The Potential of Artificial Intelligence (AI) to improve Decision Making:

Investigating the Reliance on AI Advice in a Business Context

Eva Kim Fortuin

Dissertation written under the supervision of professor Filipa de Almeida

Dissertation submitted in partial fulfilment of requirements for the MSc in International Management with a specialization in Strategy and Consulting, at the Universidade Católica Portuguesa, January 2022.

Abstract

Title: The Potential of Artificial Intelligence (AI) to improve Decision Making: Investigating the Reliance on AI Advice in a Business Context

Author: Eva Kim Fortuin

With the intention of overcoming human decision making biases, organizations are increasingly using AI as decision support. However, to unlock the full potential of AI-based advice to improve decision making, users must be willing to rely on it in the first place. To better understand people's readiness to accept advice from AI, two experimental studies were conducted in the scope of this research.

Study 1 examined whether people rely more on advice coming from AI or a human. People showed algorithm appreciation in both tasks – the performance evaluation of an employee and the closing price prediction of a stock. The effect was fully mediated by people's trust in the source and varied across different levels of confidence in one's own decision. Study 2 examined whether people also choose AI advice over human advice when presented with both options and whether they choose equally for themselves and for others. In this setting, algorithm appreciation persisted only for the stock price prediction task, irrespectively of who the decision was made for. Furthermore, several influencing factors were identified that point to domains where AI is most likely to be accepted and ways in which its benefits can be maximized.

The results from these studies have clear implications for organizations that turn to Big Data and AI-generated advice to improve decision making, suggesting that AI might be a good addition to their daily operations.

Keywords: Artificial Intelligence, AI Advice, Strategic Decision Making, Reliance on Advice, JAS Paradigm, Trust, Self-Confidence

Sumário

Título: O Potencial da Inteligência Artificial (IA) para melhorar a Tomada de Decisões: Investigar a Confiança no aconselhamento sobre IA num contexto empresarial

Autor: Eva Kim Fortuin

Com a intenção de superar enviesamentos na tomada de decisão, as empresas e organizações estão cada vez mais a usar modelos de Inteligência Artificial (AI) como suporte aos processos de tomada de decisão. Contudo, de modo a atingir o seu potencial máximo, os usuários devem confiar na IA em primeiro lugar.

Para melhor compreender a aceitação de conselhos de IA, dois estudos foram conduzidos no âmbito desta dissertação. O Estudo 1 examinou se as pessoas se baseiam mais nos conselhos da IA ou de humanos. As pessoas valorizaram o algoritmo para ambas as tarefas – a avaliação de desempenho de um funcionário e a previsão do valor final de uma ação. O efeito foi totalmente mediado pela confiança das pessoas na fonte e variou de acordo com o nível de confiança nas suas decisões. O Estudo 2 examinou se as pessoas preferem conselhos da IA ou de humanos, e se escolhem igualmente para si e para outros. Nesse cenário, a valorização do algoritmo persistiu apenas para a previsão de preço das ações, independentemente de quem seria influenciado pela mesma. Também, foram identificados diversos fatores que apontam para domínios em que a IA tem maior probabilidade de ser aceite e formas pelas quais as suas vantagens e benefícios podem ser maximizadas.

Os resultados destes estudos têm implicações evidentes para as empresas e organizações que recorrem a Big Data e a conselhos de AI para melhorar a tomada de decisão, sugerindo que esta tecnologia pode ser um bom complemento para suas operações diárias.

Palavras-chave: Inteligência Artificial, Aconselhamento Artificial, Tomada de Decisão Estratégica, Confiança de Aconselhamento, Paradigma JAS, Confiança, Autoconfiança

Acknowledgements

This dissertation marks the end of an exciting chapter in my educational career, in which I have been able to deepen my knowledge and experience personal growth. Especially the experience of starting all over again in a foreign country, establishing new friendships, learning a new language, and mastering all the related challenges have helped me prepare for my upcoming professional career. This would not have been possible without all the people who supported me along the way.

I would like to thank my professor Filipa de Almeida, who supported and mentored me throughout this dissertation. Thank you for always taking the time to answer my questions and for guiding me through the entire process. Above all, you were always a good motivator and encouraged me to do a second study to gain even more insights. I would also like to thank my friends and family for their constant support, as well as everyone who participated in my surveys.

Thanks to all of you!

Table of Contents

List of Figures

List of Tables

List of Abbreviations

&	And
AI	Artificial Intelligence
ANOVA	Analysis of Variance
b	Regression coefficient
df	Degrees of freedom
F	F-statistic
H1	Hypothesis 1 (2-5 respectively)
IMM	Index of Moderated Mediation
M	Sample mean
ML	Machine Learning
N	Total number of cases
p	p-value
R^2	Multiple correlation squared; measure of strength of association
SD	Standard Deviation
SE	Standard Error
t	t-statistic

Model 14 of Hayes PROCESS macro for SPSS

M	Mediator
MW	Interaction term between mediator and moderator
W	Moderator
X	Independent variable
Y	Dependent variable

1 Introduction

“***If you want the bias out, get the algorithms in.***”

Andrew McAfee of MIT (2018)

1.1 Opening Thought

Biases and heuristics have a huge impact on human decision making. Their importance is hard to overstate when aiming at improving decision outcomes. Tversky and Kahneman (1974) first introduced the term *cognitive bias* describing a systematic pattern of deviation from rationality in judgment. Consistent with the theory of bounded rationality, research proves that decision makers are prone to biases when making decisions. This also applies to contexts such as strategic decision making and planning which are often influenced by a high level of uncertainty (Barnes, 1984; Hodgkinson et al., 1999).

Another perspective was offered by Thaler (2015), who defined a choice as systematically biased if the judgment made by a decision maker differs from the judgment that a robot would have made. One of such instances is when asked to estimate a certain quantity, humans are most likely to give estimates that are systematically biased toward an anchor that they have been randomly exposed to previously, a phenomenon called anchoring bias (Tversky & Kahneman, 1974). A robot on the other hand would not be influenced by a small or large numeric anchor but would provide exactly the same estimate regardless (Bellé et al., 2018).

1.2 Relevance of the Topic

Due to the continuous rise of Big Data, companies in almost every industry are looking for ways to exploit the vast amount of data available and turn it into a competitive advantage (Provost & Fawcett, 2013). This has led to an increased deployment of Machine Learning (ML) models in decision making systems, where Artificial Intelligence (AI) helps humans to take better decisions in domains such as finance, human resources, healthcare, criminal justice, among others (Lai et al., 2021). In these cases, AI systems are used to improve decision outcomes by making use of the available data and by counteracting cognitive biases in decisions (Wang et al., 2019). So far, the main role of these systems is to augment or assist humans in their decision making by providing predictions or recommendations for certain tasks. Humans, however, can choose to follow or ignore the advice. Thus, in order to turn AI-based solutions into a competitive advantage and significantly improve strategic decision making, it is important to ensure the acceptance and usage of AI. Along with the introduction of new AI recommendation systems, there has been a call for research to better explain people's reliance on advice from algorithms (Kleinberg et al., 2018; Logg, 2017). While in the past this advice was typically obtained from humans, both research and practice show that there is a shift towards AI-based advice (Schemmer et al., 2022).

1.3 Problem Statement and Research Objective

Several streams of research have aimed to narrow the gap between AI and human decision-makers and the results are somewhat contradictory. Previous research looked at how people respond to advice from algorithms compared to their own, self-generated decisions. Even though algorithms often outperform human judgment, people are still cautious when relying on them and tend to distrust algorithmic output – also referred to as *algorithm aversion* (Dawes, 1979). This is reinforced by the findings that individuals typically tend to be overly confident about their own predictions and estimates and inaccurately discount advice when making quantitative judgements (Cain et al., 2015; Yaniv & Kleinberger, 2000). On the other hand, a more recent stream of studies shows a preference for algorithmic advice – also referred to as *algorithm appreciation* (Logg et al., 2019). According to Lee and See (2004), the intention to rely on the automation is strongly determined by the level of trust in AI, combined with other attitudes such as self-confidence.

However, there has been little research comparing how people respond to advice from algorithms versus advice from other humans. Literature further raises the question of whether people generally trust advisors too little or themselves too much, which are often hard to disentangle (Logg et al., 2019). These are important topics to be discussed as companies increasingly adopt AI as decision aids with the promise of overcoming biases in human decision making. The aim of this research is to bridge this gap and directly compare the reliance on AI advice vs. human advice. This allows for consideration of the fact that people generally disregard advice in favor of their own judgment. Furthermore, it is critical to understand whether there are self-other differences, as managers need to choose AI advice not only for themselves but also for others (Kray & Gonzalez, 1999). However, only a few authors have investigated the prevalence of self–other differences in the context of AI advice (Gai & Klesse, 2019). At this point in time, algorithms do not leave out the human decision maker (yet). Instead, AI-based decision making emerges from an interaction between the two. Thus, in order to reduce cognitive biases in decision making, humans need to be willing to accept advice from AI in the first place. Therefore, this dissertation intends to provide an answer to the following research question:

Research question: *Does AI have the potential to improve strategic decision making?*

To address the identified research gap, the central research question was divided into five sub-questions reading as follows:

1. *Do individuals generally show more trust towards AI advisors or human advisors?*
2. *Do high levels of trust increase the reliance on advice and thus, lead to a less biased decision outcome?*
3. *How does the confidence individuals have in their own decisions impact their reliance on advice?*
4. *Are individuals more likely to rely on advice from AI or from humans?*
5. *Do individuals show a different preference when choosing the source of advice for themselves vs. for others?*

To answer these questions two experimental studies were conducted. In Study 1, I measured the reliance on advice and tested whether participants relied more on human advice or AI advice. In Study 2, I measured the source preference when presented with both options and tested whether it was the same when deciding for oneself vs. for others. In doing so, this research intends to contribute to the growing body of literature on the interaction between humans and AI and provides insights into factors that are critical in determining whether AI has the potential to improve decision making. This lays the foundation for an effective commercialization of new AI systems as their success depends on users' willingness to rely on them.

1.4 Structure of the Dissertation

The structure of this dissertation follows the classical structure of empirical research papers. The introduction has already defined the general topic, the problem statement, and the research question(s). To justify the latter and derive respective hypotheses, Chapter 2 reviews existing literature and summarizes relevant concepts and paradigms related to cognitive biases, advice taking, and AI-based decision making. Chapter 3 and 4 describe the two studies that were conducted to answer the research question(s). Chapter 5 discusses the results in relation to existing literature, derives theoretical and practical implications, and points out limitations of this research. Finally, the conclusion wraps up this dissertation by summarizing the findings in a way that is actionable for managers.

2 Literature Review

The following chapter deals with the theoretical foundations of cognitive biases in strategic decision making and the potential of AI to improve such decisions with a special focus on people's reliance on AI advice.

2.1 Cognitive Biases in Strategic Decision Making

2.1.1 Strategic Decision Making and the Concept of Bounded Rationality

Strategic decision making is a dynamic process that falls into the category of decisions that are subject to uncertainty. It refers to "the process by which top management makes its most fundamental decisions" (Das & Teng, 1999, p. 758), which starts by acquiring and analyzing all relevant information. In today's complex world, managers often deal with a vast amount and variety of data, which must be interpreted correctly. To opt for the best strategic decisions, managers must pay particular attention to the collection and selection of data (Acciarini et al., 2020). Strategic decisions are "important, in terms of the action taken, the resources committed, or the precedents set" (Mintzberg et al., 1976, p. 246). Thus, this process should be as efficient as possible so that decisions can be effectively translated into strategy.

Due to uncertainty, decision makers need to rely on heuristics which can lead to errors in their decisions. According to Metzger and Spengler (2019), the degree of rationality within a decision can vary, as some aspects may be influenced more intuitively than others. Challenging the concept of homo economicus, Simon (1957) was the first to introduce the term of *bounded rationality,* which refers to the idea that rationality is limited when individuals make decisions due to limitations in thinking capacity, available information, and time. Thus, individuals rely on simplifying strategies that they use to cope with complex and uncertain decisions (Busenitz & Barney, 1997). Along with it comes the risk of bias, which can result in incorrect problem definitions or erroneous alternative assessments (Boone et al., 2019). Thus, to improve strategic decision making, it is important to look at the origin of such deviations from rationality.

2.1.2 System 1 and System 2

"Cognitive biases are an ever-present ingredient of strategic decision making" (Das & Teng, 1999, p. 757). To better understand cognitive biases, I will start with the dual process theory by Kahneman (2011). In his book, *Thinking Fast and Slow*, Daniel Kahneman introduced the metaphor of two systems: System 1 (thinking fast) is effortless and operates automatically, while System 2 (thinking slow) is effortful and consumes a great deal of energy as it involves careful reasoning. For example, for most people it is easy to drive a car on an empty road, but it requires effort to park a car in a narrow space. This energy is only used when necessary and much less often than we might think. System 1 works without voluntary effort and relies on mental shortcuts to reach a conclusion. It can be very useful, particularly in dangerous situations, as it allows us to act rapidly and without having to think too much. But it can also lead to systematic biases, impacting our ability to make rational decisions (Tversky & Kahneman, 1974).

2.1.3 Cognitive Biases

Heuristics give rise to a plethora of cognitive biases that affect human judgment in somewhat predictable ways and are often used to explain deviations from rational decision making (Kahneman et al. 1982; Schwenk 1988). Busenitz and Barney (1997, p. 12) define biases and heuristics as "decision rules, cognitive mechanisms, and subjective opinions people use to assist in making decisions".

Decades-old research on nonrational decision making has examined a large number of biases and heuristics (e.g., Epley & Gilovich, 2006; Nickerson, 1998). However, Tversky and Kahneman (1974) state that biases result from three fundamental heuristics: representativeness, availability, as well as anchoring and adjustment (also called anchoring bias). Out of all these biases and heuristics, this study primarily focuses on anchoring bias, as it is frequently found in strategic decision making. Anchoring occurs when individuals base their decisions on an initial numerical assessment – the anchor – and then fail to make sufficient adjustments from this anchor before providing their final answer (Das & Teng, 1999). Several researchers have demonstrated the robustness of this bias (George et al., 2000), which calls for new ways to overcome this bias.

