W., and Fairhurst, A. 2010. “The Influence of Consumer Traits and Demographics on Intention to Use Retail Self-service Checkouts,” *Marketing Intelligence & Planning* (28:1), Emerald Group Publishing Limited, pp. 46–58. (https://doi.org/10.1108/02634501011014606).

Lee, H.-J. 2017. “Personality Determinants of Need for Interaction with a Retail Employee and Its Impact on Self-Service Technology (SST) Usage Intentions,” *Journal of Research in Interactive Marketing* (11:3), Bradford, United Kingdom: Emerald Group Publishing Limited, pp. 214–231. (https://doi.org/10.1108/JRIM-04-2016-0036).

Lee, H.-J., and Lyu, J. 2019. “Exploring Factors Which Motivate Older Consumers’ Self-Service Technologies (SSTs) Adoption,” *The International Review of Retail, Distribution and Consumer Research* (29:2), London: Routledge, pp. 218–239. (https://doi.org/10.1080/09593969.2019.1575261).

Lewin, K. 1943. *Bulletin of the National Research Council*, National Research Council of the National Academy of Sciences.

Libert, K., Mosconi, E., and Cadieux, N. 2020. *Human-Machine Interaction and Human Resource Management Perspective for Collaborative Robotics Implementation and Adoption*. (http://hdl.handle.net/10125/63805).

Ligaya, A. 2018. *TD Acquires Toronto-Based AI Startup Layer 6*. (https://www.theglobeandmail.com/report-on-business/td-acquires-toronto-based-ai-startup-layer-6/article37537469/).

Lim, B. Y., Dey, A. K., and Avrahami, D. 2009. “Why and Why Not Explanations Improve the Intelligibility of Context-Aware Intelligent Systems,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’09, New York, NY, USA: Association for Computing Machinery, April 4, pp. 2119–2128. (https://doi.org/10.1145/1518701.1519023).

Lin, H.-F. 2011. “An Empirical Investigation of Mobile Banking Adoption: The Effect of Innovation Attributes and Knowledge-Based Trust,” *International Journal of Information Management* (31:3), pp. 252–260. (https://doi.org/10.1016/j.ijinfomgt.2010.07.006).

Lindell, M. K., and Whitney, D. J. 2001. “Accounting for Common Method Variance in Cross-Sectional Research Designs,” *Journal of Applied Psychology* (86), US: American Psychological Association, pp. 114–121. (https://doi.org/10.1037/0021-9010.86.1.114).

Liu, C.-F., Chen, Z.-C., Kuo, S.-C., and Lin, T.-C. 2022. "Does AI Explainability Affect Physicians' Intention to Use AI?," *International Journal of Medical Informatics (Shannon, Ireland)* (168), Elsevier B.V, pp. 104884–104884. (https://doi.org/10.1016/j.ijmedinf.2022.104884).

Liu, H.-Y., and Zawieska, K. 2020. "From Responsible Robotics towards a Human Rights Regime Oriented to the Challenges of Robotics and Artificial Intelligence," *Ethics and Information Technology* (22:4), Dordrecht: Springer Netherlands, pp. 321–333. (https://doi.org/10.1007/s10676-017-9443-3).

Lundberg, S., and Lee, S.-I. 2017. "A Unified Approach to Interpreting Model Predictions," *arXiv:1705.07874 [Cs, Stat]*. (http://arxiv.org/abs/1705.07874).

Luo, X., Tong, S., Fang, Z., and Qu, Z. 2019. "Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases," *Marketing Science* (38:6), INFORMS, pp. 937–947. (https://doi.org/10.1287/mksc.2019.1192).

Lysyakov, M., and Viswanathan, S. 2022. "Threatened by AI: Analyzing Users' Responses to the Introduction of AI in a Crowd-Sourcing Platform," *Information Systems Research*. (https://doi.org/10.1287/isre.2022.1184).

Mackeprang, M., Müller-Birn, C., and Stauss, M. T. 2019. "Discovering the Sweet Spot of Human-Computer Configurations: A Case Study in Information Extraction," *Proceedings of the ACM on Human-Computer Interaction* (3:CSCW), 195:1-195:30. (https://doi.org/10.1145/3359297).

Mackworth, N. H. 1950. "Researches on the Measurement of Human Performance.," *Researches on the Measurement of Human Performance.* (268), London H M Stationery Office. (https://www.cabdirect.org/cabdirect/abstract/19502703894).

Maitra, S. 2020. *Artificial Intelligence and Indigenous Perspectives: Protecting and Empowering Intelligent Human Beings*, AIES '20, ACM, pp. 320–326. (https://doi.org/10.1145/3375627.3375845).

Markus, A. F., Kors, J. A., and Rijnbeek, P. R. 2021. "The Role of Explainability in Creating Trustworthy Artificial Intelligence for Health Care: A Comprehensive Survey of the Terminology, Design Choices, and Evaluation Strategies," *Journal of Biomedical Informatics* (113), p. 103655. (https://doi.org/10.1016/j.jbi.2020.103655).

Maxwell, J. A. 2002. "Understanding and Validity in Qualitative Research," in *The Qualitative Researcher's Companion*, 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc., pp. 36–64. (https://doi.org/10.4135/9781412986274.n2).

Mayson, S. G. 2018. "Bias in, Bias Out," *Yale Law Journal* (128), p. 2218.

Mazmanian, M., Orlikowski, W. J., and Yates, J. 2013. "The Autonomy Paradox: The Implications of Mobile Email Devices for Knowledge Professionals," *Organization Science (Providence, R.I.)* (24:5), Linthicum: INFORMS, pp. 1337–1357. (https://doi.org/10.1287/orsc.1120.0806).

McKnight, D. H., Choudhury, V., and Kacmar, C. 2002. “Developing and Validating Trust Measures for E-Commerce: An Integrative Typology,” *Information Systems Research* (13:3), INFORMS, pp. 334–359. (https://doi.org/10.1287/isre.13.3.334.81).

Meerbeek, B., Hoonhout, J., Bingley, P., and Terken, J. M. B. 2008. “The Influence of Robot Personality on Perceived and Preferred Level of User Control,” *Interaction Studies* (9:2), John Benjamins, pp. 204–229. (https://doi.org/10.1075/is.9.2.04mee).

Merriam Webster. 2023. “Definition of ARTIFICIAL INTELLIGENCE,” , August 23. (https://www.merriam-webster.com/dictionary/artificial+intelligence, accessed August 31, 2023).

Meyers, L. S., Gamst, G., and Guarino, A. J. 2016. *Applied Multivariate Research: Design and Interpretation*, SAGE Publications.

Miller, B. K. 2021. “Impact of Social Desirability and Common Method Variance on Two Measures of Entitlement,” *Psychological Reports* (124:4), SAGE Publications Inc, pp. 1845–1862. (https://doi.org/10.1177/0033294120937439).

Miller, T. 2019a. “Explanation in Artificial Intelligence: Insights from the Social Sciences,” *Artificial Intelligence* (267), Amsterdam: Elsevier B.V, pp. 1–38. (https://doi.org/10.1016/j.artint.2018.07.007).

Miller, T. 2019b. “Explanation in Artificial Intelligence: Insights from the Social Sciences,” *Artificial Intelligence* (267), pp. 1–38. (https://doi.org/10.1016/j.artint.2018.07.007).

Min, J.-A. 2016. “How Automated Candidate Screening Solves The ‘Ignore’ Problem,” *Ideal*, , December 7. (https://ideal.com/candidate-screening/, accessed January 26, 2020).

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. 2015. “Human-Level Control through Deep Reinforcement Learning,” *Nature* (518:7540), Nature Publishing Group, pp. 529–533. (https://doi.org/10.1038/nature14236).

Modern Hire. 2018. “AI Could Reduce Hiring Bias, Recruiters & Candidates Agree,” *Modern Hire*. (https://modernhire.com/ai-in-recruiting-could-reduce-bias/, accessed January 26, 2020).

Mohseni, S., Zarei, N., and Ragan, E. D. 2021. “A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems,” *ACM Transactions on Interactive Intelligent Systems* (11:3–4), 24:1-24:45. (https://doi.org/10.1145/3387166).

Molnar, C. 2019. *Interpretable Machine Learning*, Lulu.com.

Moore, J. E. 2000. “One Road to Turnover: An Examination of Work Exhaustion in Technology Professionals,” *MIS Quarterly* (24:1), Minneapolis: The Society for Information Management and The Management Information Systems Research Center of the University of Minnesota, and The Association for Information Systems, pp. 141–168. (https://doi.org/10.2307/3250982).

Motulsky, S. L. 2021. “Is Member Checking the Gold Standard of Quality in Qualitative Research?,” *Qualitative Psychology* (8:3), Educational Publishing Foundation, pp. 389–406. (https://doi.org/10.1037/qup0000215).

Mou, Y., and Xu, K. 2017. “The Media Inequality: Comparing the Initial Human-Human and Human-AI Social Interactions,” *Computers in Human Behavior* (72), pp. 432–440. (https://doi.org/10.1016/j.chb.2017.02.067).

Muggleton, S. H., Muggleton, S., Chater, Nicholas, and Chater, Nick. 2021. *Human-Like Machine Intelligence*, Oxford University Press.

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., and Yu, B. 2019. “Definitions, Methods, and Applications in Interpretable Machine Learning,” *Proceedings of the National Academy of Sciences* (116:44), pp. 22071–22080. (https://doi.org/10.1073/pnas.1900654116).

Murray, D. A., Rhymer, M. J., and Sirmon, D. D. G. 2020. “HUMANS AND TECHNOLOGY: FORMS OF CONJOINED AGENCY IN ORGANIZATIONS,” *Https://Doi.Org/10.5465/Amr.2019.0186*, , March 3. (https://doi.org/10.5465/amr.2019.0186).

Myers, M. D. 2013. *Qualitative Research in Business and Management*, SAGE.

Myers, M. D. 2019. *Qualitative Research in Business and Management*, SAGE Publications Ltd, pp. 1–364.

Namgung, M., and Akar, G. 2014. “Role of Gender and Attitudes on Public Transportation Use,” *Transportation Research Record* (2415:1), SAGE Publications Inc, pp. 136–144. (https://doi.org/10.3141/2415-15).

Nass, C., Fogg, B. J., and Moon, Y. 1996. “Can Computers Be Teammates?,” *International Journal of Human-Computer Studies* (45:6), pp. 669–678. (https://doi.org/10.1006/ijhc.1996.0073).

Nass, C., and Moon, Y. 2000. “Machines and Mindlessness: Social Responses to Computers,” *Journal of Social Issues* (56:1), pp. 81–103. (https://doi.org/10.1111/0022-4537.00153).

Nass, C., Steuer, J., and Tauber, E. R. 1994. “Computers Are Social Actors,” in *Conference Companion on Human Factors in Computing Systems*, CHI ’94, New York, NY, USA: Association for Computing Machinery, April 28, p. 204. (https://doi.org/10.1145/259963.260288).

Negnevitsky, M. 2005. *Artificial Intelligence: A Guide to Intelligent Systems*, (2nd ed.), Harlow, England ; New York: Addison-Wesley.

Newman, L., Haran, U., and Fink, L. 2022. *Let Me Decide: The Importance of User Autonomy in Accepting Online Recommendations*.

Nguyen, T. 2019. *Potential Effects of Chatbot Technology on Customer Support: A Case Study*. (https://aaltodoc.aalto.fi:443/handle/123456789/38921).

Nickerson, J. V., and Reilly, R. R. 2004. “A Model for Investigating the Effects of Machine Autonomy on Human Behavior,” in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of The*, , January, 10 pp.-. (https://doi.org/10.1109/HICSS.2004.1265325).