2.1.4 Anchoring Bias

Out of all cognitive biases that keep people from taking rational decisions, anchoring is one of the most studied (Cen et al, 2013). In various experiments, Tversky and Kahneman (1974) explore the premise that individuals commonly form estimates by starting with a readily available reference value and then adjusting from this value. For example, they set up a wheel of fortune and had people spin it. They were then asked how many African member states the UN had. People for whom the wheel of fortune had stopped on a low number indicated a lower number of member states. People who had spun a high number estimated a higher number. This shows that whenever people estimate something, they tend to use an anchor – in this case the number spun on the wheel of fortune, despite it having nothing to do with the number of African UN member states. The anchor is an initial value on which all further estimates and evaluations depend (Tversky & Kahnemann, 1974). This mechanism is problematic, as people typically fail to adequately adjust their final estimates away from the (sometimes) important but over-emphasized starting point (Cen et al, 2013). By deliberately setting anchors, in a negotiation for example, the party who moves first can influence how people evaluate subsequent information.

2.1.5 Anchoring in Strategic Decision Making

In strategic business decisions, three different types of anchoring bias are common. First, initial estimates are often best guesses and are used without questioning their accuracy (e.g., guessing cost components of a capital investment project). Second, estimates are based on historical speculations, assuming trends will continue (e.g., predicting a company's sales target by drawing a straight line). Finally, some anchors are clearly deliberate (e.g., setting a low or high base in a price negotiation). The trap of anchors is that people often think they can disregard them, but in fact they cannot (Kahneman et al., 2011).

In recent experimental research, anchoring has been consistently observed in a wide range of contexts. These contexts include general knowledge (Epley & Gilovich, 2001), economic valuations (Alevy et al., 2015), stock predictions (Duclos, 2015), employee performance evaluation (Bellé et al., 2017), recommendations for employee promotion (Chen & Kemp, 2015), and negotiations (Orr & Guthrie, 2005). For example, Bellé et al. (2017) showed how individual performance appraisal is skewed toward performance ratings from prior years.

Although decision making is often unconscious and intuitive in nature, several authors have claimed that it is becoming increasingly data driven and evidence focused (Acciarini et al., 2020; Merendino et al. 2018). Therefore, the role of new technologies influencing business decisions will be discussed in the following chapter.

2.2 AI in Strategic Decision Making

2.2.1 The Rise of Big Data and AI

The amount of available data has increased tremendously in recent years, making it increasingly important for companies to make efficient use of this data (Gentsch, 2018). The main challenge is to combine the collected data in a meaningful way and identify relevant trends and patterns. The four characteristics of volume, velocity, variety, and veracity associated with Big Data cannot contribute to sound decision making unless appropriate algorithms are developed and applied (Cheng & Hackett, 2021). An algorithm refers to the procedure used for solving a problem and performing computations, following a defined set of instructions (Gillis, 2022).

This requires technologies such as Machine Learning (ML), which belong to the field of AI (Dhall et al., 2020). Kaplan and Haenlein (2019, p. 17) define AI as "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation". Big Data thus forms the basis for the use of AI, because the larger the amount of data, the more effectively AI can analyze, learn, and improve.

This is particularly important for organizations, as the modern world is often shaped by high levels of volatility, uncertainty, complexity, and ambiguity – in short VUCA (Bennet & Lemoine, 2014). This in turn, has given rise to the 4th Industrial Revolution, also referred to as Industry 4.0. AI is a core element of Industry 4.0 and, given the advance of digitalization and constant changes in organizations that are needed to maintain competitive advantage, AI applications are becoming increasingly pervasive in all industries (Kim & Kim 2020). Scaling AI can create a massive competitive advantage but investing in cutting-edge technologies is not enough (BCG, 2022). To achieve significant benefits from AI, decision making processes must be fully rewired. Thereby, a symbiotic relationship emerges, where companies deploy different human-machine interactions for different situations, adapting to changing contexts, circumstances, and scenarios (BCG, 2022).

2.2.2 AI in Strategic Decision Making

As stated earlier, humans do not always take rational decisions as they suffer from numerous cognitive biases. For nearly as long as such biases have been recorded, psychologists and others have tried to find ways to mitigate or eliminate the effects of decision making biases and heuristics (George et al, 2000). Due to its ability to take decisions in an effective and optimized manner, integrating AI may result in less biased decisions. The extent to which AI could be used to reduce cognitive bias in strategic decision making is one of the key benefits of its deployment in strategy. Indeed, a thorough overview of cognitive and motivational biases and debiasing strategies concludes that not only is bias pervasive, but also that AI indeed has the potential to eliminate it (Montibeller & Von Winterfeldt, 2015).

With the intention of overcoming the biases of human decision makers, organizations are increasingly using AI as decision support (Alon-Barkat & Busuioc, 2022). According to numerous studies, AI algorithms are being used in a variety of fields, including policing, welfare, criminal justice, healthcare, immigration, and education (e.g., Alon-Barkat & Busuioc, 2022; Eubanks, 2018). This demonstrates a rising and intensifying reliance on AI advice in business decisions. However, according to Lai et al. (2021), current research on human-AI decision making focuses mostly on decision efficacy and performance, and places less emphasis on whether people are willing to accept the advice of AI or not. To narrow this research gap, this study focuses on the human-AI collaboration and on factors influencing the decision maker's willingness to rely on AI advice.

2.2.3 Potentials and Limitations of AI Advice

Currently, AI is viewed more as a support for important business decisions than as a decision maker itself (Claudé & Comb, 2018). This is because AI in its current form is still relatively limited in capacity, compared to anticipated advances in the future. However, as computing power increases and as the amount of data available to support these decisions grows, this may not be a limitation for much longer (Stone et al., 2020). The underlying algorithms will also continue to improve and successively acquire capabilities that go beyond their human-created models (Montal & Reich, 2017).

According to Shrestha et al. (2019), there are currently three structural categories of human-AI decision making:

- Full human to AI delegation (e.g., dynamic pricing, recommender systems)
- Hybrid sequential decision making:
    - AI to human (e.g., idea evaluation, hiring)
    - Human to AI (e.g., sports analytics, health monitoring)
- Aggregated human-AI decision making (e.g., top management teams, boards)

This research focuses on one of the most common collaborative scenarios, AI-assisted decision making – a form of hybrid sequential decision making. In this case, algorithmic decisions serve as input to human decision making (Shrestha et al., 2019). Thus, humans are still in charge of the final choice. But, to support this decision, an AI algorithm offers a recommendation which the human decision maker can either accept or reject.

Thereby, AI holds the promise of reducing time, effort, and costs involved with resource intensive decision making (Parasuraman & Riley, 1997). Human-AI decision making allows people to follow the advice of AI when completing a task and thus work more efficiently (Rastogi et al., 2020). Despite their potential, such AI-based systems can have unforeseen and costly consequences for both users and businesses. Uncertainty, accuracy, and reliability issues can limit the utilization of such systems and become obvious when ML models are applied in real-world domains (Baird & Maruping 2021). Therefore, understanding human behavior and reliance on technology has become crucial since poor partnerships between people and automation can be "costly and catastrophic" (Lee & See, 2004, p. 50).

2.3 Reliance on Advice

2.3.1 Advice Taking in Decision Making and the Judge-Advisor System (JAS)

People seek advice from others for many important decisions in life. Whereas in the past advice was typically obtained from human experts, algorithmic advisors are becoming increasingly common nowadays. The phenomenon of people giving and taking advice has often been studied under the paradigm of the Judge-Advisor System (JAS) (Bonaccio & Dalal, 2006; Tauchert & Mesbah, 2019). This paradigm consists of two roles – the judge and the advisor. While the advisor merely provides some input to the decision, the judge holds the full power and is responsible for the final decision (Sniezek & Buckley, 1995).

Several factors influence the judge's reliance on the advice given to them, more specifically the weight that decision makers place on advice they receive. One of the main findings is that individuals often inaccurately discount advice when making decisions (Bonaccio & Dalal, 2006). Thus, individuals tend to assign more weight to their own opinion than to the opinion of others (Yaniv & Kleinberger, 2000). This has been attributed to concepts like differential information, anchoring, and egocentrism. Indeed, individuals have privileged access to their internal reasons for holding their opinions but not to the internal reasons of the advisors (Yaniv, 2004). Furthermore, the initial estimate of the decision maker serves as an anchor that is adjusted when receiving advice. However, this adjustment is usually insufficient and thus results in egocentric discounting (Harvey & Fischer, 1997; Gino & Moore, 2007). According to egocentric discounting, individuals favor their own opinions because they believe them to be superior to those of others, and thus tend to shift their initial estimate only up to 30 percent towards the advice (Yaniv & Kleinberger, 2000). Additionally, factors such as competence, power of advice, distance from advice, trust, and overconfidence have been found to influence the advice-taking behavior of individuals (Bonaccio & Dalal, 2006; Logg et al., 2018; Sniezek & Buckley, 1995; Tauchert, 2022; Van Swol & Sniezek, 2005). According to Mosier and Skitka (2018) people rely on decision aids when they trust them more than themselves. Therefore, the present research focuses on the role of trust between the judge and advisor in relation to the confidence of the judge in his or her own decision.

2.3.2 The Role of Trust in the Source

One of the most recognized attributes influencing the reliance on AI advice is trust (Tauchert & Mesbah, 2019). Trust is "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer et al. 1995, p. 712).

Mosier and Skitka (2018) hypothesize that people rely on automated decision aids when they believe that it is more reliable than their own opinion and thus place greater trust in it. However, when people distrust an automated decision aid, they are more likely to rely on themselves. Many other studies have found that trust plays a major role in understanding the adoption and use of AI advice (e.g., Lee & Moray, 1994; Muir & Moray, 1996). Considering the complexity and lack of control that are related to interactions with AI-based systems, the importance of trust in such interactions becomes clear. In an AI-assisted decision making setting, people might assume a favorable behavior of the AI although there is a possibility of the AI offering incorrect recommendations (Mayer et al., 1995).

On the other hand, Söllner et al. (2016) stress that a lack of trust in recommendations from AI can hinder successful adoption and deployment. According to prior research, there is a widespread distrust in AI advice. Algorithmic error has been identified as a key factor in explaining lower levels of trust in algorithmic advice (Dzindolet et al. 2002; Hoff & Bashir 2015). However, people also sometimes fail to recognize erroneous recommendations and follow incorrect algorithmic advice (Parasuraman & Manzey, 2010; Lee & See, 2004). This literature identified inappropriate trust as a reason for humans accepting or rejecting AI advice when they should not since their trust in AI does not match its trustworthiness – resulting in an over- or under-reliance on AI.

The conceptual framework by Lee and See (2004) provides a theoretical foundation to differentiate between trust as an attitude and reliance as a behavioral outcome (Schmitt et al 2021). Based on that, Dietvorst et al. (2015) considered trust as a mediator for behavioral outcomes. Similarly, Yin et al. (2019) measure both self-reported levels of trust and conformity to algorithmic advice. Following their example, this study distinguishes between trust in the source as a perceptual outcome and reliance on advice as a behavioral outcome variable. Prior research showed that trust enhances behavioral outcomes, whereby greater levels of trust lead to higher reliance on advice and vice-versa (Chua et al., 2022; Schaffer et al., 2015). Thus, the following hypothesis is proposed:

H1: *High levels of trust in the source increase the reliance on advice.*

Individuals tend to rely on automation that they trust and reject automation that they do not (Lee & See, 2004). Accordingly, reliance is guided by trust, but it is not entirely determined by it. Reliance on advice was found to be related to participants' perception of the advisor's ability relative to the perception of their own ability. Improving the reliability of the advisor will not increase the reliance on advice unless the advisor's perceived reliability surpasses that of one's own decisions (Dzindolet et al., 2002). Thus, besides trust, it is important to consider the confidence people have in their own decision.

2.3.3 The Role of Confidence in Oneself

Self-confidence is a critical factor in decision making in general (Bandura, 1982). When individuals have high confidence in their own decision and low trust in the decision of the advisor, they are more likely to rely on their own estimates and vice versa (Lee & Moray, 1994). While trust in AI (or humans) derives from the trustor's perception of the trustee's ability to perform a certain task, self-confidence influences the trustor's willingness to rely on the trustee (Chong et al., 2021). Individuals frequently report high confidence in their own judgment compared to the judgement of others (Logg et al., 2018). For example, Muir and Moray (1996) showed that individuals were more confident in their own abilities than in the ones from an automation, even though the automation performed at the same level. Biased levels of self-confidence, such as overconfidence, can have a significant impact on the appropriate reliance on automation (Lee & See, 2004). Thus, instead of solely focusing on trust in the source, this study also considers the role of confidence in oneself when it comes to accepting or rejecting advice – more specifically, the trade-off between the two variables. Therefore, the next hypothesis reads as follows:

> **H2***: The relationship of trust in the source and reliance on advice is moderated by the confidence individuals have in their own decision, such than high levels of confidence decrease the positive effect of trust on the reliance on advice.*

According to the framework proposed by Lee and See (2004), trust combined with other attitudes such as self-confidence form and determine the intention to rely on the automation. Furthermore, past studies raise the question of whether "individuals insufficiently trust algorithms (relative to human advisors) or merely overly trust themselves" (Logg et al., 2019, p. 91). Thus, following the example of Logg et al. (2019), I included a direct comparison of the reliance on advice coming from an algorithm versus a human.

2.3.4 Reliance on AI advice: Algorithm Aversion vs. Algorithm Appreciation

It is often found that people either rely too much or too little on AI advice (Van Dongen & Van Maanen, 2013). Thus, there have been opposing findings regarding the adoption and use of AI advice which can be classified into two streams of research – *algorithm aversion* and *algorithm appreciation*.