Norman, D. 2017. “Design, Business Models, and Human-Technology Teamwork: As Automation and Artificial Intelligence Technologies Develop, We Need to Think Less about Human-Machine Interfaces and More about Human-Machine Teamwork,” *Research-Technology Management* (60:1), pp. 26–30. (https://doi.org/10.1080/08956308.2017.1255051).

Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., and Patton, M. W. 2011. “Embodied Conversational Agent-Based Kiosk for Automated Interviewing,” *Journal of Management Information Systems* (28:1), Abingdon: Routledge, pp. 17–48. (https://doi.org/10.2753/MIS0742-1222280102).

Nunnally, J. C. 1978. “An Overview of Psychological Measurement,” in *Clinical Diagnosis of Mental Disorders: A Handbook*, B. B. Wolman (ed.), Boston, MA: Springer US, pp. 97–146. (https://doi.org/10.1007/978-1-4684-2490-4_4).

Ochmann, J., and Laumer, S. 2019. “Fairness as a Determinant of AI Adoption in Recruiting: An Interview-Based Study,” *DIGIT 2019 Proceedings*. (https://aisel.aisnet.org/digit2019/16).

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., and Pollock, A. 2009. “The Relative Influence of Advice from Human Experts and Statistical Methods on Forecast Adjustments,” *Journal of Behavioral Decision Making* (22:4), pp. 390–409. (https://doi.org/10.1002/bdm.637).

Orlikowski, W. J., and Baroudi, J. J. 1991. “Studying Information Technology in Organizations: Research Approaches and Assumptions,” *Information Systems Research* (2:1), pp. 1–28. (https://doi.org/10.1287/isre.2.1.1).

Orman, L. V. 2002. “Electronic Markets, Hierarchies, Hubs, and Intermediaries,” *Information Systems Frontiers* (4:2), pp. 213–228. (https://doi.org/10.1023/A:1016007020846).

Ozturk, A. B., Bilgihan, A., Nusair, K., and Okumus, F. 2016. “What Keeps the Mobile Hotel Booking Users Loyal? Investigating the Roles of Self-Efficacy, Compatibility, Perceived Ease of Use, and Perceived Convenience,” *International Journal of Information Management* (36:6), Kidlington: Elsevier Ltd, pp. 1350–1359. (https://doi.org/10.1016/j.ijinfomgt.2016.04.005).

Papachristos, E., Skov Johansen, P., Møberg Jacobsen, R., Bjørn Leer Bysted, L., and Skov, M. B. 2021. “How Do People Perceive the Role of AI in Human-AI Collaboration to Solve Everyday Tasks?,” in *CHI Greece 2021: 1st International Conference of the ACM Greek SIGCHI Chapter*, CHI Greece 2021, New York, NY, USA: Association for Computing Machinery, November 25, pp. 1–6. (https://doi.org/10.1145/3489410.3489420).

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. 2000. “A Model for Types and Levels of Human Interaction with Automation,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* (30:3), pp. 286–297. (https://doi.org/10.1109/3468.844354).

Peng, A., Nushi, B., Kıcıman, E., Inkpen, K., Suri, S., and Kamar, E. 2019. "What You See Is What You Get? The Impact of Representation Criteria on Human Bias in Hiring," *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* (7), pp. 125–134.

Pentina, I., Zhang, L., Bata, H., and Chen, Y. 2016. "Exploring Privacy Paradox in Information-Sensitive Mobile App Adoption: A Cross-Cultural Comparison," *Computers in Human Behavior* (65), pp. 409–419. (https://doi.org/10.1016/j.chb.2016.09.005).

Peukert, C., Bechtold, S., Batikas, M., and Kretschmer, T. 2022. "Regulatory Spillovers and Data Governance: Evidence from the GDPR," *Marketing Science*, Mksc.2021.1339. (https://doi.org/10.1287/mksc.2021.1339).

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. 2003. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology* (88:5), pp. 879–903. (https://doi.org/10.1037/0021-9010.88.5.879).

Polli, F. 2019. "Using AI to Eliminate Bias from Hiring," *Harvard Business Review*. (https://hbr.org/2019/10/using-ai-to-eliminate-bias-from-hiring).

Posard, M. N., and Gordon Rinderknecht, R. 2015. "Do People like Working with Computers More than Human Beings?," *Computers in Human Behavior* (51), pp. 232–238. (https://doi.org/10.1016/j.chb.2015.04.057).

Press, G. 2019. "AI Stats News: Chatbots Increase Sales By 67% But 87% Of Consumers Prefer Humans," *Forbes*. (https://www.forbes.com/sites/gilpress/2019/11/25/ai-stats-news-chatbots-increase-sales-by-67-but-87-of-consumers-prefer-humans/, accessed April 6, 2023).

PricewaterhouseCoopers. 2017. "PwC's Global Artificial Intelligence Study: Sizing the Prize," *PwC*. (https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html, accessed August 13, 2018).

Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P., and Mazurek, G. 2019. "In Bot We Trust: A New Methodology of Chatbot Performance Measures," *Business Horizons* (62:6), Digital Transformation & Disruption, pp. 785–797. (https://doi.org/10.1016/j.bushor.2019.08.005).

Rader, E., Cotter, K., and Cho, J. 2018. "Explanations as Mechanisms for Supporting Algorithmic Transparency," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA: Association for Computing Machinery, April 19, pp. 1–13. (https://doi.org/10.1145/3173574.3173677).

Rai, A. 2020. "Explainable AI: From Black Box to Glass Box," *Journal of the Academy of Marketing Science* (48:1), pp. 137–141. (https://doi.org/10.1007/s11747-019-00710-5).

Rai, A., Constantinides, P., and Sarker, S. 2019. "Editor's Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids," *Management Information Systems Quarterly* (43:1), pp. iii–ix.

Rajaobelina, L., Prom Tep, S., Arcand, M., and Ricard, L. 2021. "Creepiness: Its Antecedents and Impact on Loyalty When Interacting with a Chatbot," *Psychology & Marketing* (38:12), pp. 2339–2356. (https://doi.org/10.1002/mar.21548).

Ray, T. 2023. "ChatGPT Is 'not Particularly Innovative,' and 'Nothing Revolutionary', Says Meta's Chief AI Scientist," *ZDNET*. (https://www.zdnet.com/article/chatgpt-is-not-particularly-innovative-and-nothing-revolutionary-says-metas-chief-ai-scientist/, accessed June 20, 2023).

Reeves, B., and Nass, C. 1997. "The Media Equation: How People Treat Computers, Television, ? New Media Like Real People ? Places," *Computers and Mathematics with Applications* (5:33), p. 128.

Reinders, M. J., Dabholkar, P. A., and Frambach, R. T. 2008. "Consequences of Forcing Consumers to Use Technology-Based Self-Service," *Journal of Service Research* (11:2), SAGE Publications Inc, pp. 107–123. (https://doi.org/10.1177/1094670508324297).

Richardson, H. A., Simmering, M. J., and Sturman, M. C. 2009. "A Tale of Three Perspectives: Examining Post Hoc Statistical Techniques for Detection and Correction of Common Method Variance," *Organizational Research Methods* (12:4), SAGE Publications Inc, pp. 762–800. (https://doi.org/10.1177/1094428109332834).

Rock, D., and Grant, H. 2016. *Why Diverse Teams Are Smarter*, p. 4.

Rodríguez Cardona, D., Werth, O., Schönborn, S., and Breitner, M. H. 2019. "A Mixed Methods Analysis of the Adoption and Diffusion of Chatbot Technology in the German Insurance Sector," *AMCIS 2019 Proceedings*. (https://aisel.aisnet.org/amcis2019/adoption_diffusion_IT/adoption_diffusion_IT/18).

Rogers, E. M. 2003. *Diffusion of Innovations*, (5th ed.), New York ; Free Press.

Roldán, J. L., and Sánchez-Franco, M. J. 2012. "Variance-Based Structural Equation Modeling: Guidelines for Using Partial Least Squares in Information Systems Research," in *Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems*, IGI Global, pp. 193–221. (https://doi.org/10.4018/978-1-4666-0179-6.ch010).

Rosas, J., and Camarinha-Matos, L. M. 2010. "Assessment of the Willingness to Collaborate in Enterprise Networks," in *Emerging Trends in Technological Innovation*, IFIP Advances in Information and Communication Technology, L. M. Camarinha-Matos, P. Pereira, and L. Ribeiro (eds.), Springer Berlin Heidelberg, pp. 14–23.

Rosenblatt, F. 1959. *Perceptron Simulation Experiments (Project Para)*, Buffalo, Cornell Aeronautical Laboratory, Inc.,.

Rosnow, R. L., and Rosenthal, R. 2003. "Effect Sizes for Experimenting Psychologists," *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale* (57:3), Canada: Canadian Psychological Association, pp. 221–237. (https://doi.org/10.1037/h0087427).

Rossi, F. 2019. "BUILDING TRUST IN ARTIFICIAL INTELLIGENCE," *Journal of International Affairs* (72:1), pp. 127–134.

Rousseau, D. M., and Fried, Y. 2001. "Location, Location, Location: Contextualizing Organizational Research*," *Journal of Organizational Behavior* (22:1), pp. 1–13. (https://doi.org/10.1002/job.78).

Rudin, C. 2019. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead," *arXiv:1811.10154 [Cs, Stat]*. (http://arxiv.org/abs/1811.10154).

Rzepka, C., Berger, B., and Hess, T. 2020. *Why Another Customer Channel? Consumers' Perceived Benefits and Costs of Voice Commerce*, , January 7. (https://doi.org/10.24251/HICSS.2020.499).

SAE International. 2014. "SAE International Technical Standard Provides Terminology for Motor Vehicle Automated Driving Systems," *India Automobile News*, New Delhi: Athena Information Solutions Pvt. Ltd. (https://search.proquest.com/docview/1566879442?pq-origsite=primo).

Sandelowski, M. 1986. "The Problem of Rigor in Qualitative Research," *Advances in Nursing Science* (8:3), United States: Ovid Technologies Wolters Kluwer Health, pp. 27–37. (https://doi.org/10.1097/00012272-198604000-00005).

Sarker, S., Xiao, X., and Beaulieu, T. 2013. "Guest Editorial: Qualitative Studies in Information Systems: A Critical Review and Some Guiding Principles," *MIS Quarterly* (37:4), Management Information Systems Research Center, University of Minnesota, pp. iii–xviii.

Sarker, S., Xiao, X., Copenhagen Business School, Beaulieu, T., Utah State University, Lee, A. S., and Virginia Commonwealth University. 2018. "Learning from First-Generation Qualitative Approaches in the IS Discipline: An Evolutionary View and Some Implications for Authors and Evaluators (PART 2/2)," *Journal of the Association for Information Systems* (19), pp. 909–923. (https://doi.org/10.17705/1jais.00512).

Schuetz, S., and Venkatesh, V. 2020. "Research Perspectives: The Rise of Human Machines: How Cognitive Computing Systems Challenge Assumptions of User-System Interaction ," *Journal of the Association for Information Systems* (21:2), Atlanta: Association for Information Systems, pp. 460–482. (https://doi.org/10.17705/1jais.00608).

Seeber, I., Bittner, E., Briggs, R. O., Vreede, G.-J. de, Vreede, T. de, Druckenmiller, D., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., and Söllner, M. 2018. "Machines as Teammates: A Collaboration Research Agenda," *Hawaii International Conference on System Sciences 2018 (HICSS-51)*. (https://aisel.aisnet.org/hicss-51/cl/processes_and_technologies_for_team/5).

Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., and Medow, M. A. 2013. "Why Do Patients Derogate Physicians Who Use a Computer-Based Diagnostic Support System?," *Medical Decision Making* (33:1), Los Angeles, CA: SAGE Publications, pp. 108–118. (https://doi.org/10.1177/0272989X12453501).

Shin, D. 2021a. "The Effects of Explainability and Causability on Perception, Trust, and Acceptance: Implications for Explainable AI," *International Journal of Human-Computer Studies* (146), p. 102551. (https://doi.org/10.1016/j.ijhcs.2020.102551).

Shin, D. 2021b. "The Effects of Explainability and Causability on Perception, Trust, and Acceptance: Implications for Explainable AI," *International Journal of Human-Computer Studies* (146), p. 102551. (https://doi.org/10.1016/j.ijhcs.2020.102551).

Simmering, M. J., Fuller, C. M., Richardson, H. A., Ocal, Y., and Atinc, G. M. 2015. "Marker Variable Choice, Reporting, and Interpretation in the Detection of Common Method Variance: A Review and Demonstration," *Organizational Research Methods* (18:3), SAGE Publications Inc, pp. 473–511. (https://doi.org/10.1177/1094428114560023).

Skinner, E. A. 1996. "A Guide to Constructs of Control," *Journal of Personality and Social Psychology* (71), US: American Psychological Association, pp. 549–570. (https://doi.org/10.1037/0022-3514.71.3.549).

Smithers, T. 1997. "Autonomy in Robots and Other Agents," *Brain and Cognition* (34:1), pp. 88–106. (https://doi.org/10.1006/brcg.1997.0908).

Sommer. 2017. "Artificial Intelligence, Machine Learning and Cognitive Computing," *IBM Nordic Blog*, , November 20. (https://www.ibm.com/blogs/nordic-msp/artificial-intelligence-machine-learning-cognitive-computing/, accessed October 14, 2022).

Sowa, K., Przegalinska, A., and Ciechanowski, L. 2021. "Cobots in Knowledge Work: Human – AI Collaboration in Managerial Professions," *Journal of Business Research* (125), pp. 135–142. (https://doi.org/10.1016/j.jbusres.2020.11.038).

Spector, P. E., Rosen, C. C., Richardson, H. A., Williams, L. J., and Johnson, R. E. 2019. "A New Perspective on Method Variance: A Measure-Centric Approach," *Journal of Management* (45:3), SAGE Publications Inc, pp. 855–880. (https://doi.org/10.1177/0149206316687295).

Stemler, S. 2001. "An Overview of Content Analysis," *Practical Assessment, Research, and Evaluation* (7:1). (https://doi.org/10.7275/z6fm-2e34).

Sternberg, H. S., Chen, H., Hofmann, E., and Prockl, G. 2020. "Autonomous Trucks: A Supply Chain Adoption Perspective," *Hawaii International Conference on System Sciences 2020 (HICSS-53)*. (https://aisel.aisnet.org/hicss-53/in/digital_supply_chain/5).

Stevens, J. 2002. *Applied Multivariate Statistics for the Social Sciences*, (4th ed.), Mahwah, N.J: Lawrence Erlbaum Associates.

Straub, D., Boudreau, M.-C., and Gefen, D. 2004. "Validation Guidelines for IS Positivist Research," *Communications of the Association for Information Systems* (13:1). (https://doi.org/10.17705/1CAIS.01324).

Symeonidis, P., Nanopoulos, A., and Manolopoulos, Y. 2009. "MoviExplain: A Recommender System with Explanations," in *Proceedings of the Third ACM Conference on Recommender Systems - RecSys '09*, New York, New York, USA: ACM Press, p. 317. (https://doi.org/10.1145/1639714.1639777).

Szulanski, G. 1996. "Exploring Internal Stickiness: Impediments to the Transfer of Best Practice within the Firm: Exploring Internal Stickiness," *Strategic Management Journal* (17:S2), pp. 27–43. (https://doi.org/10.1002/smj.4250171105).

Tamagnini, P., Krause, J., Dasgupta, A., and Bertini, E. 2017. "Interpreting Black-Box Classifiers Using Instance-Level Visual Explanations," in *Proceedings of the 2nd Workshop on Human-In-the-Loop Data Analytics - HILDA'17*, Chicago, IL, USA: ACM Press, pp. 1–6. (https://doi.org/10.1145/3077257.3077260).

Tehseen, S., Ramayah, T., and Sajilan, S. 2017. “Testing and Controlling for Common Method Variance: A Review of Available Methods,” *Journal of Management Sciences* (4:2), pp. 142–168. (https://doi.org/10.20547/jms.2014.1704202).

van Teijlingen, E., and Hundley, V. 2002. “The Importance of Pilot Studies,” *Nursing Standard* (16:40), England: BMJ Publishing Group LTD, pp. 33–36. (https://doi.org/10.7748/ns2002.06.16.40.33.c3214).

Tenenhaus, M., Amato, S., and Esposito Vinzi, V. 2004. “A Global Goodness-of-Fit Index for PLS Structural Equation Modelling,” *Proceedings of the XLII SIS Scientific Meeting*, pp. 739–742.

Teubner, T., Flath, C. M., Weinhardt, C., Van Der Aalst, W., and Hinz, O. 2023. “Welcome to the Era of ChatGPT et al.: The Prospects of Large Language Models,” *Business & Information Systems Engineering* (65:2), pp. 95–101. (https://doi.org/10.1007/s12599-023-00795-x).

Trevelyan, R. 2001. “The Paradox of Autonomy: A Case of Academic Research Scientists,” *Human Relations (New York)* (54:4), Thousand Oaks, CA: SAGE Publications, pp. 495–525. (https://doi.org/10.1177/0018726701544005).

Turban, E., and Watkins, P. R. 1986. “Integrating Expert Systems and Decision Support Systems,” *MIS Quarterly* (10:2), Minneapolis: The Society for Information Management and The Management Information Systems Research Center of the University of Minnesota, pp. 121–136. (https://doi.org/10.2307/249031).

Urbach, N., and Ahlemann, F. 2010. “Structural Equation Modeling in Information Systems Research Using Partial Least Squares,” *Journal of Information Technology Theory and Application (JITTA)* (11:2). (https://aisel.aisnet.org/jitta/vol11/iss2/2).

Vassilakopoulou, P., Haug, A., Salvesen, L. M., and Pappas, I. O. 2023. “Developing Human/AI Interactions for Chat-Based Customer Services: Lessons Learned from the Norwegian Government,” *European Journal of Information Systems* (32:1), pp. 10–22. (https://doi.org/10.1080/0960085X.2022.2096490).

Venkatesh, V., Brown, S. A., and Bala, H. 2013. “Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems,” *MIS Quarterly* (37:1), pp. 21–54.

Venkatesh, V., Brown, S. A., and Sullivan, Y. W. 2016. “Guidelines for Conducting Mixed-Methods Research: An Extension and Illustration,” *J. AIS* (17), p. 2. (https://doi.org/10.17705/1jais.00433).

Visani, G., Bagli, E., Chesani, F., Poluzzi, A., and Capuzzo, D. 2020. “Statistical Stability Indices for LIME: Obtaining Reliable Explanations for Machine Learning Models,” *arXiv:2001.11757 [Cs, Stat]*. (http://arxiv.org/abs/2001.11757).

Wakefield, K. 2023. “A Guide to the Types of Machine Learning Algorithms,” *SAS UK*. (https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html, accessed August 31, 2023).

Walsham, G. 1997. "Actor-Network Theory and IS Research: Current Status and Future Prospects," in *Information Systems and Qualitative Research*, A. S. Lee, J. Liebenau, and J. I. DeGross (eds.), Boston, MA: Springer US, pp. 466–480. (https://doi.org/10.1007/978-0-387-35309-8_23).

Wang, D., Churchill, E., Maes, P., Fan, X., Shneiderman, B., Shi, Y., and Wang, Q. 2020. "From Human-Human Collaboration to Human-AI Collaboration: Designing AI Systems That Can Work Together with People," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu HI USA: ACM, April 25, pp. 1–6. (https://doi.org/10.1145/3334480.3381069).

Wang, D., Yang, Q., Abdul, A., and Lim, B. Y. 2019. "Designing Theory-Driven User-Centric Explainable AI," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA: Association for Computing Machinery, May 2, pp. 1–15. (https://doi.org/10.1145/3290605.3300831).

Wang, L., Huang, N., Hong, Y., Liu, L., Guo, X., and Chen, G. 2020. "Effects of Voice-Based AI in Customer Service: Evidence from a Natural Experiment," *ICIS 2020 Proceedings*. (https://aisel.aisnet.org/icis2020/hci_artintel/hci_artintel/21).

Wang, W., and Benbasat, I. 2007. "Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs," *Journal of Management Information Systems* (23:4), Taylor & Francis, Ltd., pp. 217–246.

Wang, W., and Benbasat, I. 2008. "Attributions of Trust in Decision Support Technologies: A Study of Recommendation Agents for E-Commerce," *Journal of Management Information Systems* (24:4), Taylor & Francis, Ltd., pp. 249–273.

Weber, R. 2003. "Editor's Comments," *MIS Quarterly* (27:3), pp. iii–xii. (https://doi.org/10.2307/30036536).

Weber, R. P. 1990. *Basic Content Analysis*, SAGE.

Weber, S., Klesel, M., Oschinsky, F. M., and Niehaves, B. 2020. "How Autonomy Is Used in Information Systems Research: Status Quo and Prospective Opportunities," *Hawaii International Conference on System Sciences 2020 (HICSS-53)*. (https://aisel.aisnet.org/hicss-53/os/design/2).

Welfare, K., Hallowell, M., Shah, J., and Riek, L. 2019. *Consider the Human Work Experience When Integrating Robotics in the Workplace*, HRI '19, IEEE Press, pp. 75–84. (https://doi.org/10.1109/HRI.2019.8673139).

Wertenbroch, K., Schrift, R., Alba, J. W., Barasch, A., Bhattacharjee, A., Giesler, M., Knobe, J., Lehmann, Matz, S., Nave, G., Parker, J. R., Puntoni, S., Zweber, Y., and Zheng, Y. Y. 2020. "Autonomy in Consumer Choice," *Marketing Letters* (31:4), New York: Springer US, pp. 429–439. (https://doi.org/10.1007/s11002-020-09521-z).

Wetzels, M., Odekerken-Schröder, G., and van Oppen, C. 2009. "Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration," *MIS Quarterly* (33:1), Management Information Systems Research Center, University of Minnesota, pp. 177–195. (https://doi.org/10.2307/20650284).

Whang, C., and Im, H. 2018. "Does Recommendation Matter for Trusting Beliefs and Trusting Intentions? Focused on Different Types of Recommender System and Sponsored Recommendation," *International Journal of Retail & Distribution Management* (46:10), Emerald Publishing Limited, pp. 944–958. (https://doi.org/10.1108/IJRDM-06-2017-0122).

Whetten, D. A. 2009. "An Examination of the Interface between Context and Theory Applied to the Study of Chinese Organizations," *Management and Organization Review* (5:1), pp. 29–56. (https://doi.org/10.1111/j.1740-8784.2008.00132.x).

Whetten, D. A., and Mackey, A. 2002. "A Social Actor Conception of Organizational Identity and Its Implications for the Study of Organizational Reputation," *Business & Society* (41:4), SAGE Publications Inc, pp. 393–414. (https://doi.org/10.1177/0007650302238775).

White, D. W., and Lean, E. 2008. "The Impact of Perceived Leader Integrity on Subordinates in a Work Team Environment," *Journal of Business Ethics* (81:4), pp. 765–778. (https://doi.org/10.1007/s10551-007-9546-6).