Algorithm aversion refers to findings suggesting that human decision makers tend to rely on human advice rather than advice from AI, even if such ML-based models have been shown to outperform humans in simple as well as in complex decision tasks (Castelo et al., 2019; Dietvorst et al., 2015). The psychological distrust in AI algorithms can be traced all the way back to 1954, when Meehl demonstrated the predictive superiority of algorithms over humans, but experts were skeptical that a linear model could exceed their own evaluations. This view has been supported by many other scholars (e.g., Hastie, 2001). Numerous reasons for the distrust in AI advice have been identified, including the error rate of the algorithm (Dzindolet et al., 2002), the difficulty of the task (Castelo et al., 2019), as well as the human confidence in their own reasoning (Whitecotton, 1996). Prior performance of the algorithm plays an important role in the acceptance of AI advice. Especially since people are more forgiving of their own or another person's mistakes than if this mistake comes from an algorithm (Dietvorst et al., 2015). This study specifically addresses the role of confidence in one's own decision. Furthermore, different mechanisms and design considerations have been suggested to overcome algorithm aversion. Enhancing the transparency and providing explanations on the functioning of an algorithm can improve the adoption of AI-based advice (Yeomans et al. 2019)

While most studies support the concept of algorithm aversion, a more recent stream of research shows a preference for AI advice. In various experiments conducted by Logg et al. (2019) they contradict the assumption of algorithmic aversion by demonstrating that people tend to follow advice more when they believe it comes from an AI algorithm than from a person. This phenomenon has been explained by the fact that people attribute more objectivity and rationality to AI as compared to humans (Beerbaum & Puaschunder, 2019). Therefore, it is not surprising that previous studies found a preference for AI advice especially in objective domains, such as those related to logic problems (Logg, 2017). Furthermore, Lee and See (2004) found that people sometimes erroneously rely on algorithms. In this case, the trust in AI exceeded the true capabilities of the AI-based system leading to an overreliance on AI advice. Some researchers warn of such overreliance and refer to this phenomenon as *automation bias*

(Goddard et al., 2012; Cummings, 2017). Thus, following the literature on algorithm appreciation, I hypothesize that:

H3: *Individuals show higher levels of trust in AI advice compared to human advice.*

H4: *Individuals rely more on AI advice compared to human advice.*

Together, these four hypotheses lead to a mediated moderation model (see Figure 1). The model aims at revealing insights about the effects of different sources of advice on the reliance on advice through trust in the source, whereby its strength depends on the confidence in one's own performance.

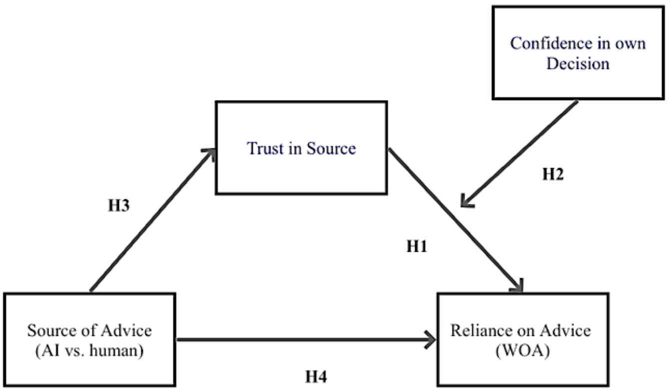

Figure 1: Conceptual Model 1

2.3.5 Self-Other Differences

Most of the decisions taken by a manager, do not only affect themselves but also other employees. Some literature suggests that people tend to choose differently for themselves vs. for others (Harkness et al., 1985, Kray & Gonzalez, 1999; Beisswanger et al., 2003). However, research also revealed that self–other differences differ across domains (Araujo et al., 2020). While such differences frequently appear in low-impact contexts, they are less present in high-impact contexts such as management decisions since they involve more risk (Stone & Allgaier, 2008). For AI-based solutions to be implemented, managers need to choose AI advice not only for themselves but also for others. Thus, it is important to understand whether self–other differences are prevalent in this context. Prior literature on automated decision making took into consideration both the impact of the decision as well as the subject of the decision (Araujo et al., 2020). Results showed that in high impact decisions, AI scored higher than human experts, irrespectively of who the subject was. Therefore, the last hypothesis reads as follows:

H5: *Individuals are more likely to choose AI advice over human advice, regardless of whether they choose for themselves or for others.*

While the first four hypotheses entail a separate evaluation of advice coming either from an algorithm or a human, the last one allows people to choose between the two sources. I thereby followed Logg et al. (2019) who asked individuals to pick between sources to see whether the findings are robust in a joint evaluation of alternatives (see also Hsee, C. K., 1996). This results in a second conceptual model illustrated below:

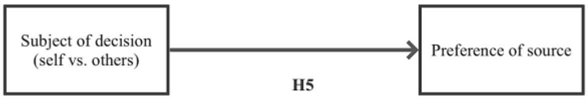

Figure 2: Conceptual Model 2

To test the five hypotheses, two studies were conducted. While Study 1 was designed to test the first conceptual model and its four hypotheses, Study 2 was designed to test the second conceptual model which includes H5. The methodological approach and the results for each of the studies will be described in the following two chapters.

3 Study 1

3.1 Research design

The aim of Study 1 was to test the causal effect of receiving advice from different sources (AI vs. human) on the final decision outcome. Therefore, an experimental study has been designed, since this is an eligible way to test for causality in hypothetical situations (Malhotra at al., 2017). Researchers frequently face a trade-off between internal validity and control versus external validity and realism when designing experiments (McDermott, 2011). In the present experiment, a controlled design with high internal validity was adopted. However, by using real life scenarios I aimed at making it as realistic as possible and thereby ensure a high level of external validity as well.

Building on previous research that studied the presence of cognitive biases in decision making as well as the reliance on advice, I used a quantitative approach to answer the research question(s). Therefore, an online study was designed with Qualtrics. The experiment consisted of two cells (human vs. AI), to better understand the influence of different sources on the reliance on advice. Participants were randomly and evenly assigned to one of the two groups. To compare participant's behavior between these two conditions while preventing knowledge and spillover effects, a between-subjects design was used (Charness et al., 2012). To further explain the causal effect, the variable trust in the source was included as a mediator and to assess a potential limit condition, confidence in one's own decision was included as a moderator

(MacKinnon, 2011). These two variables complete the research model, which ends up being a moderated mediation model.

Furthermore, a sequential decision making setup was adopted with two steps of human decision making (Friedel, 2014). First, the participant takes a decision regarding the task at hand and then receives advice from either an AI or a human. In a second step, the participant is given the opportunity to update the initial decision. Thus, the participant can either adopt the advice or stick to the previous answer. According to Schemmer et al. (2022), this allows measuring the reliance on advice in a fine-granular way. In total, this experiment included two scenarios that represent common business decisions.

3.2 Choice of decision making scenarios

Algorithms already help people to take better, less biased decisions in various domains (Lai et al., 2021). To come up with suitable scenarios, this study aimed at replicating two experiments that depict internal management practices in human resources and finance. In both domains anchoring bias has been identified as one of the most robust biases impacting decision making. Thus, the use of AI promises great potential to improve decision making.

In the first scenario participants were asked to provide feedback on a subordinate's performance in form of a performance rating ranging from 1 to 100 (Bellé et al., 2017). Such ratings are a common form of performance evaluation and are often subject to bias (Nagtegaal et al., 2020). Therefore, the performance of the employee was briefly described. Afterwards, participants were exposed to different anchors – high and low – that consisted of rankings from the previous year. Based on the results of Bellé et al. (2017), I expected participants in the high-anchor group to report significantly higher estimates than participants in the low-anchor group.

The second scenario focused on predicting the closing price of a company's stock based on the study from Duclos (2015). This task was chosen since financial decisions and planning play an important role in business but are often prone to biases (Athota et al., 2022). When processing financial information to forecast future trends, end-anchoring often led to significant investment asymmetries (Duclos, 2015). In the scenario at hand, participants were exposed to a graph showing the development of a stock's closing price over the last 30 days. However, one stock price was closing upward, the other downward. Based on previous findings, I expected participants to predict significantly higher (lower) forecasts, when a stock price was closing upward (downward).

To generate the advice for the performance rating task, average estimates from past studies were used – independently of the anchor condition to provide unbiased advice. For the stock price prediction, the mean of $60 was used since the graph was generated randomly around this value by the authors of the original study (Duclos, 2015). Using unbiased advice in both tasks further increases the validity of this study.

Although these two decision-making scenarios appeal to anchoring bias, it is important to mention that I did not expect differences between the high and low anchor groups regarding the reliance on advice. The reason I included them is to follow research's best practices on this topic, while testing for the robustness of the results. To assure there are no order effects between the first and the second tasks, I manipulated the order of the two tasks.

3.3 Sample and procedure

The required sample size was determined beforehand by running both a power analysis in G Power (at .80 power) for the moderation and a bias-corrected bootstrap test in PROCESS for the mediation. Effect size estimates were based on prior experiments assuming that both paths are medium sized, which resulted in a sample size of 148. Data was collected using a non-probability sampling technique; convenience sampling took place with voluntary participants that were recruited through social media (Malhotra et al., 2017).

After agreeing with the informed consent, participants were asked to answer demographic questions and rate their understanding of AI. Then, participants were exposed to the two decision scenarios described previously. After giving an estimate for each of the tasks, participants were asked to state how confident they were about their decision. In the next step, they received advice either from AI or from a human. In both conditions participants received the same advice, which allowed me to control for the quality of advice. Participants could then decide whether to accept this advice and adjust their previous estimate. Participants in the AI condition were given a short explanation of AI, ML, and the functioning of the advice algorithm, as this is a very recent field and not having further information might have influenced the trust component independently of the manipulation. The information given was based on the work of Sarker et al. (2018) for the first algorithm, and Ravikumar et al. (2020) for the second.

Between October 31st and November 9th, a total of 217 surveys were completed. However, from these surveys 12 were excluded from the analysis because participants failed one of the attention checks, which led to a total valid sample of 205 participants (48% male, 48% female, 4% other). For more details on the attention checks, see Appendix 1. Their age ranged from 19 to 78 years (M = 36.93, SD = 13.27) and most of the respondents had a Bachelor's or a Master's degree (N = 155). Furthermore, most respondents were either employed or worked as freelancer at the time of the survey (N = 140) and had a European nationality (N = 168). On average, participants rated themselves as having a slightly good understanding of AI (M = 4.41, SD = 1.24) and a slightly good pre-knowledge regarding the two tasks (M = 4.49, SD = 1.31). For more details on the population statistics, see Appendix 3.

3.4 Variable measurement

In the following, the concrete measurements, paradigms, and items for each variable will be described. Following Hinshaw (2007), the moderator as well as the covariates were measured before the intervention to ensure that the moderator is uncorrelated with the treatment assignment and there is no systematic bias in interpreting this interaction. The mediator, on the other hand, was measured during the active intervention to uncover the processes through which interventions exert their effects (Hinshaw, 2007). For more details on all variables, see Appendix 1.

3.4.1 Main variables

Source of advice (AI vs. human): The independent variable in this experiment was a categorical variable that represented two different sources of advice. In each of the two decision scenarios, participants either received advice from AI (experimental condition) or a human (control condition) – with the advice in both conditions being equal. Thus, differences resulting from the manipulation should not be attributed to the advice accuracy, but rather to the influence of the source (Logg et al., 2019).

Reliance on advice: The dependent variable looked at how much weight participants assigned to the different sources. To measure participants reliance on advice, the JAS paradigm was employed which requires participants to make a first decision under uncertainty, and then receive advice before making a second, perhaps revised decision (Sniezek & Buckley, 1995). The difference between the first and revised decision divided by the difference between the first decision and the advice represents the dependent variable, the Weight on Advice (WOA). Participants have a WOA of 0% when the final estimate is equal to the initial estimate, and a WOA of 100% when the final estimate is equal to the received advice. Intermediate values suggest that participants weight both their own initial estimate and the advice received.

Confidence in oneself: To measure participants confidence in their first estimate, a modification of Dzindolet et al. (2003) items was used. Minor changes were made to adapt the items to the hypothetical situation at hand. An example item reads *"I am confident in my answer"* (1 = *Strongly disagree;* 7 = *Strongly agree)*. All three items were measured using a seven-point Likert scale, which allowed participants to be indifferent between the two descriptors.

Trust in Source: The level of trust was measured using a mix of items from Jamaludin and Ahmad (2013) and Gold et al. (2015). An example item reads *"I believe this recommendation is trustworthy"* (1 = *Strongly disagree;* 7 = *Strongly agree)*. Like confidence, trust was measured with three items using a seven-point Likert scale.

3.4.2 Covariates

Understanding of AI: The familiarity with and perception of AI has been identified as important factor influencing the reliance on advice, which is why I included it as a covariate in the model (Alon-Barkat & Busuioc, 2022). To measure participant's understanding of AI, a scale from De Vries et al. (2021) was adapted. For consistency, a seven-point Likert scale was used (1 = *Strongly disagree;* 7 = *Strongly agree)*.

Distance from advice: Following the example Logg et al. (2019), I also included the distance from advice as covariate to my model and measured the variable by calculating the difference between the first estimate and the advice received.

Pre-knowledge: A pretest revealed that the previous knowledge regarding the two tasks also played a role in the decision to rely on advice. Therefore, participants were asked to rate their pre-knowledge on a seven-point Likert scale (1 = *Extremely bad;* 7 = *Extremely good)*.

Demographics: Following previous literature, participants' gender, age, nationality, educational background, and current employment status were also included as covariates (e.g., Chua et al., 2022; Hou & Jung, 2021). Gender was captured as either male, female, or other. Age was captured in years. The remaining three variables were measured in a single choice format using the options provided by Qualtrics. Minor adjustments were made, after a pretest revealed missing options (see Appendix 3).

3.5 Data preparation and scale reliability

Before conducting the analysis, I ran all necessary tests that ensure a valid analysis which can be found in Appendices 4-8. As no significant difference between the two decision making tasks was found, the two tasks were combined meaning that the respective variables were aggregated by their means. Furthermore, I tested whether the different anchors impacted the dependent variable. As expected, results showed that there were no significant differences between the high (upward) and low (downward) anchor groups, which is why I aggregated across them. As mentioned earlier, the order of the tasks was randomized as well. Since this randomization had no impact on the participants' answers, the following analyses will also not distinguish between the different orders. Finally, a Fisher's Exact Test of Independence was used to examine whether the manipulations of the source of advice worked as intended (Freeman & Halton, 1951). The results showed that there was a significant dependence between relative proportions of the different sources of advice and the responses to the manipulation check ($p < .001$). Thus, the null hypothesis of independence can be rejected, meaning that the manipulation was successful as most participants recognized the respective agents and answered correctly ($N = 135$).

Although the scales used in this experiment have proved reliable in past studies (Chua et al., 2022; Dzindolet et al., 2003), a reliability analysis was conducted to test for Cronbach's alpha. The scale for confidence in oneself as well as the scale for trust in the source both yielded a Cronbach's alpha of 0.96 which is considered as excellent according to Gliem and Gliem (2003). Thus, both constructs showed high internal consistency between the items and are reliable enough to predict the respective variables.

3.6 Testing for Anchoring

To emphasize the presence of cognitive biases in strategic decision making, this experiment was designed to replicate two studies with anchoring paradigms. Therefore, participants were exposed to different anchors in each of the two tasks – high vs. low in the first task (T1) and upward vs. downward in the second task (T2). To measure the anchoring effect, an independent samples t-test comparing the two groups was used (see Appendix 9). The results show that the replication was successful since there was a significant difference between the estimates of the two groups ($t(203) = 14.13$, $p < .001$ for T1; $t(203) = 13.09$, $p < .001$ for T2). For descriptive information, see Table 1.

	Anchor	*N*	*M*	*SD*
Task 1	High anchor	101	82.02	10.34
	Low anchor	104	61.29	10.66
Task 2	Upward anchor	103	72.99	14.74
	Downward anchor	102	49.00	11.26

Table 1: Anchoring Effect

3.7 Hypotheses testing

To test the four hypotheses of the first conceptual model, a statistical analysis was conducted using Hayes' *PROCESS* macro for SPSS. The macro is based on regression-path analyses to reveal moderation and mediation effects using a bootstrapping approach (Hayes, 2018). As the hypotheses of this study aimed to test a moderated mediation effect, with the moderation happening on path b and not on the direct effect, the appropriate model for the analysis is number 14 (see Figure 3).

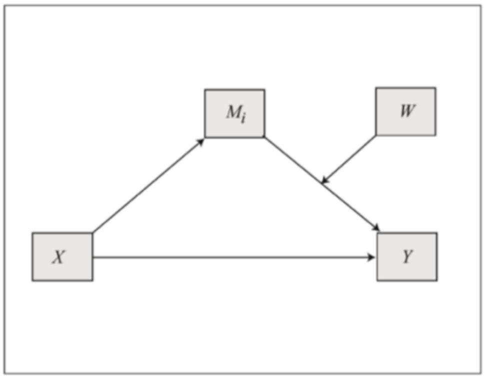

Figure 3: Model 14 of Hayes PROCESS macro for SPSS

For the following hypotheses tests, a 5% significance level with 5,000 bootstrap replications was chosen. The conducted tests include simple and multiple linear regressions. While traditional procedures require all regression assumptions to be met, the bootstrapping approach used by this macro takes the original sample data and then resamples it to generate a large number of simulated samples. Because the sample distribution can be observed and this approach does not rely on theory, the assumptions become obsolete (Joseph, 2020).

The chosen model involves testing for the indirect effect of X on Y via the proposed mediator M, whereby the indirect effect is influenced by the moderator W. We are talking about a second-stage moderation as the moderation takes place at the second path. To test whether the effects are significant, two regression sub-models were conducted. It is important to mention that the source of advice has been coded as a dummy variable with 1 representing AI advice and 0 representing human advice. Furthermore, all covariates described in Chapter 3.5.5 were included in the analysis.

The first sub-model entailed regressing the M onto X and showed a positive and significant effect of the source of advice (Human vs. AI) on the trust in the source ($t(9, 192) = 4.31$, $b = 0.71$, $p < .001$). Thus, the relation between the independent variable and the mediator is positive and significant, which supports **H3,** such that participants on average trusted advice from an AI more than advice from a human. None of the covariates showed a significant effect. The overall sub-model is significant and explains 10% of the variance (R^2 = .12, F(9,192) = 2.94, < .001).

The second sub-model entailed regressing Y onto X, M, W, and the interaction term MW, which captured the moderating effect of confidence. When looking at the coefficients of the predictor variables, the results showed that both trust in the source and confidence in oneself were significant. While on average trust increased the weight participants placed on advice (t(12, 189) = 10.17, b .17, p < .001), confidence decreased it (t(12, 189) = 3.85, b = -.06, p < .001). **H1** is thereby supported, such that high levels of trust in the source increase the reliance on advice. Furthermore, the interaction term was also significant with a negative sign and thus served as evidence of a moderation effect t(12, 189) = 3.97, b = -.05, p < .001). More specifically, this means that the effect of trust on the weight on advice varies across levels of confidence (i.e., the effect is not constant). The higher the confidence, the lower the weight on advice, despite higher levels of trust, thereby supporting **H2.** This described relationship can be seen in Figure 4. With higher levels of confidence, the slopes for the effect of trust on WOA become less positive at -1sd (blue line: b = .23), at the mean (green line: b = .17) and at +1sd (red line: b = .11).

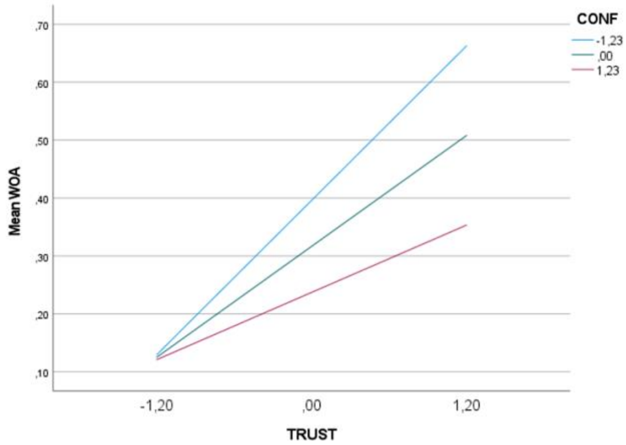

Figure 4: Moderation Effect of Confidence

However, when looking at the relationship between X and Y, the model suggests the absence of a direct effect of the source of advice on the reliance on advice (t(12, 189) = .55, b = .02, p = .58). Thus, the path between X and Y is not significant and **H4** is not supported, which indicates that this effect is fully mediated through the trust in the source. In addition, I looked at the Index of Moderated Mediation (IMM), which quantifies the degree to which the indirect effect in the model is moderated. It can be treated as an omnibus test for moderated mediation using bootstrap confidence intervals. Since 0 falls outside the lower and upper interval bounds, the effect is statistically significant, which confirms the case of moderated mediation.

Whereas in the first model, none of the covariates were significant, in the second model some of them were. Participants' understanding of AI, distance from advice, and their pre-knowledge showed a significant effect. While greater distance from the advice ($t(12, 189) = 3.36$, $b = .01$, $p < .001$) and pre-knowledge regarding the tasks ($t(12, 189) = 2.85$, $b = .05$, $p < .001$) positively influenced participants' reliance on advice, greater understanding of AI showed the opposite effect ($t(12, 189) = -2.15$, $b = -.03$, $p = .03$). The overall second model is significant and explains 52% of the variance ($R^2 = 0.52$, $F(12, 189) = 16.75$, $p < .001$). A correlation matrix including all variables support the findings described above and can be found in Appendix 11. For the PROCESS output, see Appendix 10.

4 Study 2

4.1 Research design

The aim of Study 2 was to test the robustness of the findings from Study 1, as well as uncover the reasons behind them. Therefore, another experimental study was designed, in which participants were asked to choose between the two sources of advice and state the reasons behind their decision. Furthermore, it was examined whether participants chose differently for themselves vs. for others by manipulating the subject of the decision, to test for a possible limit condition of the findings.

Therefore, a second online study was designed with Qualtrics, in which participants were randomly and evenly assigned to one of the two groups (self vs. others). To compare participant's behavior a between-subjects design was adopted (Charness et al., 2012). Building on the first study, the same two decision making scenarios were used, with the order of the two scenarios being manipulated to account for order effects.

While descriptive models are well suited to determine whether the outcome variable differs systematically across experimental circumstances, they are poorly suited to determine the underlying mechanisms of this effect (Bellé et al., 2018). Thus, to uncover the reasons behind the respondents' decisions, the format of open-ended questions was used. Open-ended questions have the advantage of not forcing respondents into one of the provided categories, but rather allow them to qualify and quantify their answers (Emde, 2014). In this way, new insights can be obtained. Thus, to strengthen and extend the findings of Study 1, I followed a hybrid approach in Study 2 using both quantitative and qualitative methods.

4.2 Sample and procedure

The sample was determined in advance by running a power analysis in G Power (at .80 power) assuming a medium sized effect. This resulted in a sample size of 128 participants. Data was gathered via the Laboratory of Experimental Research in Economics and Management. All participants were Master students at a European business school who participated in exchange for course credit. Therefore, this research center was especially useful as business students will be affected the most by the trend towards AI-based solutions and thus, were the main target group of this research.

The procedure followed a similar structure as Study 1. After agreeing with the informed consent, participants were asked to answer demographic questions and rate their understanding of AI. Then, they were exposed to the two decision scenarios. To help participants with their decision, they were offered advice from another person and from an algorithm. However, in one condition they were asked to pick the preferred source for themselves and in the other condition, they had to pick for a colleague. In both conditions they received a short description of the two sources, according to which they performed equally well in the past. Thus, participants' preferences were based only on the conclusions they drew about the two sources of advice. In the last step, participants were asked to state the main reason behind their choice. They went through this procedure for both the HR and the finance task.

Between November 22nd and 30th, a total of 239 completed surveys were completed, of which 34 surveys were excluded from the analysis as participants failed the attention check, that asked participants to select "*Strongly agree*". This led to a final sample of 205 Master students (53.7% female, 46.3% male), aged between 19 and 30 years (M = 23.43; SD = 1.85). On average, participants rated themselves as having a slightly good understanding of AI (M = 8.82, SD = 1.55). For more details, see Appendix 13.

4.3 Variable measurement

In the following, the concrete measurements and constructs for each variable of Study 2 will be described. For details on the variables used in this Study, see Appendix 2.

4.3.1 Variables

Subject of decision (self vs. others): As in Study 1, the independent variable was categorical representing two conditions. In the experiment, participants either decided for themselves (control condition) or for others (experimental condition) which source of advice they pick –

everything else being equal. Participants were given the same information on AI, ML, and the algorithm as in the previous study.

Preference of source: The dependent variable examined which source of advice participants preferred. To measure their preference, a seven-point Likert scale was used (1 = *Definitely the colleague;* 7 = *Definitely the algorithm*. This format was chosen for consistency with Study 1 and because it allowed participants to be indifferent between the two options.

4.3.2 Reasons

Participants were asked to state the main reason for their decision in an open-ended question. For the analysis of the responses, data was coded following Grounded Theory (Corbin & Strauss, 1990). In order to be open to new findings, the influencing factors were derived inductively (not deductively) from the collected data. A common approach to analyzing qualitative data is thematic analysis (Braun & Clarke, 2006). This method is used for identifying, analyzing, and capturing different themes within the data. Therefore, each response was assigned one or more codes. Over time, recurring patterns in the codes were identified, and codes were assigned to broader themes (see Appendix 18). The analysis was done iteratively until all responses were assigned to the different themes, which will be explained in more detail in Chapter 4.6.

4.4 Data preparation

As in Study 1, I tested whether the effect of the manipulation (self vs. others) differed between the two decision-making scenarios. However, this time a significant difference between the two scenarios was found regarding the effect of the manipulation on the preference of advice. Thus, for the following analysis, the two scenarios will *not* be aggregated but will be treated separately. Furthermore, no significant differences between the two order groups could be identified in either of the two tasks. Thus, the following analysis will not distinguish between different orders. For test results, see Appendices 14 and 15.

4.5 Hypotheses testing

To test whether participants chose differently for themselves vs. for others, I ran an independent samples t-test for each of the two tasks. As anticipated, results were not significant which suggests that, on average, participants chose the same sources of advice for others as for themselves ($t(203) = .93$, $p = .25$ for T1, ($t(202) = .64$, $p = .62$ for T2). To analyze participants preferred source of advice, a one-sided t-test against the midpoint value of the scale (= 4) was conducted for each of the two tasks. Results showed a significant preference of algorithmic advice for the stock price prediction task ($t(203) = 11.29$, $p < .001$). However, they showed a significant preference for human advice for the performance evaluation task ($t(204) = -3.36$, $p < .001$). **H5** is thereby only supported for the second task. For more details on the two tests, see Appendices 16 and 17.

Both the repeated measures ANOVA as well as a paired t-test revealed significant differences between the two tasks in Study 2 (see Appendix 14). When comparing these results to the results of the first study, the data suggests that participants trusted algorithms when they received advice from them (Study 1) but when asked to choose between the two sources, participants only acted in line with their trust towards algorithms for the finance task (Study 2). The reasons behind their respective choices will become clearer in the following chapter.

4.6 Qualitative analysis

While Study 1 examined the effects of trust and self-confidence on the reliance of AI advice, Study 2 shed light on additional factors that seem to influence this decision. The results of the qualitative analysis reinforce the importance of trust, which was found to be the key factor behind participants' decisions. While some participants expressed strong opinions either for the AI or the human, others stated that their decision was based on gut-feeling. However, they mentioned different advantages and concerns regarding each of the two sources.

One of the main reasons why participants chose the AI was their belief that it would perform better compared to humans. Performance included the ability of AI to complete the tasks in a consistent, reliable, and accurate manner drawing on a large amount of data and information. Thereby, the past performance and error rate of the algorithm mattered. Furthermore, participants considered the output more objective and less biased, which was an important factor as well since many participants were aware of human decision-making biases. To avoid bias in algorithmic decision-making, participants were particularly interested in the process of the output generation (i.e., how the output AI was trained, and which data was used). They also mentioned greater efficiency as a reason, since AI advice is faster, cost-saving, and allows for a better resource allocation. Finally, some of them chose the AI because they believed it to be the future.

However, participants also stressed the importance of emotional intelligence, which they assigned to the human advisor. Accordingly, some tasks require empathy, feelings, and "a human touch". They expressed concerns that AI models may not be able to consider special circumstances or assess personal and interpersonal characteristics that are not reflected in historical data. For some participants, the possibility of contacting the advisor in cases of doubt or mistakes was an important factor, which they felt is not given with an AI advisor. Furthermore, some participants even viewed AI as a threat, mentioning that jobs are being replaced by AI and humans are no longer in control.

As the quantitative analysis showed, there was a significant difference between the two tasks. This is supported by the qualitative analysis, as many participants stated that they prefer algorithmic advice for numbers and data driven tasks (i.e., prediction tasks) and human advice for people related tasks involving (inter)personal components (i.e., performance evaluations). Furthermore, many participants pledged for a combination of both sources, whereby the human should still oversee the final decision but use the AI as assistance. For an overview of the factors mentioned above, see Table 2.