Whitman, M., and Woszczynski, A. 2003. "The Problem of Common Method Variance in IS Research," in *Handbook of Information Systems Research*, Panama: IGI Global, pp. 66–78. (https://doi.org/10.4018/978-1-59140-144-5.ch005).

Wiener, N. 1950. "The Human Use of Human Beings: Cybernetics and Society," *Eweb:611*. (https://repository.library.georgetown.edu/handle/10822/762048, accessed January 11, 2023).

Wiggers, K. 2018. "Plum Uses AI to Hire People 'That Never Would Have Been Discovered through a Traditional Hiring Process,'" *VentureBeat*, , June 13. (https://venturebeat.com/2018/06/13/plum-uses-ai-to-hire-people-that-never-would-have-been-discovered-through-a-traditional-hiring-process/, accessed November 22, 2019).

Wilkinson, D., Alkan, Ö., Liao, Q. V., Mattetti, M., Vejsbjerg, I., Knijnenburg, B. P., and Daly, E. 2021. "Why or Why Not? The Effect of Justification Styles on Chatbot Recommendations," *ACM Transactions on Information Systems* (39:4), 42:1-42:21. (https://doi.org/10.1145/3441715).

Williams, L. J., Hartman, N., and Cavazotte, F. 2010. "Method Variance and Marker Variables: A Review and Comprehensive CFA Marker Technique," *Organizational Research Methods* (13:3), SAGE Publications Inc, pp. 477–514. (https://doi.org/10.1177/1094428110366036).

Wilson, H. J., and Daugherty, P. R. 2018. "Collaborative Intelligence: Humans and AI Are Joining Forces," *Harvard Business Review* (July–August 2018). (https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces).

Wright, K. B. 2005. "Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services," *Journal of Computer-Mediated Communication* (10:3), p. JCMC1034. (https://doi.org/10.1111/j.1083-6101.2005.tb00259.x).

Wu, J.-H., and Wang, S.-C. 2005. "What Drives Mobile Commerce?: An Empirical Evaluation of the Revised Technology Acceptance Model," *Information & Management* (42:5), pp. 719–729. (https://doi.org/10.1016/j.im.2004.07.001).

Wu, L., Hitt, L., and Lou, B. 2020. "Data Analytics, Innovation, and Firm Productivity," *Management Science* (66:5), INFORMS, pp. 2017–2039. (https://doi.org/10.1287/mnsc.2018.3281).

Yang, L., Wang, H., and Deleris, L. A. 2021. "What Does It Mean to Explain? A User-Centered Study on AI Explainability," in *Artificial Intelligence in HCI*, Lecture Notes in Computer Science, H. Degen and S. Ntoa (eds.), Cham: Springer International Publishing, pp. 107–121. (https://doi.org/10.1007/978-3-030-77772-2_8).

Yin, S. 2019. "Where Chatbots Are Headed," *The Intercom Blog*, , November 18. (https://www.intercom.com/blog/the-state-of-chatbots/, accessed April 6, 2023).

Zhang, F., Pan, Z., Lu, Y., and Hu, Q. 2022. "The Dark Side of Artificial Autonomy: Formation and Mitigation of AI Technostress in Smart Voice Assistants (SVAs)," *PACIS 2022 Proceedings*. (https://aisel.aisnet.org/pacis2022/211).

Appendices

Appendix A: Description of Interviewees' Profiles

		Occupation/ Job Title	Gender	Age	Hiring Experience (in years)
Pilot Interviews	1.	Digital Transformation Director at Google	Man	41-50	10
	2.	Sales Analyst at Siemens	Man	20-30	1
	3.	Public Health officer	Woman	41-50	5
Main Interviews	4.	Senior Manager at Hatch	Man	61-70	21
	5.	Senior Manager at Hockey Canada	Man	31-40	5
	6.	Product Manager at PWC	Man	31-40	5
	7.	Senior Manager	Woman	41-50	20
	8.	Senior Manager at Deloitte	Man	31-40	7
	9.	Senior HR manager	Man	41-50	11
	10.	Contract inspector at City of Hamilton	Man	41-50	14
	11.	Director, Development and Operations	Woman	31-40	21
	12.	Technology Transformation Consultant at Deloitte	Man	31-40	5
	13.	Senior Manager at EY	Woman	31-40	15
	14.	Strategic Business Consultant at Deloitte	Man	31-40	5
	15.	Program Manager	Woman	20-30	1
	16.	Regional Director	Woman	31-40	5
	17.	Senior HR Manager	Woman	31-40	7
	18.	Senior HR Manager	Woman	20-30	2
	19.	Senior Manager at Accenture	Man	41-50	5
	20.	Project Manager	Woman	41-50	15
	21.	Director of Operations	Man	31-40	5
	22.	Manager	Man	31-40	4
	23.	Senior Manager	Woman	41-50	11
	24.	Director of HR	Woman	51-60	18
	25.	Senior Consultant at KPMG	Woman	20-30	2

Appendix B: Interview Sample Excerpts

Benefits:

Theme	Supporting Quotes
Efficiency	- *"Huge opportunity for initial screening with the basics"* - *"Frees up time of employees."* - *"It is really a time saver"* - *"Time is the biggest benefit. It takes us several hours to see who might work best"* - *"One benefit maybe is to exclude humans from screening candidates.. but the focus should be to enable humans to focus on tasks that AI is not good at."* - *" The good thing is that it can be used anytime, anywhere to target global candidates….and that it can be used on demand."* - *"One benefit is that there is no need to be there physically for every task which can save a lot of time."* - *"Ability to reach to more people at convenient times"* - *"Reaching out to a broader audience because we can go only to one school, so widening the pool allows us to have people who are international or live in farther cities."*

	- *"Processing a lot of applications faster…..screening hundreds of resumes is hard for a person to process."* - *"Over time and with practice, AI would be 100% more accurate than a human."* - *"AI would be more accurate when looking for certain keywords… we ourselves look for specific keywords when screening resumes."* - *"AI would be beneficial in avoiding many human errors and simplifying things."* - *"Many paperwork will be removed."* - *"You can lead it easily rather than a human who might be in a clash with what you are looking for."* - *"AI can even add more factors to the decision-making process that are beyond the capability of humans."* - *"AI can look at many things beyond the verbal language, such as energy levels, gestures, and body language."*
Integrity	- *"Taking out unconscious bias is a huge thing."* - *"Creates some objectivity for everybody when evaluating them as some hiring managers might be unconsciously biased."* - *"AI does not care about skin colour or ethnicity….it skips all the biases the human might have .. but the final decision should still be mine."* - *"AI would be less biased and more fair … because it will look for specific human qualities treating all applicants equally."* - *"Provides a sort of objectivity to inform decisions."*

	- *"Removes inherent bias so you do not miss good candidates."* - *"Can potentially eliminate bias for things such as accents, if the algorithm can capture word pronounce accurately and interpret it correctly."* - *"AI removes many biases and has an impartial perspective… because we do not know when we apply our own biases."* - *"It is a huge benefit that AI can provide an explanation of why they like or dislike certain candidates."*

Concerns:

Theme	Supporting Quotes
Lack of Human Interaction	- *"The feel that you get when talking to somebody … I mean a human."* - *"Loosing some of the human interaction is a concern here."* - *"The fact it is called "HR" means that we are getting away from the human element."* - *"I prefer having myself in the loop as these are people I am going to work with in the future, and it is important to know them personally."* - *"I am concerned about the human aspect.. for example, when I first came to Canada, in my first interview, I was very nervous because I did not know how to speak English well… but the interviewee could sense my potential and understood that it is ok to be nervous."* - *"I would prefer that it does the screening only and then take my opinion. I would still prefer the human element."*

	- *"How can the AI ask follow-up questions as humans do?..it is hard to do using AI."*
Incompatibility	- *"It is difficult to create a feature in AI to filter for culture fit or things like getting well with my team."* - *"If you are hiring a technician or a cleaner, for example, how will AI measure work ethics?"* - *"It does not take into account the behavioural profiles."* - *"Compatibility with the values and morals of the company would be a concern. It could be detected by asking AI certain questions to make sure it aligns with the company's values."* - *"how about the culture-fit and team-fit… how AI would assess someone who is appreciative or ungrateful?….someone who is great for one company might not be good in another company."* - *"Compatibility with values and ethics is, of course, a concern."*

Appendix C: Screening Questions

Role in the current company/organization: ☐ Upper Management ☐ Middle Management ☐ Junior Management ☐ Administrative Staff ☐ Support Staff ☐ Trained Professional ☐ Skilled Laborer ☐ Consultant ☐ Temporary Employee ☐ Researcher ☐ Student ☐ Self-employed/Partner ☐ Other: Please Specify ….
Department: ☐ Accounting ☐ Administration ☐ Customer Service ☐ Human Resources ☐ Inventory ☐ IT ☐ Logistics ☐ Manufacturing ☐ Marketing and Sales ☐ Procurement ☐ Quality Assurance ☐ Research & Development ☐ Other. Please specify ……

Industry:

- ☐ Agriculture, Forestry, Fishing and Hunting
- ☐ Utilities
- ☐ Computer and Electronics Manufacturing
- ☐ Wholesale
- ☐ Transportation and Warehousing
- ☐ Software
- ☐ Broadcasting
- ☐ Other Information Industry
- ☐ Real Estate, Rental and Leasing
- ☐ Primary/Secondary (K-12) Education
- ☐ Health Care and Social Assistance
- ☐ Hotel and Food Services
- ☐ Legal Services
- ☐ Homemaker
- ☐ Religious
- ☐ Mining
- ☐ Construction
- ☐ Other Manufacturing
- ☐ Retail
- ☐ Publishing
- ☐ Telecommunications
- ☐ Information Services and Data Processing
- ☐ Finance and Insurance
- ☐ College, University, and Adult Education
- ☐ Other Education Industry
- ☐ Arts, Entertainment, and Recreation
- ☐ Government and Public Administration
- ☐ Scientific or Technical Services
- ☐ Military
- ☐ Other Industry.

Please specify …..

In my current or previous jobs, I have (check all that apply):

- ☐ Been a member of a team
- ☐ Managed a team
- ☐ Worked part-time
- ☐ Hired personnel
- ☐ Commuted more than an hour to work
- ☐ Travelled on work assignments
- ☐ Dismissed an employee

Appendix D: Measurement Scales

Construct	Adapted Scale Items	Adapted from
Efficiency	1.Working with AI-Assist would… 1-save me time 2-speed up decision making 3-make me more efficient	(Wilkinson et al. 2021)
Integrity	2.I believe working with AI-Assist would … 1-provide unbiased recommendations 2-make honest decisions 3-make decisions with integrity	(Wilkinson et al. 2021)
Lack of Human Interaction	3.In completing this task, … 1-working with another human would be important to me 2-I would rather interact with a person than with AI-Assist 3-it would bother me to use AI-Assist when I could work with a human instead	(Dabholkar and Bagozzi 2002)
Incompatibility	In completing this task...	(Hari et al. 2021)

	1-I am uncertain that AI-Assist would recommend candidates that reflect my organization's needs 2-I am uncertain that AI-Assist would enable us to select candidates that are a good fit for my organization 3-I am not confident that AI-Assist would select candidates that match my organization's needs	
Explainability	In this scenario, - I felt I was provided with an explanation for the AI-Assist recommendation of which candidates to interview. - The AI-Assist provided a justification for its recommendation for the shortlisted candidates to interview	(Jabagi et al. 2021)
Autonomy	In this scenario, - I felt the AI was in control of deciding who would be invited for an interview - I felt the AI had authority over deciding who would be invited for an interview	(Ahuja and Thatcher 2005)
Willingness to Collaborate	1.In completing this task….	(Gursoy et al. 2019)

	1- I would be willing to work with AI-Assist 2-I would be happy to collaborate with AI-Assist 3-I would be likely to interact with AI-Assist	
Marker Variable (Attitude toward blue colour)	- I prefer blue to other colours - I like the colour blue - I like blue clothes.	(Miller 2021)
Marker Variable (Attitude towards public transit)	- If the buses come often enough, I would use the bus more often; - Where I choose to live is affected by transit service availability - During the period of heavy snow or rain, I prefer transit over	(Namgung and Akar 2014)
Attention Check Questions	In completing this task... - I am not paying any attention to this survey - I am answering questions in this survey without thinking	
Open-Ended Questions	- When collaborating with AI-Assist, do you think it is important that AI provides an explanation for its recommendation? Please, explain your response.	

	- When collaborating with AI-Assist, do you think it is important that the AI consults you before making a recommendation? Please, explain your response. - Is there any other thing you wish to add?

Appendix E: Demographic Questions

1. I identify as (Check one):
   [ ] Man

   [ ] Woman

   [ ]Non-gender-binary, two-spirit,or similar

   [ ] Prefer not to say

   [ ] Others

2. I'm (Check one):
   [ ] 20-30

   [ ] 31-40

   [ ] 41-50

   [ ] 51-60

   [ ] 61-70

   [ ] > 70

   [ ] Prefer not to answer

3. I have (Check one):
   [ ] High school diploma

   [ ] Some college degree

   [ ] Bachelors

   [ ] Master's

   [ ] Ph.D

   [ ] Other

   [ ] Prefer not to answer

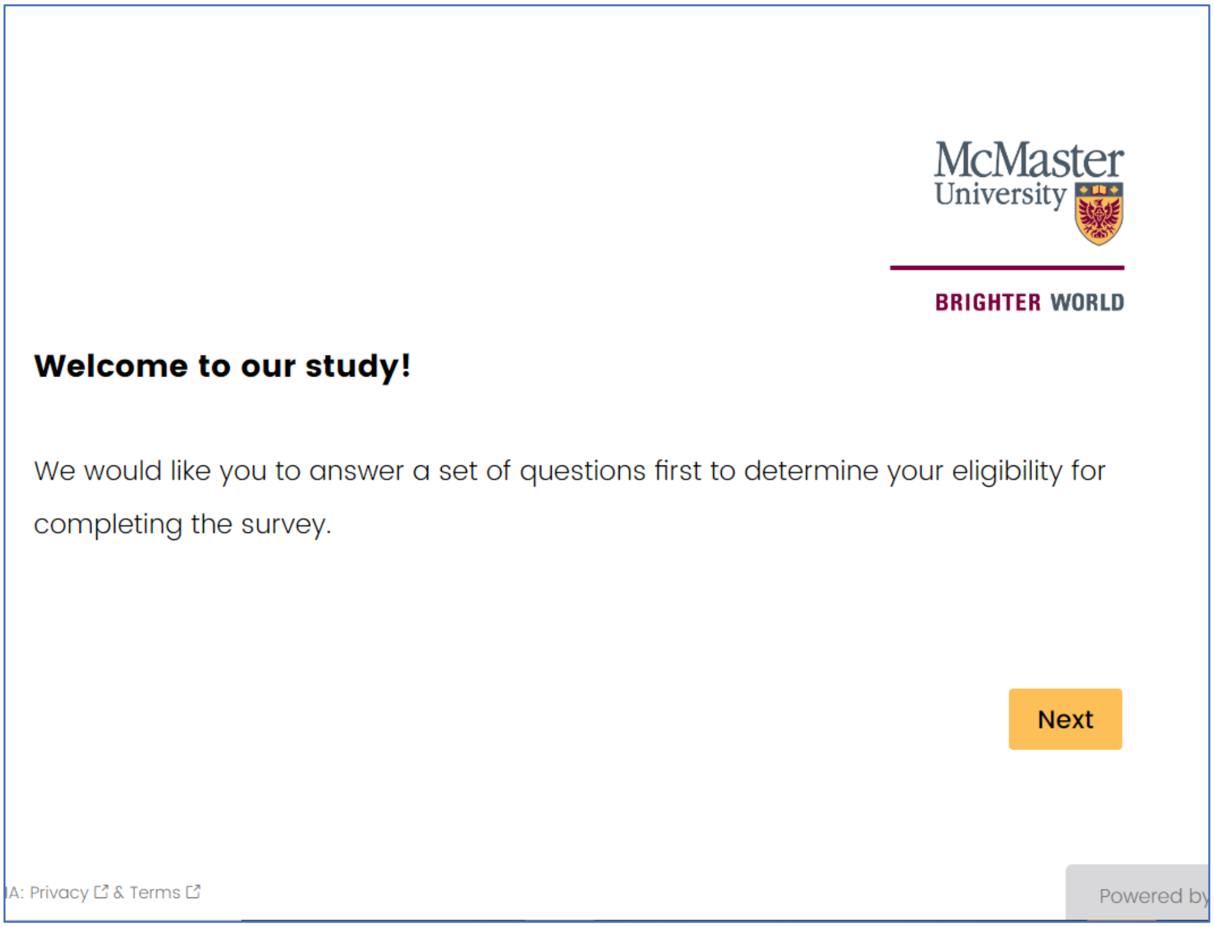
4. Total years of hiring experience:

5. How would you describe your knowledge level of AI:

- ○ No knowledge
- ○ Very basic knowledge
- ○ Medium knowledge level
- ○ Very good knowledge level
- ○ Excellent

Appendix F: Experimental Design

1- Participants were welcomed to the study.



2- Participants were then directed to the screening questions to determine their eligibility for taking the survey.

BRIGHTER WORLD

Role in the current company/organization:

- ◯ Upper Management
- ◯ Middle Management
- ◯ Junior Management
- ◯ Administrative Staff
- ◯ Support Staff
- ◯ Trained Professional
- ◯ Skilled Laborer
- ◯ Consultant
- ◯ Temporary Employee
- ◯ Researcher
- ◯ Student
- ◯ Self-Employed/Partner
- ◯ Other: Please Specify ....

Department

- Accounting
- Administration
- Customer Service
- Human Resources
- Inventory
- IT
- Logistics
- Manufacturing
- Marketing and Sales
- Procurement
- Quality Assurance
- Research & Development
- Other: Please Specify

Industry

- ◯ Agriculture, Forestry, Fishing and Hunting
- ◯ Utilities
- ◯ Computer and Electronics Manufacturing
- ◯ Wholesale
- ◯ Transportation and Warehousing
- ◯ Software
- ◯ Broadcasting
- ◯ Other Information Industry
- ◯ Real Estate, Rental and Leasing
- ◯ Primary/Secondary (K-12) Education
- ◯ Health Care and Social Assistance
- ◯ Hotel and Food Services
- ◯ Legal Services
- ◯ Homemaker
- ◯ Religious

- ○ Government and Public Administration
- ○ Scientific or Technical Services
- ○ Military
- ○ Other: Please Specify...

In my current or previous jobs, I have (check all that apply):

- ☐ Been a member of a team
- ☐ Managed a team
- ☐ Worked part-time
- ☐ Hired personnel
- ☐ Commuted more than an hour to work
- ☐ Dismissed an employee

Next

3- Participants who did not pass the screening phase and were ineligible to continue with the survey were provided with the below message:

Thank you for taking time to answer a few questions. Unfortunately, you do not qualify for this particular survey, but we do hope to hear from you in the future.

4- After screening participants and eliminating those who do not have hiring experience, demographic information was then collected from eligible participants.

5- After collecting demographic information, participants were directed to the consent form below.

Letter of Consent

A Study about Artificial Intelligence (AI)

• This research seeks to understand the potential that artificial intelligent (AI) technologies can bring to the workplace.

• AI can imitate intelligent human behaviour. They can perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting, problem-solving, and decision-making.

• In the recruitment context, AI can screen candidate applications, assess their profiles, and recommend candidates to support HR decision-makers.

Potential Harms, Risks or Discomforts: We do not foresee any significant risk or discomfort from your participation in this research. However, this study will collect data through an online survey, which is an externally hosted cloud-based service. Please note that whilst this service is approved for collecting data in this study by the McMaster Research Ethics Board, there is a small risk with any platform such as this of data that is collected on external servers falling outside the control of the research team.

Potential Benefits: This study will not benefit you directly. However, by participating in this study, you will help to provide designers and developers of AI with guidelines to improve the outcomes of using AI in the workplace.

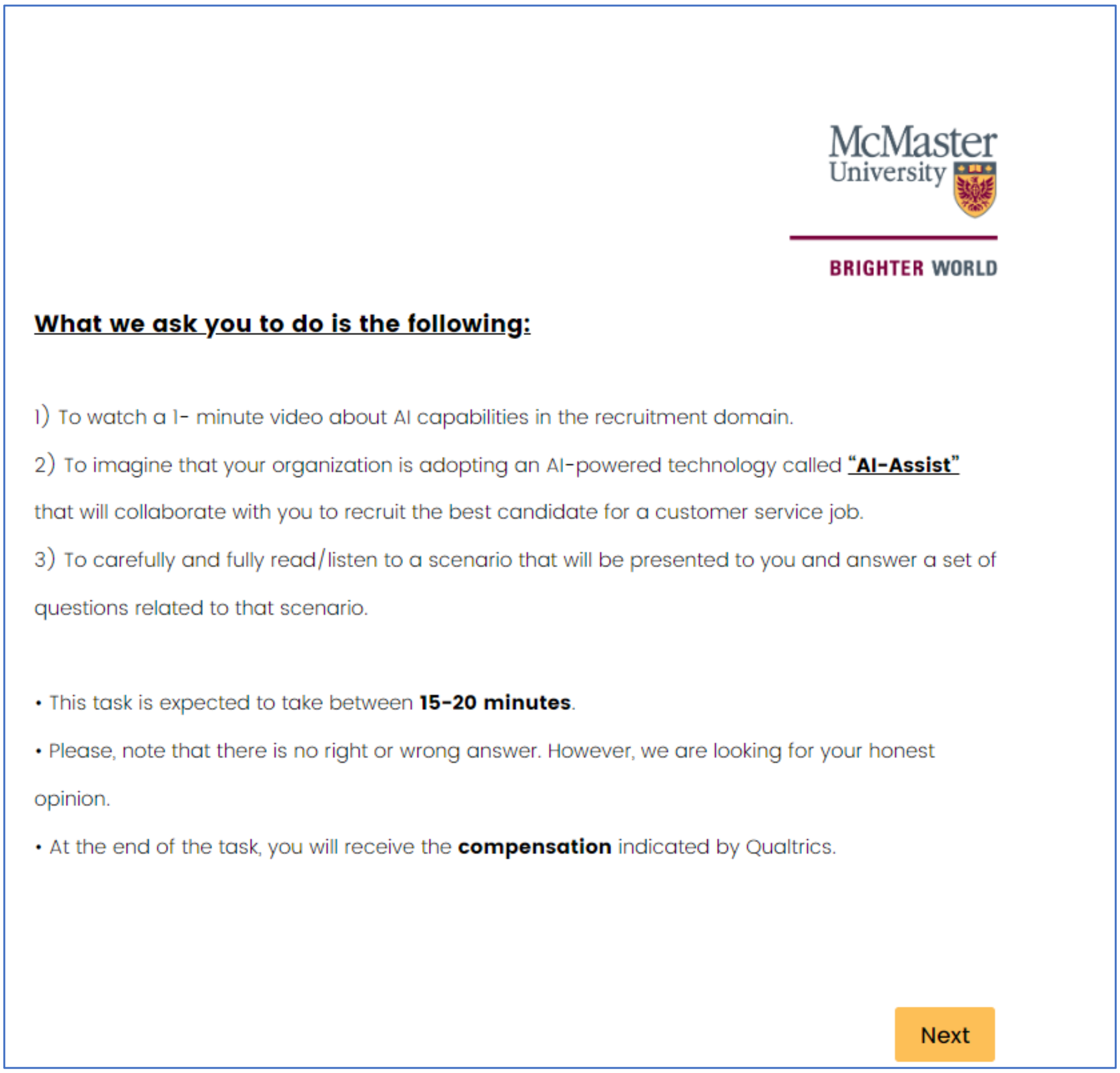
Payment or Reimbursement: if you agree to participate in this study, you will be compensated according to the scheme you agreed to with Qualtrics before you entered into the survey.