REASONS FOR CHOOSING THE AI ADVICE	REASONS FOR CHOOSING THE HUMAN ADVICE
• General trust towards algorithms • Performance • Objectivity and rationality • Transparency regarding the process • Efficiency • Future oriented mindset • Assessment of quantifiable data • Domain: Numbers, Mathematic question, Data driven, Calculations, Logic problems, Prediction tasks	• General trust towards humans • Emotional intelligence of human advisor • Direct point of contact • No previous knowledge of AI • Fear of job replacement • Assessment of non-quantifiable data • Domain: People, Personal question, Interpersonal skills, Employee evaluations

Table 2: Overview of Themes

5 Discussion

Spanning a variety of functions, AI-based solutions are becoming increasingly common in organizations. Most importantly, it has become relevant for decision making in organizations, which was the main motivator for this research. However, to unlock the potential of AI to improve decisions outcomes, users must be willing to rely on it in the first place. To better understand people's readiness to accept advice from AI, two experimental studies were performed in the scope of this research. Thereby, five hypotheses were tested, each of which corresponded to a research sub-question. In the following, the results will be discussed in more detail and practical and theoretical implications will be derived.

5.1 Research Findings

5.1.1 Study 1

The first experiment tested the moderated mediation model. The goal was to understand how receiving advice from AI or from a human affected people's trust and reliance on it. The results supported **H1**, which predicted that high levels of trust in the source increase the reliance on advice. These results are consistent with the large body of previous literature suggesting that greater levels of trust lead to a higher reliance on advice and vice-versa (e.g., Chua et al., 2022; Schaffer et al., 2015). In fact, the results show no direct effect between the source of advice on the reliance of advice as **H4** is not supported, which suggest that the entire effect is indirectly transmitted through the mediator, trust in the source. This emphasizes the importance of trust for a successful adoption and deployment of AI advice.

Although trust is a crucial determinant of the reliance on advice, it is not the only one. In fact, it is closely related to how confident people are in themselves. Indeed, **H2** was supported, according to which the relationship between trust and reliance on advice is moderated by the confidence that individuals have in their own decision. High levels of confidence in their own decision and low levels of trust in the decision of the advisor led people to rely more on their own estimates and vice versa, which is in line with previous literature (e.g., Lee & Moray, 1994). Since people tend to be overly confident, they may not rely on advice although it would be reasonable and potentially lead to a better decision **(**Malmendier & Tate, 2005). While a healthy amount of self-confidence is good, it is important to be able to properly assess one's own abilities and to know when to rely on advice. But whose advice do individuals trust more?

In contrast to the widely held belief that people tend to distrust algorithms (see Chapter 2.5), the results of this study indicate that individuals show higher levels of trust in AI advice compared to human advice. Thus, **H3** was also supported. When facing identical advice, participants trusted advice more when it came from an algorithm than from human. This was the case when evaluating an employee's performance and when predicting the closing price of a stock. Interestingly, participants displayed algorithm appreciation regardless of their demographics which is in line with the recent findings from Logg et al. (2019). What mattered however, was the understanding of AI and thus, the familiarity with algorithms as well as the distance from advice (see also Alon-Barkat & Busuioc, 2022; Minson et al., 2011). The greater the distance, the greater the extent to which participants adjusted their initial estimate towards the advice. Finally, participants' previous knowledge regarding the two tasks showed a significant effect as well. Unexpectedly, greater knowledge led to a stronger reliance on advice which contradicts the findings of Sniezek and Van Swol (2001). A potential reason for this discrepancy might be that people with greater knowledge understand that these tasks are prone to biases and thus are more open to rely on advice.

5.1.2 Study 2

The second experiment tested whether algorithm appreciation could also be found in a joint evaluation of alternatives and examined whether people choose differently for themselves vs. others. The results show that when individuals are presented with both options, they are more likely to choose AI advice for the stock prediction and human advice for the performance evaluation of an employee – regardless of who the decision is made for. This serves as validation for **H5** in the second scenario, and a rejection of **H5** in the first. These findings are in line with previous literature suggesting that people prefer different sources of advice for different domains. Logg (2017) found that people agreed more with advice coming from AI than from a human in objective domains. However, in subjective domains people preferred human advice (Logg, 2017; Yeomans et al., 2019). Including two different scenarios allowed me to understand whether there are different preferences depending on the domain or if the preferences are robust across both. While in Study 1 algorithm appreciation was prevalent in both scenarios, in Study 2 it was only prevalent in the second, more objective scenario for the reasons outlined in the previous chapter. Perhaps another reason for this discrepancy has to do with the different paradigms used in the two studies. Providing individuals with a direct alternative to the AI advice might have reduced the reliance on it, as information on the human alternative became more available (Bazerman et al., 1992; Hsee, 1996). Thus, people were willing to trust AI advice when they received it but when it came to choosing between AI or human advice, had different preferences depending on the task at hand. These two different paradigms mimic the real-world case in which managers need to choose AI advice at the point of the implementation decision (Study 2) and that employees need to use the AI advice afterwards (Study 1).

Moreover, individuals demonstrated the same preferences when deciding for others. According to Stone and Allgaier (2008), self–other differences are typically found in low-impact contexts which suggests that participants considered the two decision scenarios as high-impact contexts. It appears that individuals investigate their options more carefully in high-impact scenarios, regardless of who the decision is made for. Thus, further aspects regarding the two sources were considered in the decision-making process and explain why individuals did not take different, potentially riskier decisions for others (Wray & Stone, 2005).

5.2 Theoretical implications

The increasing number of studies on AI advice in recent years illustrates the growing relevance of the topic. However, there has been little research focusing on the trade-off between trust in algorithms and confidence in oneself. Building on the previous literature and the research gaps it reveals, the reliance on such advice was examined more closely, offering several theoretical implications.

First, the findings of the two studies question the widespread belief of algorithm aversion (e.g., Castelo et al., 2019, Dietvorst et al., 2015) and suggest that the story is not as straightforward. Especially Study 1 contributes to the comparatively smaller, more recent stream of research on algorithm appreciation (Logg et al., 2019). Results showed that providing AI advice increases the compliance with advice. Study 2, in turn, suggests that algorithm appreciation and aversion may be domain specific, as people preferred human advice for rather subjective and AI advice for rather objective domains. Thus, people seem to be more sensitive to the decision domain when being asked to choose between the two sources of advice than when being in a situation in which they might be influenced by them. To guide organizations towards a successful implementation of AI-based systems, various influencing factors have been identified, complementing existing literature.

Second, two previous experiments on anchoring bias were successfully replicated, which empirically generalizes and reinforces the confidence in the original findings (Bellé et al., 2017; Nagtegaal et al., 2020). The present findings showed even stronger effects than those of the original study, emphasizing the robustness of anchoring bias in strategic decision making. As AI has the potential to provide unbiased results, this research also contributes to the literature looking for ways to overcome decision making biases, in particular anchoring bias (e.g., Montibeller & Von Winterfeldt, 2015). It does so by suggesting that people generally trust AI and that this source of advice (AI) works better than the traditional one (human).

Finally, by using the JAS paradigm, this research connects the literature on AI advice with the literature on advice taking in general. Accordingly, people should average their own estimates with that of others to maximize accuracy, resulting in a WOA of 0.5 (Soll & Larrick, 2009). But since people tend to discount advice, the WOA is usually too low. By showing that people underweighted advice, this study supports previous research on advice taking. Therefore, the results show similar effects to the literature (see Appendix 12). Thus, although people weighted AI advice more than human advice, they did not adjust their estimates enough, especially since the advice given was correct. This implies that, despite the positive findings on algorithm appreciation, there is still room for improvement (Logg et al., 2019).

5.3 Managerial implications

In addition to the theoretical contributions, this research also offers managerial implications. These are particularly relevant for any organization that wishes to improve decision making with the aid of Big Data, following success cases such as Amazon and Google. As they invest in the exploitation of ever-increasing amounts of data, algorithms are used to analyze the data and offer guidance for business decisions. As technology advances, the speed and accuracy of algorithms continue to increase, and with it the potential benefits of AI advice. The sample of the two conducted experiments mainly consisted of business students, exactly those who will be affected the most by the trend towards AI-based solutions and will have the means to affect others using these solutions. Indeed, some of them may even become managers responsible for the implementation of such systems. They will not only choose for themselves but also for others whether to be influenced by AI or not. Indeed, findings suggest that future managers will choose equally for themselves and for others and will enable AI-based decision making primarily in objective domains. Regarding the adoption of AI advice, managers should be encouraged by the results, as they show that individuals are more likely to improve their decisions by listening to algorithms.

However, a successful implementation of such systems strongly depends on people's willingness to rely on them. Concrete suggestions for the implementation and expansion of AI advice can be derived from the factors that influenced participants' decisions in Study 2. They can serve as orientation for companies and managers to counter employees' concerns and increase their acceptance. Despite the many benefits, especially in terms of cost and efficiency, companies should, above all, ensure users' trust in AI-based systems. Trust strongly depends on the performance of the algorithms, which emphasizes the importance of keeping the error rates as low as possible, especially since people are less forgiving of mistakes made by algorithms (Dietvorst et al., 2015). In this context, a close control of the output is crucial to ensure that high-quality advice is generated. Furthermore, participants expressed mixed thoughts about the lack of emotions and feelings of AI. Some participants thought of this as a positive factor as the emotions did not bias decisions. Other participants, however, were concerned and felt that AI should not be used for assessing human behavior and skills as emotions are involved. This suggests that AI advice should primarily be implemented in objective and data-driven domains. Results also showed that keeping the process transparent increases people's trust in the source. This means disclosing the functioning of the algorithm and the data selection to reduce concerns about possible bias (see also Bonaccio & Dalal, 2006). Although most AI-based systems are designed to complete tasks autonomously, we will most likely witness a human-AI partnership in the short term before AI becomes more autonomous. Participants embraced such a partnership and frequently indicated that a mix of the two sources would lead to the most-trusted decision outcome. Such an interaction combines the emotional intelligence of humans and the machine intelligence of algorithms and thus optimizes the results. Finally, people are aware that AI can outperform humans in certain tasks and eventually replace them. Concerns about AI seizing jobs and replacing employees impacted participants' decisions regarding their preferred source of advice. Given that AI will continue to improve and penetrate more and more areas in the future, it becomes critical to provide retraining and training to employees that are affected by this change to limit this effect (Siau & Wang, 2018).

5.4 Limitations and future research

While this research has important implications, there are certain limitations, especially with respect to the generalization of the results. First, the data was gathered using a non-probability sampling technique. Due to time and financial constraints this technique was reasonable, however it led to a non-representative sample. Specifically, most of the participants had tertiary education and originated from Europe, which is why this study should be replicated with a larger, and more representative sample to improve the reliability of results.

Second, although I sought to increase external validity by using real work scenarios, the setup was still artificial. As a result, answering to the questions required participants to imagine how they would react in a real-life situation. Even though this is a common and widely used predictor of the actual behavior, discrepancies may exist, especially regarding participants' involvement in the task (Ajzen & Fishbein, 1980). Also, the sequential decision-making setup used in this study is not necessarily representative of all real-world situations and may not always be practicable. Thus, future studies should repeat this experiment in a real-world setting that more closely resembles the overall decision making context.

The biggest limitation of this study, however, is that it only included two decision scenarios, limiting the extent to which one can talk about an overall reliance on AI advice. Although the tasks represent different domains – human resources and finance –, there exist additional fields of practice in which AI could be used to improve decision making. Indeed, some literature suggests that people prefer different sources of advice for different tasks. Therefore, this research should be extended, including a variety of objective and subjective tasks, to understand if and how this affects employees' reliance on AI advice.

While this study focused on the trade-off between trust and confidence, future studies could expand the model by adding further variables suggested by literature, such as competence, power of advice, and transparency (Bonaccio & Dalal, 2006; Van Swol & Sniezek, 2005). For example, they could examine how different levels of transparency impact the reliance on AI advice, including black box algorithms which might be more representative of AI applications in people's everyday lives (e.g., weather forecast). The qualitative results of Study 2 suggest that this may well be the case. Nevertheless, it remains an empirical question for which I hope this study serves as a basis. Whether results hold for other biases is another empirical question that would be interesting to explore. Such studies would allow to increase the generalizability of the results and simultaneously stress the potential of AI to reduce cognitive biases in decision making.

This research focused on AI-assisted decision making, with humans still overseeing the final decision. Since algorithms will continue to improve and successively acquire capabilities that go beyond their human-created models, it would be interesting to examine people's attitudes towards a full human to AI delegation (Shrestha et al., 2019). One main reason for people's distrust in AI advice was the lack of emotional intelligence required to properly assess human behavior. However, a new concept referred to as *Artificial Emotional Intelligence* is already being used to develop systems that are capable of recognizing and mimicking human emotions (Kaur & Sharma, 2021). This could be the basis for alleviating this concern. Thus, a potential follow-up study could investigate how AI agents with varying levels of emotional intelligence influence people's attitudes and decisions.

6 Conclusion

Technology is increasingly being used to deliver effective and low-cost advice to support humans in making decisions, thereby substituting human advisors. The current thesis reveals that individuals generally trust algorithmic advisors more than human advisors and hence seek advice from AI – thereby suggesting a new way to overcome cognitive biases in strategic decision making. Given the constantly improving quality of algorithmic advice, which is not subject to cognitive biases, the message for organizations becomes clear: It's time to get the algorithms in.

References

Acciarini, C., Brunetta, F., & Boccardelli, P. (2020). Cognitive biases and decision-making strategies in times of change: a systematic literature review. *Management Decision*. *59*(3), 638-652.

Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Alevy, J. E., Landry, C. E., & List, J. A. (2015). Field experiments on the anchoring of economic valuations. *Economic Inquiry*, *53*(3), 1522-1538.

Alon-Barkat, S., & Busuioc, M. (2022). Human-AI Interactions in Public Sector Decision-Making:'Automation Bias' and 'Selective Adherence'to Algorithmic Advice.

Araujo, T., Helberger, N., Kruikemeier, S., & De Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY*, *35*(3), 611-623.

Artificial Intelligence. BCG Global. (2022). Retrieved October 11, 2022, from https://www.bcg.com/capabilities/digital-technology-data/artificial-intelligence.

Athota, V. S., Pereira, V., Hasan, Z., Vaz, D., Laker, B., & Reppas, D. (2022). Overcoming financial planners' cognitive biases through digitalization: A qualitative study. *Journal of Business Research*, *154*, Article 113291.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*(2), 122-147.

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, *45*(1), 315-341.

Barnes Jr, J. H. (1984). Cognitive biases and their impact on strategic planning. *Strategic Management Journal*, 5(2), 129-137.

Bazerman, M. H., Loewenstein, G. F., & White, S. B. (1992). Reversals of preference in allocation decisions: Judging an alternative versus choosing among alternatives. *Administrative Science Quarterly*, 220-240.

Beerbaum, D., & Puaschunder, J. M. (2019). A Behavioral Approach to Irrational Exuberances–An Artificial Intelligence Roboethics Taxonomy. *Scientia Moralitas-International Journal of Multidisciplinary Research*, 4(1), 1-30.