Confidentiality: Every effort will be made to protect (guarantee) your confidentiality and privacy. All information you supply during the research will be held in confidence. No personally-identifying information (e.g., name, social insurance number) will be required or collected before, during, or after the study, and therefore, your name will not appear in any report or publication of the research. The data will be collected through an online survey without the need to record any audios or videos. Your data will be safely stored on a password protected computer and only the student researcher and the supervisors will have access to this information. Data will be kept for approximately 4 years, after which the data will be completely deleted from any computer or storage drive. Confidentiality will be provided to the fullest extent possible by law.

Participation and Withdrawal: Your participation in this study is voluntary. It is your choice to be part of the study or not. If you decide to be part of the study, you can stop (withdraw) from whatever reason, even after giving consent or part-way through the study. Once you have submitted your responses for this anonymous survey: your answers will be put into a database and will not be identifiable. This means that once you have submitted your survey, your responses cannot be withdrawn from the study because it will not be possible for us to identify which responses are yours. Your decision whether or not to be part of the study will not affect your continuing access to services from Qualtrics.

This study has been reviewed by the McMaster University Research Ethics Board and received ethics clearance. If you have concerns or questions about your rights as a participant or about the way the study is conducted, please contact: McMaster Research Ethics Secretariat Telephone: (905) 525-9140 ext. 23142 C/o Research Office for Administrative Development and Support E-mail: ethicsoffice@mcmaster.ca.

6- After getting participants' consent, they were directed to a page that described to them what is expected from them during the experiment. On this page, I was keen on providing participants in a point format the steps they are going to go through next, as well as the expected time duration to complete their task.


McMaster University

BRIGHTER WORLD

What we ask you to do is the following:

1) To watch a 1- minute video about AI capabilities in the recruitment domain.

2) To imagine that your organization is adopting an AI-powered technology called **"AI-Assist"** that will collaborate with you to recruit the best candidate for a customer service job.

3) To carefully and fully read/listen to a scenario that will be presented to you and answer a set of questions related to that scenario.

• This task is expected to take between **15-20 minutes**.

• Please, note that there is no right or wrong answer. However, we are looking for your honest opinion.

• At the end of the task, you will receive the **compensation** indicated by Qualtrics.

Next

7- Next, participants were notified that they are going to watch the 1-minute video and were reminded of the importance of watching the entire video before moving forward to the next part of the survey.

McMaster University

BRIGHTER WORLD

- In the next page, please watch a 1-minute video in its entirety that gives you an idea about how AI-powered technologies can assist in the recruitment process.

Next

McMaster University

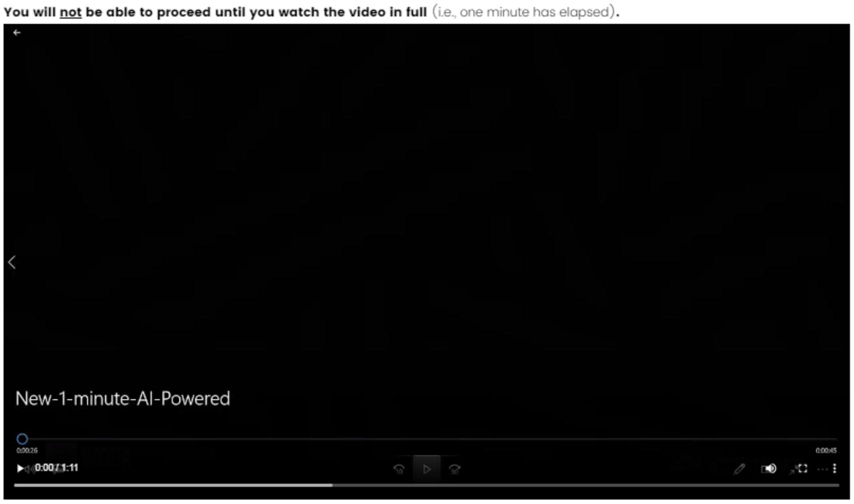
BRIGHTER WORLD

Timing

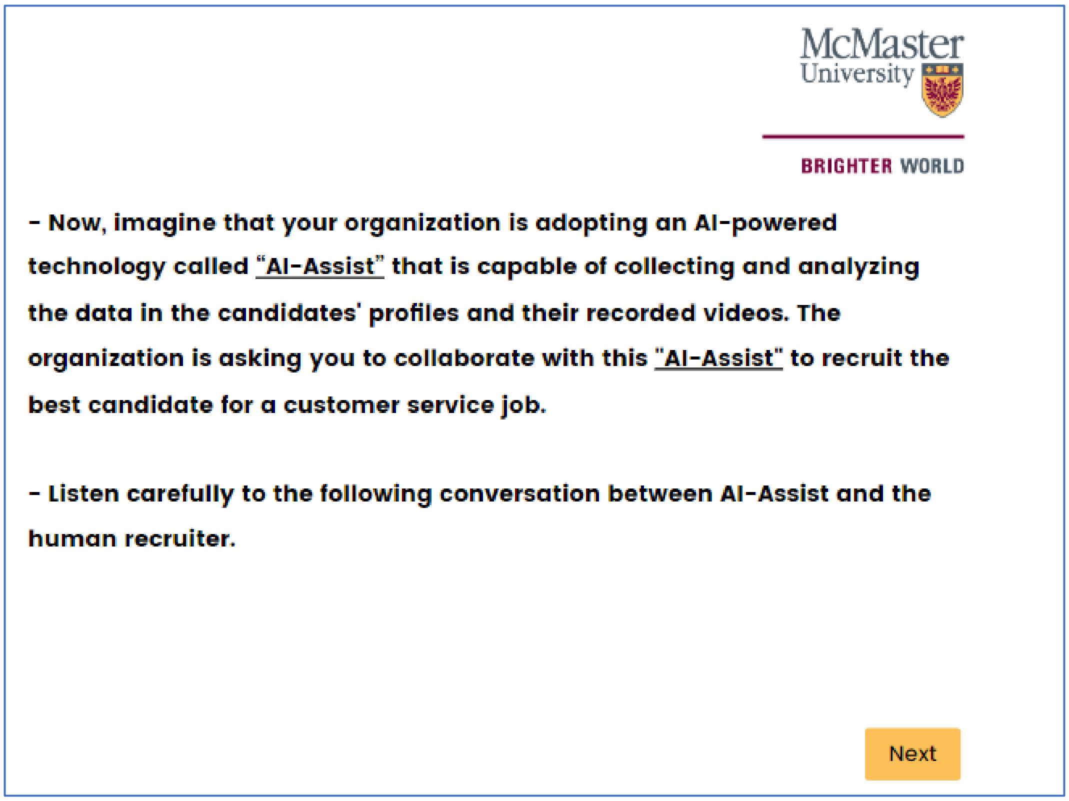
These page timer metrics will not be displayed to the recipient.

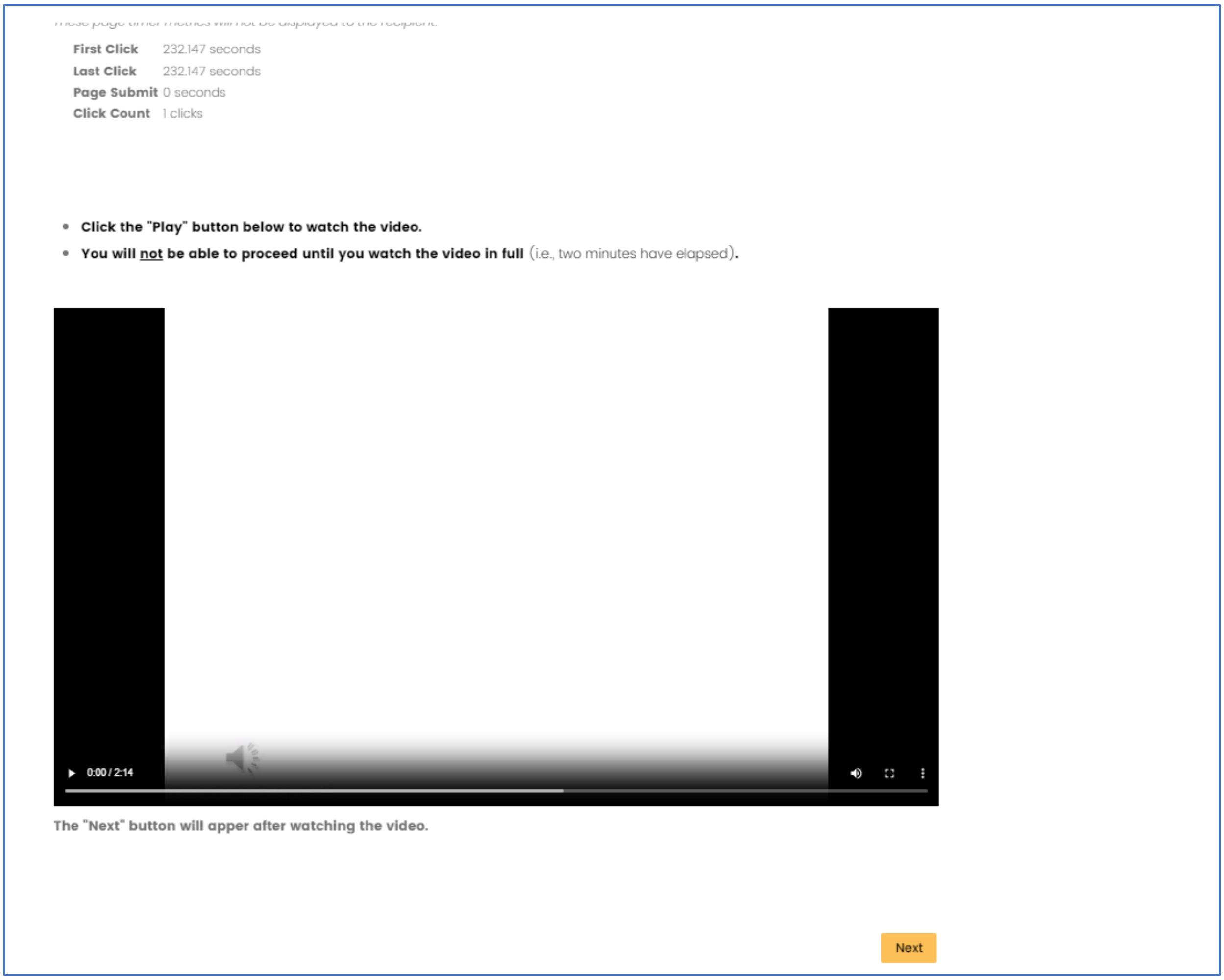
First Click 20.66 seconds
Last Click 20.66 seconds
Page Submit 0 seconds
Click Count 1 clicks

- **Click the "Play" button below to watch the video.**
- **You will not be able to proceed until you watch the video in full** (i.e., one minute has elapsed).