Bellé, N., Cantarelli, P., & Belardinelli, P. (2018). Prospect theory goes public: Experimental evidence on cognitive biases in public policy and management decisions. *Public Administration Review*, 78(6), 828-840.

Beisswanger, A. H., Stone, E. R., Hupp, J. M., & Allgaier, L. (2003). Risk taking in relationships: Differences in deciding for oneself versus for a friend. *Basic and Applied Social Psychology*, 25(2), 121-135.

Bennett, N., & Lemoine, J. (2014). What VUCA really means for you. *Harvard Business Review*, *92*(1/2).

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, *101*(2), 127-151.

Boone, L. E., Kurtz, D. L., & Berston, S. (2019). *Contemporary Business*. John Wiley & Sons.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Busenitz, L. W., & Barney, J. B. (1997). Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making. *Journal of Business Venturing*, 12(1), 9-30.

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, *81*(5), 825-836.

Cain, D. M., Moore, D. A., & Haran, U. (2015). Making sense of overconfidence in market entry. *Strategic Management Journal*, *36*(1), 1-18.

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, *56*(5), 809-825.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of economic behavior & organization*, *81*(1), 1-8.

Chen, Z., & Kemp, S. (2015). Anchoring effects in simulated academic promotion decisions: How the promotion criterion affects ratings and the decision to support an application. *Journal of Behavioral Decision Making*, *28*(2), 137-148.

Cheng, M. M., & Hackett, R. D. (2021). A critical review of algorithms in HRM: Definition, theory, and practice. *Human Resource Management Review*, *31*(1), Article 100698.

Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. *Computers in Human Behavior*, *127*, Article 107018.

Chua, A. Y., Pal, A., & Banerjee, S. (2023). AI-enabled investment advice: Will users buy it? *Computers in Human Behavior*, *138*, Article 107481.

Claudé, M., & Combe, D. (2018). The roles of artificial intelligence and humans in decision making: Towards augmented humans? A focus on knowledge-intensive firms.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3-21.

Cummings, M. L. (2017). Automation bias in intelligent time critical decision support systems. In *Decision Making in Aviation* (pp. 289-294). Routledge: London, UK.

Das, T. K., & Teng, B. S. (1999). Cognitive biases and strategic decision processes: An integrative perspective. *Journal of Management Studies*, *36*(6), 757-778.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34(7), 571-582.

De Vries, C. F., Morrissey, B. E., Duggan, D., Staff, R. T., & Lip, G. (2021). Screening participants' attitudes to the introduction of artificial intelligence in breast screening. *Journal of Medical Screening,* 28(3), 221–222.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114-126.

Dhall, D., Kaur, R., & Juneja, M. (2020). Machine learning: a review of the algorithms and its applications. *Proceedings of ICRIC 2019*, 47-63.

Duclos, R. (2015). The psychology of investment behavior:(De) biasing financial decision-making one graph at a time. *Journal of Consumer Psychology*, *25*(2), 317-325.

Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors*, *44*(1), 79-94.

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, *58*(6), 697-718.

Emde, M. (2014). Open-ended questions in Web Surveys-Using visual and adaptive questionnaire design to improve narrative responses.

Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science*, *12*(5), 391-396.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, *17*(4), 311-318.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Freeman, G. H., & Halton, J. H. (1951). Note on an exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika*, *38*(1/2), 141-149.

Friedel, E., Koch, S. P., Wendt, J., Heinz, A., Deserno, L., & Schlagenhauf, F. (2014). Devaluation and sequential decisions: linking goal-directed and model-based behavior. *Frontiers in Human Neuroscience*, *8*, Article 587.

Gai, P. J., & Klesse, A. K. (2019). Making recommendations more effective through framings: Impacts of user-versus item-based framings on recommendation click-throughs. *Journal of Marketing*, *83*(6), 61-75.

Gentsch, P. (2018). *AI in marketing, sales and service: How marketers without a data science degree can use AI, big data and bots*. Springer.

George, J. F., Duffy, K., & Ahuja, M. (2000). Countering the anchoring and adjustment bias with decision support systems. *Decision Support Systems*, *29*(2), 195-206.

Gillis, A. S. (2022, May 19). *algorithm*. TechTarget. Retrieved October 23, 2022, from https://www.techtarget.com/whatis/definition/algorithm.

Gino, F., & Moore, D. A. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making*, *20*(1), 21-35.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. 2003 Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, Columbus, 82-88.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121-127.

Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle. *Procedia Manufacturing*, *3*, 3025-3032.

Harkness, A. R., DeBono, K. G., & Borgida, E. (1985). Personal involvement and strategies for making contingency judgments: A stake in the dating game makes a difference*. Journal of Personality and Social Psychology*, 49(1), 22-32.

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational behavior andHhuman Decision Processes*, *70*(2), 117-133.

Hastie, R. (2001). Problems for judgment and decision making. *Annual Review of Psychology*, *52*(1), 653-683.

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, *85*(1), 4-40.

Hayes, A. F. (2018). PROCESS macro for SPSS and SAS. The PROCESS macro for SPSS and SAS. *Introduction to mediation, moderation, and conditional PROCESS analysis, second edition: A regression-based approach*.

Hinshaw, S. P. (2007). Moderators and mediators of treatment outcome for youth with ADHD: Understanding for whom and how interventions work. *Journal of Pediatric Psychology*, *32*(6), 664-675.

Hodgkinson, G. P., Bown, N. J., Maule, A. J., Glaister, K. W., & Pearman, A. D. (1999). Breaking the frame: An analysis of strategic cognition and decision making under uncertainty. *Strategic Management Journal*, 20(10), 977-985.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407-434.

Hou, Y. T. Y., & Jung, M. F. (2021). Who is the expert? Reconciling algorithm aversion and algorithm appreciation in AI-supported decision making. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1-25.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67(3), 247-257.

Jamaludin, A., & Ahmad, F. (2013). Investigating the relationship between trust and intention to purchase online. *Business and Management Horizons*, *1*(1), 1-9.

Joseph, T. (2020, June 22). *Bootstrapping Statistics. What it is and why it's used.* Towards Data Science. Retrieved November 12, 2022, from https://towardsdatascience.com/bootstrapping-statistics-what-it-is-and-why-its-used-e2fa29577307

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D., Lovallo, D., & Sibony, O. (2011). Before you make that big decision. *Harvard Business Review*, *89*(6), 50-60.

Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press, Cambridge.

Kaplan, A., & 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), 15-25.

Kaur, S., & Sharma, R. (2021). Emotion AI: integrating emotional intelligence with artificial intelligence in the digital workplace. In *Innovations in Information and Communication Technologies (IICT-2020)* (pp. 337-343). Springer, Cham.

Kim, S., & Kim, B. (2020). A decision-making model for adopting al-generated news articles: Preliminary results. *Sustainability*, *12*(18), 7418.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, *133*(1), 237-293.

Kray, L., & Gonzalez, R. (1999). Differential weighting in choice versus advice: I'll do this, you do that. *Journal of Behavioral Decision Making*, 12(3), 207-218.

Lai, V., Chen, C., Liao, Q.V., Smith-Renner, A., & Tan, C. (2021). Towards a Science of Human-AI Decision Making: A Survey of Empirical Studies. *ArXiv, abs/2112.11471*.

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, *40*(1), 153-184.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50-80.

Logg, J. M. (2017). Theory of machine: When do people rely on algorithms? *Harvard Business School working paper,* No. 17-086.

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103.

MacKinnon, D. P. (2011). Integrating mediators and moderators in research design. *Research on Social Work Practice*, *21*(6), 675-681.

Malhotra, N., Nunan, D., & Birks, D. (2017). *Marketing research: An applied approach*. Pearson.

Malmendier, U., & Tate, G. (2005). Does overconfidence affect corporate investment? CEO overconfidence measures revisited. *European Financial Management*, *11*(5), 649-659.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709-734.

McDermott, R. (2011). Internal and external validity. *Cambridge Handbook of Experimental Political Science*, 27-40.

Meehl, P. E. (1954). Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.

Mosier, K. L., & Skitka, L. J. (2018). Human decision makers and automated decision aids: Made for each other? In *Automation and Human Performance: Theory and Applications* (pp. 201-220). CRC Press.

Metzger, O., & Spengler, T. (2019). Modeling rational decisions in ambiguous situations: a multi-valued logic approach. *Business Research*, *12*(1), 271-290.

Mintzberg, H., Raisinghani, D., & Theoret, A. (1976). The structure of" unstructured" decision processes. *Administrative Science Quarterly*, *21*(2), 246-275.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2), 502-517.

Montal, T., & Reich, Z. (2017). I, robot. You, journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. *Digital Journalism*, *5*(7), 829-849.

Montibeller, G., & Von Winterfeldt, D. (2015). Cognitive and motivational biases in decision and risk analysis. *Risk Analysis*, *35*(7), 1230-1251.

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. (2017). Automation bias: Decision making and performance in high-tech cockpits. In *Decision Making in Aviation* (pp. 271-288). Routledge: London, UK.

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, *39*(3), 429-460.

Nagtegaal, R., Tummers, L., Noordegraaf, M., & Bekkers, V. (2020). Designing to debias: Measuring and reducing public managers' anchoring bias. *Public Administration Review*, *80*(4), 565-576.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175-220.

Nordic Business Forum. (2018, November 20). *Andrew McAfee: Why we are still underestimating Artificial Intelligence - Nordic Business Report*. Retrieved October 12, 2022, from https://www.nbforum.com/nbreport/andrew-mcafee-why-we-are-still-underestimating-artificial-intelligence/.

Orr, D., & Guthrie, C. (2005). Anchoring, information, expertise, and negotiation: New insights from meta-analysis. *Ohio St. J. on Disp. Resol.*, *21*, 597-628.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230-253

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59.

Rastogi, C., Zhang, Y., Wei, D., Varshney, K. R., Dhurandhar, A., & Tomsett, R. (2020). Deciding fast and slow: The role of cognitive biases in ai-assisted decision-making. *ArXiv, abs/2010.07938.*

Ravikumar, S., & Saraf, P. (2020). Prediction of stock prices using machine learning (regression, classification) Algorithms. In *2020 International Conference for Emerging Technology (INCET)* (pp. 1-5). IEEE.

Sarker, A., Shamim, S. M., Zama, M. S., & Rahman, M. M. (2018). Employee's performance analysis and prediction using K-means clustering & decision tree algorithm. *Global Journal of Computer Science and Technology*, 18(1), 1–6.

Schaffer, J., O'Donovan, J., Michaelis, J., Raglin, A., & Höllerer, T. (2019, March). I can do better than your AI: expertise and explanations. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (pp. 240-251).

Schemmer, M., Hemmer, P., Kühl, N., Benz, C., & Satzger, G. (2022). Should I Follow AI-based Advice? Measuring Appropriate Reliance in Human-AI Decision-Making. *ArXiv abs/2204.06916.*

Schmitt, A., Wambsganss, T., Söllner, M., & Janson, A. (2021). Towards a Trust Reliance Paradox? Exploring the Gap Between Perceived Trust in and Reliance on Algorithmic Advice. In *International Conference on Information Systems (ICIS)* (pp. 1-17).

Schwenk, C. R. (1988). The cognitive perspective on strategic decision making. *Journal of Management Studies*, *25*(1), 41-55.

Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47-53.

Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, *61*(4), 66-83.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge-advisor decision making. *Organizational Behavior and Human Decision Processes*, *62*(2), 159-174.

Sniezek, J. A., & Van Swol, L. M. (2001). Trust, confidence, and expertise in a judge-advisor system. *Organizational Behavior and Human Decision Processes*, *84*(2), 288-307.

Simon, H. A. (1957). Models of man; social and rational.

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 780-805.

Söllner, M., Hoffmann, A., & Leimeister, J. M. (2016). Why different trust relationships matter for information systems users. *European Journal of Information Systems*, *25*(3), 274-287

Stone, E. R., & Allgaier, L. (2008). A social values analysis of self–other differences in decision making involving risk. *Basic and Applied Social Psychology*, 30(2), 114-129.

Stone, M., Aravopoulou, E., Ekinci, Y., Evans, G., Hobbs, M., Labib, A., ... & Machtynger, L. (2020). Artificial intelligence (AI) in strategic marketing decision-making: a research agenda. *The Bottom Line*, *33*(2), 183-200.

Tauchert, C., & Mesbah, N. (2019). Following the Robot? Investigating Users' Utilization of Advice from Robo-Advisors. In *ICIS*.

Thaler, R. H., & Ganser, L. J. (2015). Misbehaving: The making of behavioral economics.

Tauchert, C. (2022). Adoption of AI-based Information Systems from an Organizational and User Perspective.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124-1131.

Van Dongen, K., & Van Maanen, P. P. (2013). A framework for explaining reliance on decision aids. *International Journal of Human-Computer Studies*, *71*(4), 410-424.

Van Swol, L. M., & Sniezek, J. A. (2005). Factors affecting the acceptance of expert advice. *British Journal of Social Psychology*, 44(3), 443-461.

Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019, May). Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

Whitecotton, S. M. (1996). The effects of experience and a decision aid on the slope, scatter, and bias of earnings forecasts. *Organizational Behavior and Human Decision Processes*, *66*(1), 111-121.

Wray, L. D., & Stone, E. R. (2005). The role of self-esteem and anxiety in decision making for self versus others in relationships. *Journal of Behavioral Decision Making*, *18*(2), 125-144.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes*, 83(2), 260-281.

Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, *93*(1), 1-13.

Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making*, *32*(4), 403-414.

Yin, M., Wortman Vaughan, J., & Wallach, H. (2019, May). Understanding the effect of accuracy on trust in machine learning models. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

.

Appendix

Appendix 1: Survey 1

Informed consent

Welcome and thank you for participating in this experiment as part of my Master Thesis at Católica Lisbon School of Business and Economics.
This study consists of **two decision scenarios** and multiple questions related to them. It will take around 5 minutes to complete. I kindly ask you to answer as honestly as possible. All answers are anonymous and confidential, which means that there will be no way to link your responses to your identity. The collected data will exclusively be used for research purposes.
If you have any questions regarding this study, please do not hesitate to contact me: Kim Fortuin (152121474@alunos.lisboa.ucp.pt).

By continuing you agree to participate.
Thank you!

Demographical questions

Q1 What is your gender?

- ◯ Female (1)
- ◯ Male (2)
- ◯ Others (3)
- ◯ Prefer not to say (4)

Q3 How old are you?

____

Q4 Where are you from?

▼ Drop-down menu from Qualtrics

Q5 What is you highest level of education?

- ◯ Less than Secondary education (1)
- ◯ Secondary education (2)
- ◯ Bachelor's degree (3)
- ◯ Master's degree (4)
- ◯ Doctoral degree (5)
- ◯ Other (Please specify) (6)

Q6 What is your current employment status?

- ◯ Employed (1)
- ◯ Freelancer (2)
- ◯ Unemployed (3)
- ◯ Student (4)
- ◯ Worker and Student (5)
- ◯ Retired (6)
- ◯ Other (Please specify) (7)

Q7 How would you rate your understanding of Artificial Intelligence?

- ◯ Extremely bad (1)
- ◯ Moderately bad (2)
- ◯ Slightly bad (3)
- ◯ Neither good nor bad (4)
- ◯ Slightly good (5)
- ◯ Moderately good (6)
- ◯ Extremely good (7)

You will now be presented with two scenarios. Please make an effort to imagine yourself in the described situations and answer as realistically as possible. Thank you!

Task 1: Employee rating

Imagine that you are a manager, and you are asked to assess this year's performance of a subordinate of yours on a scale from 1 to 100. In the following you can see a summary of Charlie's performance this year:
"During this year, Charlie met the majority of goals, had good interpersonal skills with their colleagues, and showed moderate creativity in proposing new ideas for the improvement of the services."

If Anchor 1 = High anchor
The previous year, you assigned Charlie a performance rating of 51/100.

If Anchor 1 = Low anchor
The previous year, you assigned Charlie a performance rating of 91/100.

Q8 Based on the description above, what performance rating do you assign Charlie this year?

____

Q9 Please state your agreement with the following three statements regarding the previous task:

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I believe I performed well in this task. (1)	○	○	○	○	○	○	○
I am confident in my answer. (2)	○	○	○	○	○	○	○
I believe my estimate is accurate. (3)	○	○	○	○	○	○	○

Q10 To make sure you read this question carefully, please select "Strongly agree".

○ Strongly disagree (1)
○ Disagree (2)
○ Somewhat disagree (3)
○ Neither agree nor disagree (4)
○ Somewhat agree (5)
○ Agree (6)
○ Strongly agree (7)

If Condition = AI

Now let's assume...
A machine learning algorithm named **Taylor*** assessed Charlie's performance based on various data from performance and project reviews. In the past, Taylor has proven highly effective in measuring employee performance.
The rating that Taylor estimated for Charlie's performance was: **71/100**

**Machine learning is a discipline of artificial intelligence that provides machines the ability to automatically learn from data and past experiences. Artificial intelligence leverages computers and machines to mimic cognitive skills that are associated with the human mind, such as learning, problem-solving, and decision-making.*
The algorithm Taylor uses data clustering to find groups in the data and separates employees' performance into clusters such as excellent, good, average, and poor according to their performance. It works iteratively to assign each data point to one group based on the features that are provided.

If Condition = Human

Now let's assume...
A colleague of yours named **Taylor** assessed Charlie's performance based on the collected performance and project reviews. In the past, Taylor has proven highly effective in measuring employee performance.
The rating that Taylor estimated for Charlie's performance was: **71/100**

Q11 Please state your agreement with the following three statements regarding Taylor's recommendation:

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I believe this recommendation is trustworthy. (1)	○	○	○	○	○	○	○
I consider Taylor to be reliable. (2)	○	○	○	○	○	○	○
I trust the decision of Taylor. (3)	○	○	○	○	○	○	○

Q12 After receiving this information, would you like to reconsider your previous estimate?

If yes, please state your final estimate below, otherwise leave it empty.

____

Task 2: Stock price prediction

Imagine you are asked to predict the closing price of a stock, whereby the closing price of a stock refers to the "last" price at which a company's stock is traded when the market closes at the end of the day.
The graph below reports the closing price of a company's stock for the last 30 days. Please, consider all further information encountered in this graph.

If Anchor 2 = Upward anchor

If Anchor 2 = Downward anchor

Q13 Looking at the graph above, what do you think the closing price of this stock will be by the end of the day?

____

Q14 Please state your agreement with the following three statements regarding the previous task:

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I believe I performed well in this task. (1)	○	○	○	○	○	○	○
I am confident in my answer. (2)	○	○	○	○	○	○	○
I believe my estimate is accurate. (3)	○	○	○	○	○	○	○

If Condition = AI

Now let's assume...
A machine learning algorithm named **Revel*** performed calculations based on extensive stock data, both historical and current data. In the past Revel has proven highly effective in forecasting the closing price of a company's stock.
The closing price that Revel estimated was: **$60**

**Machine learning is a discipline of artificial intelligence that provides machines the ability to automatically learn from data and past experiences. Artificial intelligence leverages computers and machines to mimic cognitive skills that are associated with the human mind, such as learning, problem-solving, and decision-making.*
The algorithm Revel is based on a deep learning framework for time-series. The system uses regression and classification to predict the closing price of stock of a company and whether it will increase or decrease the next day.

If Condition = Human

Now let's assume...
A colleague of yours named **Revel** performed calculations based on extensive stock data, both historical and current data. In the past Revel has proven highly effective in forecasting the closing price of a company's stock.
The closing price that Revel estimated was: **$60**

Q15 Please state your agreement with the following three statements regarding Revel's recommendation:

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I believe this recommendation is trustworthy. (1)	○	○	○	○	○	○	○
I consider Revel to be reliable. (2)	○	○	○	○	○	○	○
I trust the decision of Revel. (3)	○	○	○	○	○	○	○

Q16 After receiving this information, would you like to reconsider your previous estimate?

If yes, please state your final estimate below, otherwise leave it empty.

____

Final questions

Final questions

Q16 Which agent provided you with a recommendation in the previously described scenarios?

○ A colleague (1)
○ An algorithm (2)
○ Other (Please specify) (3)

Q17 Answering the questions of this study was...

○ Extremely difficult (1)
○ Moderately difficult (2)
○ Slightly difficult (3)
○ Neither easy nor difficult (4)
○ Slightly easy (5)
○ Moderately easy (6)
○ Extremely easy (7)

Q18 Imagining the previously described scenarios was...

- Extremely difficult (1)
- Moderately difficult (2)
- Slightly difficult (3)
- Neither easy nor difficult (4)
- Slightly easy (5)
- Moderately easy (6)
- Extremely easy (7)

Q19 How do you assess your previous knowledge regarding the two tasks?

	Extremely bad (1)	Moderately bad (2)	Slightly bad (3)	Neither good nor bad (4)	Slightly good (5)	Moderately good (6)	Extremely good (7)
Employee Performance Rating (1)	○	○	○	○	○	○	○
Stock Price Prediction (2)	○	○	○	○	○	○	○

Q20 How much attention did you pay during this survey?

- None at all (1)
- A little (2)
- A moderate amount (3)
- A lot (4)
- A great deal (5)

Q21 Do you have any comments you would like to share with the researcher?

If so, please write them in the box below. Otherwise, just leave it blank.

____

End of Survey text

Thank you for participating. Your responses have been transmitted.
The goal of this study was to measure the reliance on advice from different sources. Therefore, participants were randomly assigned to either a human advisor or an algorithmic advisor.
If you have any questions or comments, do not hesitate to send me an email via
152121474@alunos.lisboa.ucp.pt
Have a nice day.

Appendix 2: Survey 2

Informed consent

Welcome and thank you for participating in this experiment at Católica Lisbon School of Business and Economics.
This study consists of **two scenarios** and multiple questions related to them. It will take around 3 minutes to complete. I kindly ask you to answer as honestly as possible. All answers are anonymous and confidential, which means that there will be no way to link your responses to your identity. The collected data will exclusively be used for research purposes.
If you have any questions regarding this study, please do not hesitate to contact me: Filipa de Almeida at filipadealmeida@ucp.pt.

By continuing you agree to participate.
Thank you!

Q1 What is your participant ID?

____

Demographical questions

Q1 What is your gender?

- ○ Female (1)
- ○ Male (2)
- ○ Others (3)
- ○ Prefer not to say (4)

Q2 How old are you?

____

Q3 Where are you from?

▼ Drop-down menu provided by Qualtrics

Q4 How would you rate your understanding of Artificial Intelligence?

- ○ Extremely bad (1)
- ○ Moderately bad (2)
- ○ Slightly bad (3)
- ○ Neither good nor bad (4)
- ○ Slightly good (5)
- ○ Moderately good (6)
- ○ Extremely good (7)

You will now be presented with two scenarios. Please make an effort to imagine yourself in the described situations and answer as realistically as possible. Thank you!

Task 1: Employee rating

If Condition = Self

Imagine that you are a manager, and you are asked to assess this year's performance of a subordinate of yours on a scale from 1 to 100. Next, you can see a summary of Charlie's performance this year: *"During this year, Charlie met the majority of goals, had good interpersonal skills with their colleagues, and showed moderate creativity in proposing new ideas for the improvement of the services."*

To support you with this decision a machine learning algorithm* assessed Charlie's performance based on various data from performance and project reviews.
So did a colleague of yours using the same data.
In the past both - the algorithm and your colleague - have proven equally effective in measuring employee performance.

**Same info on ML, AI, algorithms as in Study 1*

If Condition = Others

Imagine that you are a manager, and you asked your employee **Kris** to assess this year's performance of his/her subordinate Charlie on a scale from 1 to 100. Next, you can see a summary of Charlie's performance this year:
"During this year, Charlie met the majority of goals, had good interpersonal skills with their colleagues, and showed moderate creativity in proposing new ideas for the improvement of the services."

To support your employee Kris with this decision, a machine learning algorithm* assessed Charlie's performance based on various data from performance and project reviews.
So did a colleague of yours using the same data.
In the past both - the algorithm and your colleague - have proven equally effective in measuring employee performance.

**Same info on ML, AI, algorithms as in Study 1*

Q5 To make sure you read this question carefully, please select "Strongly agree".

○ Strongly disagree (1)
○ Disagree (2)
○ Somewhat disagree (3)
○ Neither agree nor disagree (4)
○ Somewhat agree (5)
○ Agree (6)
○ Strongly agree (7)

If Condition = Self

Q6 If you had to choose between the two sources of advice, which one would you pick?

- ◯ Definitely the human (1)
- ◯ (2)
- ◯ (3)
- ◯ Indifferent between the two (4)
- ◯ (5)
- ◯ (6)
- ◯ Definitely the algorithm (7)

If Condition = Others

Q7 If you had to choose between the two sources of advice for your employee Kris to follow, which one would you pick for him/her?

- ◯ Definitely the human (1)
- ◯ (2)
- ◯ (3)
- ◯ Indifferent between the two (4)
- ◯ (5)
- ◯ (6)
- ◯ Definitely the algorithm (7)

Q8 Please state the main reason for your decision in the box below:

____

Task 2: Stock price prediction

If Condition = Self

Imagine you are asked to predict the closing price of a stock, whereby the closing price of a stock refers to the "last" price at which a company's stock is traded when the market closes at the end of the day.
The graph below reports the closing price of a company's stock for the last 30 days. Please, consider all further information encountered in this graph.

To support you with this decision a machine learning algorithm* performed calculations based on extensive stock data, both historical and current data.
So did a colleague of yours using the same data.
In the past both - the algorithm and your colleague - have proven equally effective in forecasting the closing price of a company's stock.

**Same info on ML, AI, algorithms as in Study 1*

If Condition = Others

Imagine that you are a manager, and you asked your employee **Jamie** to predict the closing price of a stock, whereby the closing price of a stock refers to the "last" price at which a company's stock is traded when the market closes at the end of the day.

The graph below reports the closing price of a company's stock for the last 30 days. Please, consider all further information encountered in this graph.

To support your employee Jamie with this decision a machine learning algorithm* performed calculations based on extensive stock data, both historical and current data.
So did a colleague of yours using the same data.
In the past both - the algorithm and your colleague - have proven equally effective in forecasting the closing price of a company's stock.

**Same info on ML, AI, algorithms as in Study 1*

If Condition = Self

Q9 If you had to choose between the two sources of advice, which one would you pick?

- Definitely the human (1)
- (2)
- (3)
- Indifferent between the two (4)
- (5)
- (6)
- Definitely the algorithm (7)

If Condition = Others

Q10 If you had to choose between the two sources of advice for your employee Jamie to follow, which one would you pick for him/her?

◯ Definitely the human (1)
◯ (2)
◯ (3)
◯ Indifferent between the two (4)
◯ (5)
◯ (6)
◯ Definitely the algorithm (7)

Q11 Please state the main reason for your decision in the box below:

____

Final questions

Final questions

Q12 How much attention did you pay during this survey?

◯ None at all (1)
◯ A little (2)
◯ A moderate amount (3)
◯ A lot (4)
◯ A great deal (5)

Q13 Do you have any comments you would like to share with the researcher?

If so, please write them in the box below. Otherwise, just leave it blank.

____

End of Survey text

Thank you for participating. Your responses have been transmitted.
The goal of this study was to explore the expectations regarding the preference for different sources of advice and the reasons behind it.
If you have any questions or comments, do not hesitate to send me an email via filipadealmeida@ucp.pt
Have a nice day!