8- Next, participants were given the scenario and were introduced to AI-Assist. They were also notified to listen to a conversation taking place between AI-Assist and the human recruiter. At this point, participants were sent randomly to watch one of four videos that represented the treatment groups of this study.





9- After participants were exposed to the scenario and the videos, they were then asked to answer the set of survey questions.

10- After participants completed the survey, they were thanked for their time and were displayed the message below:

Thank you for taking this survey. Your response has been recorded successfully. You may now close this window.

Appendix G: Treatment Group Comparisons using Tukey's Test

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Efficiency	1	2	0.22846	0.16738	0.522	-0.2036	0.6605
		3	0.31290	0.16601	0.236	-0.1156	0.7414
		4	.47673*	0.16834	**0.025**	0.0422	0.9113
	2	1	-0.22846	0.16738	0.522	-0.6605	0.2036
		3	0.08443	0.16601	0.95 7	-0.3441	0.5130
		4	0.24827	0.16834	0.454	-0.1863	0.6828
	3	1	-0.31290	0.16601	0.236	-0.7414	0.1156
		2	-0.08443	0.16601	0.957	-0.5130	0.3441
		4	0.16383	0.16698	0.760	-0.2672	0.5949
	4	1	-.47673*	0.16834	**0.025**	-0.9113	-0.0422
		2	-0.24827	0.16834	0.454	-0.6828	0.1863
		3	-0.16383	0.16698	0.760	-0.5949	0.2672
Integrity	1	2	0.29213	0.17926	0.363	-0.1706	0.7549
		3	0.42342	0.17779	0.082	-0.0355	0.8824
		4	0.39115	0.18029	0.134	-0.0742	0.8565
	2	1	-0.29213	0.17926	0.363	-0.7549	0.1706
		3	0.13129	0.17779	0.882	-0.3276	0.5902
		4	0.09901	0.18029	0.947	-0.3664	0.5644
	3	1	-0.42342	0.17779	0.082	-0.8824	0.0355
		2	-0.13129	0.17779	0.882	-0.5902	0.3276
		4	-0.03228	0.17883	0.998	-0.4939	0.4293
	4	1	-0.39115	0.18029	0.134	-0.8565	0.0742
		2	-0.09901	0.18029	0.947	-0.5644	0.3664
		3	0.03228	0.17883	0.998	-0.4293	0.4939
Lack_of_Human	1	2	-.62921*	0.23195	**0.035**	-1.2279	-0.0305
		3	-0.48870	0.23005	0.147	-1.0825	0.1051
		4	-.69379*	0.23328	**0.017**	-1.2960	-0.0916
	2	1	.62921*	0.23195	0.035	0.0305	1.2279
		3	0.14051	0.23005	0.929	-0.4533	0.7343
		4	-0.06457	0.23328	0.993	-0.6667	0.5376
	3	1	0.48870	0.23005	0.147	-0.1051	1.0825
		2	-0.14051	0.23005	0.929	-0.7343	0.4533

		4	-0.20508	0.23139	0.812	-0.8024	0.3922
	4	1	.69379*	0.23328	0.017	0.0916	1.2960
		2	0.06457	0.23328	0.993	-0.5376	0.6667
		3	0.20508	0.23139	0.812	-0.3922	0.8024
Incompatibility	1	2	-0.48315	0.21155	0.104	-1.0292	0.0629
		3	-0.43775	0.20982	0.160	-0.9794	0.1038
		4	-0.54707	0.21276	**0.051**	-1.0963	0.0021
	2	1	0.48315	0.21155	0.104	-0.0629	1.0292
		3	0.04539	0.20982	0.996	-0.4962	0.5870
		4	-0.06393	0.21276	0.991	-0.6131	0.4853
	3	1	0.43775	0.20982	0.160	-0.1038	0.9794
		2	-0.04539	0.20982	0.996	-0.5870	0.4962
		4	-0.10932	0.21104	0.955	-0.6541	0.4354
	4	1	0.54707	0.21276	0.051	-0.0021	1.0963
		2	0.06393	0.21276	0.991	-0.4853	0.6131
		3	0.10932	0.21104	0.955	-0.4354	0.6541
Willingness	1	2	.49438*	0.18663	**0.042**	0.0126	0.9761
		3	0.30488	0.18511	0.354	-0.1729	0.7827
		4	0.37638	0.18770	0.188	-0.1081	0.8609
	2	1	-.49438*	0.18663	0.042	-0.9761	-0.0126
		3	-0.18950	0.18511	0.736	-0.6673	0.2883
		4	-0.11800	0.18770	0.923	-0.6025	0.3665
	3	1	-0.30488	0.18511	0.354	-0.7827	0.1729
		2	0.18950	0.18511	0.736	-0.2883	0.6673
		4	0.07151	0.18618	0.981	-0.4091	0.5521
	4	1	-0.37638	0.18770	0.188	-0.8609	0.1081
		2	0.11800	0.18770	0.923	-0.3665	0.6025
		3	-0.07151	0.18618	0.981	-0.5521	0.4091

Appendix H: One-Way ANOVAs for Control Variable Analysis

Gender[5]

Variable	Groups	N	Mean	95% Confidence Interval for Mean		ANOVA			
				Lower Bound	Upper Bound	Sum of Squares	Mean Square	F	Sig.
Efficiency	Men	208	5.6635	5.5033	5.8236	0.07	0.07	0.05	0.82
	Women	148	5.6913	5.5197	5.8629				
	Total	356	5.6751	5.5580	5.7922				
Integrity	Men	208	5.3221	5.1503	5.4939	1.36	1.36	0.94	0.33
	Women	148	5.4474	5.2657	5.6291				
	Total	356	5.3744	5.2493	5.4996				
Lack_of_Human	Men	208	4.1250	3.9136	4.3364	0.02	0.02	0.01	0.92
	Women	148	4.1409	3.8829	4.3990				
	Total	356	4.1317	3.9688	4.2945				
Incompatibility	Men	208	3.8269	3.6294	4.0245	0.01	0.01	0.00	0.95
	Women	148	3.8367	3.6112	4.0622				
	Total	356	3.8310	3.6830	3.9790				
Willingness	Men	208	5.5144	5.3358	5.6931	0.097	0.097	0.062	0.804
	Women	148	5.4810	5.2903	5.6717				
	Total	356	5.5005	5.3700	5.6309				

[5] Please note that Tukey's HSD test does not apply for groups that are less then three. Thus, no Tukey'HSD test results were reported for "Gender".

Age

Variable	Groups	N	Mean	95% Confidence Interval for Mean		ANOVA			
				Lower Bound	Upper Bound	Sum of Squares	Mean Square	F	Sig.
Efficiency	1	108	5.904	5.696	6.112	36.099	18.049	15.417	0.000
	2	98	5.745	5.529	5.961				
	3	79	5.426	5.174	5.678				
	4	72	5.509	5.232	5.787				
	Total	357	5.675	5.558	5.792				
Integrity	1	108	5.904	5.696	6.112	34.196	17.098	12.596	0.000
	2	98	5.745	5.529	5.961				
	3	79	5.426	5.174	5.678				
	4	72	5.509	5.232	5.787				
	Total	357	5.675	5.558	5.792				
Lack_of_Human	1	108	5.904	5.696	6.112	43.314	21.657	9.259	0.000
	2	98	5.745	5.529	5.961				
	3	79	5.426	5.174	5.678				
	4	72	5.509	5.232	5.787				
	Total	357	5.675	5.558	5.792				
Incompatibility	1	108	5.904	5.696	6.112	40.125	20.062	10.454	0.000
	2	98	5.745	5.529	5.961				
	3	79	5.426	5.174	5.678				
	4	72	5.509	5.232	5.787				
	Total	357	5.675	5.558	5.792				
Willingness	1	108	5.904	5.696	6.112	63.580	31.790	22.714	0.000
	2	98	5.745	5.529	5.961				

	3	79	5.426	5.174	5.678				
	4	72	5.509	5.232	5.787				
	Total	357	5.675	5.558	5.792				

Multiple Comparisons					
Tukey's HSD					
Dependent Variable	**Group (I)**	**Group (J)**	**Mean Difference (I-J)**	**Std. Error**	**Sig.**
Efficiency	1	2	-.38679*	0.147	0.024
		3	-.77871*	0.141	0.000
	2	1	.38679*	0.147	0.024
		3	-.39193*	0.136	0.012
	3	1	.77871*	0.141	0.000
		2	.39193*	0.136	0.012
Integrity	1	2	-0.346	0.158	0.076
		3	-.75411*	0.152	0.000
	2	1	0.346	0.158	0.076
		3	-.40839*	0.147	0.016
	3	1	.75411*	0.152	0.000
		2	.40839*	0.147	0.016
Lack_of_Human	1	2	0.411	0.208	0.120
		3	.85161*	0.199	0.000
	2	1	-0.411	0.208	0.120

		3	0.441	0.192	0.059
	3	1	-.85161*	0.199	0.000
		2	-0.441	0.192	0.059
Incompatibility	1	2	0.291	0.188	0.272
		3	.80047*	0.180	0.000
	2	1	-0.291	0.188	0.272
		3	.50973*	0.174	0.010
	3	1	-.80047*	0.180	0.000
		2	-.50973*	0.174	0.010
Willingness	1	2	-.50014*	0.161	0.006
		3	-1.03203*	0.154	0.000
	2	1	.50014*	0.161	0.006
		3	-.53188*	0.149	0.001
	3	1	1.03203*	0.154	0.000
		2	.53188*	0.149	0.001
*. The mean difference is significant at the 0.05 level.					

Education Level

Variable	Groups	N	Mean	95% Confidence Interval for Mean		ANOVA			
				Lower Bound	Upper Bound	Sum of Squares	Mean Square	F	Sig.
Efficiency	1	121	5.72	5.52	5.92	2.94	1.47	1.18	0.31
	2	140	5.75	5.55	5.94				
	3	87	5.52	5.30	5.75				
	Total	348	5.68	5.57	5.80				
Integrity	1	121	5.38	5.16	5.60	2.15	1.08	0.74	0.48
	2	140	5.47	5.27	5.67				
	3	87	5.27	5.01	5.53				
	Total	348	5.39	5.26	5.52				
Lack_of_Human	1	121	4.20	3.91	4.49	7.39	3.70	1.53	0.22
	2	140	3.95	3.70	4.21				
	3	87	4.30	3.97	4.63				
	Total	348	4.13	3.96	4.29				
Incompatibility	1	121	3.85	3.60	4.09	8.89	4.45	2.25	0.11
	2	140	3.67	3.42	3.91				
	3	87	4.07	3.78	4.37				
	Total	348	3.83	3.68	3.98				
Willingness	1	121	5.47	5.24	5.71	1.39	0.69	0.45	0.64
	2	140	5.59	5.38	5.80				
	3	87	5.46	5.21	5.71				
	Total	348	5.52	5.39	5.65				

Multiple Comparisons							
Tukey HSD							
Dependent Variable	Group (I)	Group (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Efficiency	1	2	-0.026	0.139	0.981	-0.353	0.301
		3	0.197	0.157	0.423	-0.173	0.567
	2	1	0.026	0.139	0.981	-0.301	0.353
		3	0.223	0.153	0.312	-0.137	0.582
	3	1	-0.197	0.157	0.423	-0.567	0.173
		2	-0.223	0.153	0.312	-0.582	0.137
Integrity	1	2	-0.091	0.150	0.815	-0.444	0.261
		3	0.108	0.170	0.799	-0.291	0.507
	2	1	0.091	0.150	0.815	-0.261	0.444
		3	0.199	0.165	0.448	-0.188	0.587
	3	1	-0.108	0.170	0.799	-0.507	0.291
		2	-0.199	0.165	0.448	-0.587	0.188
Lack_of_Human	1	2	0.245	0.193	0.413	-0.209	0.698
		3	-0.101	0.218	0.890	-0.614	0.413
	2	1	-0.245	0.193	0.413	-0.698	0.209
		3	-0.345	0.212	0.235	-0.844	0.154
	3	1	0.101	0.218	0.890	-0.413	0.614
		2	0.345	0.212	0.235	-0.154	0.844
Incompatibility	1	2	0.179	0.175	0.561	-0.232	0.590
		3	-0.227	0.198	0.485	-0.692	0.238
	2	1	-0.179	0.175	0.561	-0.590	0.232
		3	-0.406	0.192	0.088	-0.858	0.046
	3	1.00	0.227	0.198	0.485	-0.238	0.692
		2.00	0.406	0.192	0.088	-0.046	0.858
Willingness	1	2.00	-0.122	0.155	0.711	-0.486	0.243
		3.00	0.015	0.175	0.996	-0.398	0.428
	2	1.00	0.122	0.155	0.711	-0.243	0.486

		3.00	0.137	0.170	0.701	-0.264	0.538
	3	1.00	-0.015	0.175	0.996	-0.428	0.398
		2.00	-0.137	0.170	0.701	-0.538	0.264

Knowledge Level of AI

Variable	Groups	N	Mean	95% Confidence Interval for Mean		ANOVA			
				Lower Bound	Upper Bound	Sum of Squares	Mean Square	F	Sig.
Efficiency	1.00	102	5.245	5.014	5.476	36.10	18.05	15.42	0.00
	2.00	115	5.632	5.424	5.840				
	3.00	140	6.024	5.862	6.186				
	Total	357	5.675	5.558	5.792				
Integrity	1.00	102	4.967	4.720	5.214	34.20	17.10	12.60	0.00
	2.00	115	5.313	5.108	5.518				
	3.00	140	5.721	5.531	5.912				
	Total	357	5.374	5.249	5.500				
Lack_of_Human	1.00	102	4.598	4.294	4.902	43.31	21.66	9.26	0.00
	2.00	115	4.187	3.925	4.449				
	3.00	140	3.746	3.479	4.014				
	Total	357	4.132	3.969	4.294				
Incompatibility	1.00	102	4.239	3.976	4.501	40.12	20.06	10.45	0.00
	2.00	115	3.948	3.698	4.198				
	3.00	140	3.438	3.197	3.680				
	Total	357	3.831	3.683	3.979				
Willingness	1.00	102	4.935	4.667	5.202	63.58	31.79	22.71	0.00
	2.00	115	5.435	5.227	5.643				
	3.00	140	5.967	5.786	6.148				
	Total	357	5.500	5.370	5.631				

Multiple Comparisons						
Tukey HSD						
Dependent Variable	Group (I)	Group (J)	Mean Difference (I-J)	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Efficiency	1	2	-.38679*	0.024	-0.7332	-0.040
		3	-.77871*	0.00	-1.11	-0.447
	2	1	.38679*	0.02	0.04	0.733
		3	-.39193*	0.01	-0.71	-0.071
	3	1	.77871*	0.00	0.45	1.110
		2	.39193*	0.01	0.07	0.712
Integrity	1	2	-0.3457	0.08	-0.72	0.027
		3	-.75411*	0.00	-1.11	-0.397
	2	1	0.3457	0.08	-0.03	0.719
		3	-.40839*	0.02	-0.75	-0.063
	3	1	.75411*	0.00	0.40	1.111
		2	.40839*	0.02	0.06	0.753
Lack_of_Human	1	2	0.4111	0.12	-0.08	0.901
		3	.85161*	0.00	0.38	1.320
	2	1	-0.4111	0.12	-0.90	0.079
		3	0.4405	0.06	-0.01	0.894
	3	1	-.85161*	0.00	-1.32	-0.383
		2	-0.4405	0.06	-0.89	0.012

Incompatibility	1	2	0.2907	0.27	-0.15	0.734
		3	.80047*	0.00	0.38	1.225
	2	1	-0.2907	0.27	-0.73	0.153
		3	.50973*	0.01	0.10	0.920
	3	1	-.80047*	0.00	-1.22	-0.376
		2	-.50973*	0.01	-0.92	-0.099
Willingness	1	2	-.50014*	0.01	-0.88	-0.121
		3	-1.03203*	0.00	-1.39	-0.670
	2	1	.50014*	0.01	0.12	0.879
		3	-.53188*	0.00	-0.88	-0.181
	3	1	1.03203*	0.00	0.67	1.394
		2	.53188*	0.00	0.18	0.882
*. The mean difference is significant at the 0.05 level.						

<u>**Years of Hiring Experience**</u>

Variable	Groups	N	Mean	95% Confidence Interval for Mean		ANOVA			
				Lower Bound	Upper Bound	Sum of Squares	Mean Square	F	Sig.
Efficiency	1.00	5	6.000	3.849	8.151	0.51	0.17	0.12	0.95
	2.00	15	5.711	5.214	6.209				
	3.00	7	5.810	4.997	6.622				
	4.00	51	5.895	5.550	6.240				
	Total	78	5.859	5.597	6.121				
Integrity	1.00	5	5.933	4.579	7.287	8.15	2.72	1.68	0.18
	2.00	15	5.111	4.371	5.852				
	3.00	7	5.095	3.988	6.203				
	4.00	51	5.810	5.452	6.169				
	Total	78	5.620	5.329	5.910				
Lack_of_Human	1.00	5	3.700	2.271	5.129	9.81	3.27	1.31	0.28
	2.00	15	3.200	2.397	4.003				
	3.00	7	4.571	2.812	6.331				
	4.00	51	3.863	3.412	4.314				
	Total	78	3.788	3.430	4.147				
Incompatibility	1.00	5	4.533	2.411	6.656	8.33	2.78	1.14	0.34
	2.00	15	3.467	2.550	4.384				
	3.00	7	4.048	2.771	5.324				
	4.00	51	3.366	2.931	3.801				
	Total	78	3.521	3.168	3.875				
Willingness	1.00	5	5.733	3.516	7.951	0.18	0.06	0.03	0.99
	2.00	15	5.622	5.096	6.148				

	3.00	7	5.810	4.658	6.961				
	4.00	51	5.699	5.310	6.089				
	Total	78	5.697	5.403	5.990				

Multiple Comparisons						
Tukey HSD						
Dependent Variable	**Group (I)**	**Group (J)**	**Mean Difference (I-J)**	**Sig.**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
Efficiency	1	2	0.289	0.965	-1.314	1.891
		3	0.190	0.993	-1.627	2.008
		4	0.105	0.998	-1.350	1.559
	2	1	-0.289	0.965	-1.891	1.314
		3	-0.098	0.998	-1.519	1.322
		4	-0.184	0.951	-1.096	0.727
	3	10	-0.190	0.993	-2.008	1.627
		2	0.098	0.998	-1.322	1.519
		4	-0.086	0.998	-1.337	1.165
	4	1	-0.105	0.998	-1.559	1.350
		2	0.184	0.951	-0.727	1.096
		3	0.086	0.998	-1.165	1.337
Integrity	1	2	0.822	0.596	-0.905	2.549
		3	0.838	0.675	-1.120	2.796

		4	0.123	0.997	-1.444	1.690
	2	1	-0.822	0.596	-2.549	0.905
		3	0.016	1.000	-1.515	1.547
		4	-0.699	0.249	-1.682	0.283
	3	1	-0.838	0.675	-2.796	1.120
		2	-0.016	1.000	-1.547	1.515
		4	-0.715	0.507	-2.063	0.633
	4	1	-0.123	0.997	-1.690	1.444
		2	0.699	0.249	-0.283	1.682
		3	0.715	0.507	-0.633	2.063
Lack_of_Human	1	2	0.500	0.928	-1.646	2.646
		3	-0.871	0.783	-3.305	1.562
		4	-0.163	0.996	-2.110	1.785
	2	1	-0.500	0.928	-2.646	1.646
		3	-1.371	0.239	-3.273	0.531
		4	-0.663	0.487	-1.883	0.558
	3	1	0.871	0.783	-1.562	3.305
		2	1.371	0.239	-0.531	3.273
		4	0.709	0.683	-0.966	2.384
	4	1	0.163	0.996	-1.785	2.110
		2	0.663	0.487	-0.558	1.883
		3	-0.709	0.683	-2.384	0.966
Incompatibility	1	2	1.067	0.553	-1.056	3.189
		3	0.486	0.951	-1.921	2.892
		4	1.167	0.389	-0.759	3.093
	2	1	-1.067	0.553	-3.189	1.056
		3	-0.581	0.849	-2.462	1.300
		4	0.101	0.996	-1.106	1.308
	3	1	-0.486	0.951	-2.892	1.921
		2	0.581	0.849	-1.300	2.462
		4	0.682	0.702	-0.975	2.338
	4	1	-1.167	0.389	-3.093	0.759
		2	-0.101	0.996	-1.308	1.106
		3	-0.682	0.702	-2.338	0.975
Willingness	1	2	0.111	0.998	-1.692	1.914
		3	-0.076	1.000	-2.120	1.968
		4	0.034	1.000	-1.602	1.670
	2	1	-0.111	0.998	-1.914	1.692
		3	-0.187	0.990	-1.785	1.411
		4	-0.077	0.997	-1.102	0.948
	3	1	0.076	1.000	-1.968	2.120

		2	0.187	0.990	-1.411	1.785
		4	0.110	0.997	-1.297	1.517
	4	1	-0.034	1.000	-1.670	1.602
		2	0.077	0.997	-0.948	1.102
		3	-0.110	0.997	-1.517	1.297

Appendix I: Adding Age and Knowledge Level of AI to The Model

Adding Age to the Model

Path	Path Coeffifienct	P-value
Age → Efficiency	-0.127	0.008
Age → Incompatibility	0.165	0.002
Age → Integrity	-0.106	0.021
Age → Lack-of-Human	0.114	0.017
Age → Willingness	-0.001	0.488
Efficiency → Willingness	0.464	0.000
Integrity → Willingness	0.239	0.000
Lack_of_Human → Willingness	-0.261	0.000
Incompatibility → Willingness	-0.038	0.333
Autonomy → Efficiency	0.184	0.000
Autonomy → Integrity	0.074	0.000
Autonomy → Lack_of_Human	0.096	0.035
Autonomy → Incompatibility	0.082	0.000
Explainability → Efficiency	0.502	0.000
Explainability → Integrity	0.498	0.000
Explainability → Lack_of_Human	-0.330	0.000
Explainability → Incompatibility	-0.380	0.000

Adding Knowledge Level of AI to the Model

Path	Path Coefficient	P-value
AI_Knowledge → Efficiency	0.184	0.000
AI_Knowledge → Incompatibility	-0.198	0.001
AI_Knowledge → Integrity	0.175	0.001
AI_Knowledge → Lack-of-Human	-0.169	0.003
AI_Knowledge → Willingness	0.072	0.022
Efficiency → Willingness	0.453	0.000
Integrity → Willingness	0.239	0.000
Lack_of_Human → Willingness	-0.261	0.000
Incompatibility → Willingness	-0.038	0.338
Autonomy → Efficiency	0.184	0.000
Autonomy → Integrity	0.074	0.000
Autonomy → Lack_of_Human	0.096	0.037
Autonomy → Incompatibility	0.082	0.000
Explainability → Efficiency	0.502	0.000
Explainability → Integrity	0.498	0.000
Explainability → Lack_of_Human	-0.330	0.000
Explainability → Incompatibility	-0.380	0.000