Data Analysis Study 1

Appendix 3: Population Statistics

	VALUES	FREQUENCY	PERCENT	CUMULATIVE PERCENT
GENDER	female	98	47,8	47,8
	male	98	47,8	95,6
	diverse	6	2,9	98,5
	Other	3	1,5	100,0
	Total	205	100,0	
AGE	19	1	,5	,5
	20	3	1,5	2,0
	21	2	1,0	2,9
	22	2	1,0	3,9
	23	7	3,4	7,4
	24	20	9,8	17,2
	25	20	9,8	27,0
	26	12	5,9	32,8
	27	10	4,9	37,7
	28	5	2,5	40,2
	29	6	2,9	43,1
	30	3	1,5	44,6
	31	2	1,0	45,6
	32	2	1,0	46,6
	33	6	2,9	49,5
	34	5	2,5	52,0
	35	6	2,9	54,9
	36	3	1,5	56,4
	37	5	2,5	58,8
	38	3	1,5	60,3
	39	7	3,4	63,7
	40	8	3,9	67,6
	41	1	,5	68,1
	42	1	,5	68,6
	43	2	1,0	69,6
	44	4	2,0	71,6
	45	6	2,9	74,5
	46	1	,5	75,0
	47	2	1,0	76,0
	49	1	,5	76,5
	50	4	2,0	78,4
	51	6	2,9	81,4
	52	1	,5	81,9
	53	3	1,5	83,3
	54	6	2,9	86,3
	55	3	1,5	87,7
	56	5	2,5	90,2
	57	5	2,5	92,6
	58	3	1,5	94,1
	59	3	1,5	95,6
	60	2	1,0	96,6
	61	1	,5	97,1
	63	1	,5	97,5
	66	1	,5	98,0
	72	1	,5	98,5
	76	2	1,0	99,5
	78	1	,5	100,0
	Total	204	100,0	

EDUCATION	Below Secondary	1	,5	,5
	Secondary	27	13,2	13,7
	Bachelor	82	40,0	53,7
	Master	73	35,6	89,3
	Doctor	1	,5	89,8
	Other	21	10,2	100,0
	Total	205	100,0	
EMPLOYMENT	Employed	94	46,1	46,1
	Freelancer	44	21,6	67,6
	Unemployed	4	2,0	69,6
	Student	32	15,7	85,3
	Worker and Student	19	9,3	94,6
	Retired	6	2,9	97,5
	Other	5	2,5	100,0
	Total	204	100,0	
COUNTRY	Angola	1	,5	,5
	Argentina	6	3,0	3,4
	Armenia	1	,5	3,9
	Austria	29	14,3	18,2
	Belgium	11	5,4	23,6
	Brazil	8	3,9	27,6
	Denmark	1	,5	28,1
	France	10	4,9	33,0
	Germany	88	43,3	76,4
	Hong Kong	1	,5	76,8
	Italy	11	5,4	82,3
	Mexico	2	1,0	83,3
	Netherlands	11	5,4	88,7
	Poland	1	,5	89,2
	Portugal	15	7,4	96,6
	Slovenia	1	,5	97,0
	South Africa	1	,5	97,5
	Spain	1	,5	98,0
	Switzerland	2	1,0	99,0
	United Kingdom	2	1,0	100,0
	Total	203	100,0	

Population Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	204	19	78	36,93	13,301
Understanding of Artificial Intelligence	205	1	7	4,41	1,240
Previous knowledge of the tasks	205	1	7	4,49	1,31

Appendix 4: Ordering effect on Dependent Variable (WOA)

The order of the two tasks was randomized to assure there are no order effects between the first and the second tasks. To examine whether the order of the tasks had an impact on the participants' answers a one-way ANOVA was performed measuring the effect of the different orders on the dependent variable. As expected, no significant differences between the two groups could be identified ($F(1, 204) = 1.51$; $p = .22$). Therefore, the following analyses does not distinguish between the different orders.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
WOA 1	Between Groups	,417	1	,417	2,752	,099
	Within Groups	30,771	203	,152		
	Total	31,189	204			
WOA 2	Between Groups	,041	1	,041	,245	,621
	Within Groups	33,523	203	,165		
	Total	33,564	204			
WOA	Between Groups	,179	1	,179	1,513	,220
	Within Groups	24,076	203	,119		
	Total	24,256	204			

Appendix 5: Scale Reliability

A reliability analysis was conducted to test for the Cronbach's alpha of the two scales used. The scale for confidence in oneself as well as the scale for trust in the source both yielded a Cronbach's alpha of 0.96 which is considered as excellent according to George & Gliem (2003). Thus, both constructs showed high internal consistency between the items and are reliable enough to predict the respective variables.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,960	,960	3	,962	,962	3

Item Statistics

	Mean	Std. Deviation	N		Mean	Std. Deviation	N
Trust_1	4,85	1,496	205	Conf_1	5,13	1,309	205
Trust_2	4,82	1,465	205	Conf_2	5,08	1,350	205
Trust_3	4,72	1,490	205	Conf_3	5,09	1,301	205

Appendix 6: Manipulation check

Participants were randomly distributed across conditions, where a total of 105 participants received advice from AI and 100 from a human. To examine whether the manipulations of the variable source of advice worked as intended a Fisher's Exact Test of Independence was used, as this measure involves two nominal variables with counts less than five (Freeman & Halton, 1951). The results showed that there was a significant dependence between relative proportions of the different sources of advice and the responses to the manipulation check ($p < .001$). Thus, the null hypothesis of independence can be rejected, meaning that the manipulation was successful as most participants recognized the respective agents and answered correctly ($N = 135$).

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	41,316[a]	2	<,001	<,001		
Likelihood Ratio	43,032	2	<,001	<,001		
Fisher-Freeman-Halton Exact Test	42,530			<,001		
Linear-by-Linear Association	19,413[b]	1	<,001	<,001	<,001	,000
N of Valid Cases	199					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 8,29.
b. The standardized statistic is 4,406.

Appendix 7: Comparison of the two tasks (WOA, Trust)

As mentioned previously, I expected no difference between the two decision making tasks. Before aggregating the two tasks, I verified whether the effect of the manipulation differed between the two tasks. To test for that, two-way repeated measures ANOVAs were conducted – one regarding WOA and one regarding trust. As expected, the tests show that there is no significant difference between the two tasks when looking at the effect of different sources of advice on the two variables mentioned above ($F(3, 8028701) = 1.78$; $p = .15$ for WOA and $F(3, 8028701) = .08$; $p = .97$ for trust). Thus, for the remaining analysis the two tasks were combined meaning that the respective variables were aggregated by their means.

Deskriptive Statistiken

	Condition	Mittelwert	Standardabweichung	N
WOA 1	AI	,399005830	,415500592	100
	Human	,254394328	,353906972	105
	Gesamt	,324936524	,391005340	205
WOA 2	AI	,404933824	,425510948	100
	Human	,257189542	,373663261	105
	Gesamt	,329259923	,405621382	205

Descriptive Statistics

	Condition	Mean	Std. Deviation	N
Trust 1	AI	5,06666667	1,42842712	100
	Human	4,54285714	1,38463573	105
	Total	4,79837398	1,42705631	205
Trust 2	AI	5,36000000	1,27628811	100
	Human	4,49523810	1,22222500	105
	Total	4,91707317	1,31902292	205

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[c]
Task	Pillai's Trace	,000	,024[b]	1,000	203,000	,876	,000	,024	,053
	Wilks' Lambda	1,000	,024[b]	1,000	203,000	,876	,000	,024	,053
	Hotelling's Trace	,000	,024[b]	1,000	203,000	,876	,000	,024	,053
	Roy's Largest Root	,000	,024[b]	1,000	203,000	,876	,000	,024	,053
Task * Condition	Pillai's Trace	,000	,003[b]	1,000	203,000	,955	,000	,003	,050
	Wilks' Lambda	1,000	,003[b]	1,000	203,000	,955	,000	,003	,050
	Hotelling's Trace	,000	,003[b]	1,000	203,000	,955	,000	,003	,050
	Roy's Largest Root	,000	,003[b]	1,000	203,000	,955	,000	,003	,050

a. Design: Intercept + Condition
Within Subjects Design: Task

b. Exact statistic

c. Computed using alpha = ,05

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[c]
Task	Pillai's Trace	,008	1,731[b]	1,000	203,000	,190	,008	1,731	,258
	Wilks' Lambda	,992	1,731[b]	1,000	203,000	,190	,008	1,731	,258
	Hotelling's Trace	,009	1,731[b]	1,000	203,000	,190	,008	1,731	,258
	Roy's Largest Root	,009	1,731[b]	1,000	203,000	,190	,008	1,731	,258
Task * Condition	Pillai's Trace	,016	3,333[b]	1,000	203,000	,069	,016	3,333	,443
	Wilks' Lambda	,984	3,333[b]	1,000	203,000	,069	,016	3,333	,443
	Hotelling's Trace	,016	3,333[b]	1,000	203,000	,069	,016	3,333	,443
	Roy's Largest Root	,016	3,333[b]	1,000	203,000	,069	,016	3,333	,443

a. Design: Intercept + Condition
Within Subjects Design: Task

b. Exact statistic

c. Computed using alpha = ,05

Appendix 8: Effects of different anchors on dependent variable

I tested whether the different anchors impacted the dependent variable. To do so, I ran two one-way ANOVAs with anchor type and WOA as the dependent variable. As expected, the results show that there were no significant differences between the high (upward) and low (downward) anchor groups on the WOA on a 95% confidence interval for both tasks ($F(1, 202) = .08$, $p = .77$ for T1; $F(1, 202) = 3.56$, $p = .06$ for T2).

ANOVA for Task 1

WOA 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	,013	1	,013	,082	,774
Within Groups	31,176	203	,154		
Total	31,189	204			

ANOVA for Task 2

WOA 2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	,578	1	,578	3,557	,061
Within Groups	32,986	203	,162		
Total	33,564	204			

Appendix 9: Test for Anchoring Bias

Task 1, Anchor 1:

Group Statistics for Task 1

	Anchor 1	N	Mean	Std. Deviation	Std. Error Mean
First Estimate Task 1	High anchor	101	82,02	10,337	1,029
	Low anchor	104	61,29	10,660	1,045

Independent Samples Test for Task 1

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
First Estimate T1	Equal variances assumed	,002	,964	14,130	203	<,001	<,001	20,731	1,467	17,838	23,624
	Equal variances not assumed			14,136	203,000	<,001	<,001	20,731	1,467	17,840	23,623

Task 2, Anchor 2:

Group Statistics for Task 2

	Anchor 2 coded	N	Mean	Std. Deviation	Std. Error Mean
First Estimate Task 2	Upward anchor	103	72,99	14,738	1,452
	Downward anchor	102	49,00	11,259	1,115

Independent Samples Test for Task 2

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
First Estimate T2	Equal variances assumed	,940	,333	13,087	203	<,001	<,001	23,990	1,833	20,376	27,605
	Equal variances not assumed			13,104	190,748	<,001	<,001	23,990	1,831	20,379	27,601

Appendix 10: Hypothesis testing with Hayes PROCESS macro

Model 1

```
Run MATRIX procedure:

*************** PROCESS Procedure for SPSS Version 4.2 beta ***************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3

**************************************************************************
Model  : 14
    Y  : WOA
    X  : COND
    M  : TRUST
    W  : CONF

Covariates:
 DIST     PREKNOW  UnderAI  Gender   Age      Country  Educ     Employm

Sample
Size:  202

**************************************************************************
OUTCOME VARIABLE:
 TRUST

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      ,3478      ,1210     1,3183     2,9357     9,0000   192,0000      ,0028

Model
```

	coeff	se	t	p	LLCI	ULCI
constant	-,0105	,6049	-,0174	,9861	-1,2036	1,1826
COND	,7117	,1650	4,3142	,0000	,3863	1,0371
DIST	-,0129	,0125	-1,0346	,3022	-,0376	,0117
PREKNOW	-,0758	,0705	-1,0741	,2841	-,2149	,0634
UnderAI	,0671	,0724	,9259	,3556	-,0758	,2099
Gender	-,1324	,1332	-,9938	,3216	-,3951	,1304
Age	-,0069	,0064	-1,0692	,2863	-,0196	,0058
Country	-,0001	,0020	-,0284	,9774	-,0040	,0039
Educ	,0659	,0798	,8258	,4100	-,0915	,2232
Employm	,0459	,0509	,9011	,3687	-,0546	,1463

****

Model 2

OUTCOME VARIABLE:
WOA

Model Summary

R	R-sq	MSE	F	df1	df2	p
,7179	,5154	,0616	16,7489	12,0000	189,0000	,0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	,1696	,1325	1,2799	,2022	-,0918	,4310
COND	,0205	,0375	,5478	,5845	-,0534	,0945
TRUST	,1621	,0159	10,1716	,0000	,1307	,1936
CONF	-,0643	,0167	-3,8561	,0002	-,0971	-,0314
Int_1	-,0474	,0119	-3,9720	,0001	-,0709	-,0238
DIST	,0096	,0028	3,3628	,0009	,0040	,0152
PREKNOW	,0478	,0168	2,8509	,0048	,0147	,0809
UnderAI	-,0338	,0157	-2,1527	,0326	-,0648	-,0028
Gender	-,0007	,0290	-,0257	,9795	-,0580	,0565
Age	-,0008	,0014	-,6040	,5466	-,0036	,0019
Country	,0001	,0004	,3281	,7432	-,0007	,0010
Educ	,0102	,0173	,5911	,5551	-,0239	,0444
Employm	-,0274	,0111	-2,4771	,0141	-,0492	-,0056

Product terms key:
Int_1 : TRUST x CONF

Test(s) of highest order unconditional interaction(s):

	R2-chng	F	df1	df2	p
M*W	,0405	15,7770	1,0000	189,0000	,0001

Focal predict: TRUST (M)
Mod var: CONF (W)

Conditional effects of the focal predictor at values of the moderator(s):

CONF	Effect	se	t	p	LLCI	ULCI
-1,2163	,2197	,0226	9,7310	,0000	,1752	,2643
,0000	,1621	,0159	10,1716	,0000	,1307	,1936
1,2163	,1045	,0205	5,1075	,0000	,0642	,1449

****

Data for visualization

```
Data for visualizing the conditional effect of the focal predictor:
Paste text below into a SPSS syntax window and execute to produce plot.

DATA LIST FREE/
   TRUST      CONF       WOA        .
BEGIN DATA.
    -1,1969    -1,2163      ,1342
      ,0000    -1,2163      ,3972
     1,1969    -1,2163      ,6601
    -1,1969      ,0000      ,1250
      ,0000      ,0000      ,3190
     1,1969      ,0000      ,5130
    -1,1969     1,2163      ,1157
      ,0000     1,2163      ,2408
     1,1969     1,2163      ,3660
END DATA.
GRAPH/SCATTERPLOT=
 TRUST    WITH     WOA      BY       CONF     .

**************************************************************************
```

Index of Moderated Mediation

```
****************** DIRECT AND INDIRECT EFFECTS OF X ON Y *****************

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      ,0205      ,0375      ,5478      ,5845     -,0534      ,0945

Conditional indirect effects of X on Y:

INDIRECT EFFECT:
 COND        ->    TRUST       ->    WOA

       CONF     Effect     BootSE   BootLLCI   BootULCI
    -1,2163      ,1564      ,0367      ,0885      ,2336
      ,0000      ,1154      ,0281      ,0631      ,1749
     1,2163      ,0744      ,0240      ,0330      ,1264

      Index of moderated mediation:
          Index     BootSE   BootLLCI   BootULCI
CONF     -,0337      ,0109     -,0568     -,0146

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95,0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  5000

W values in conditional tables are the mean and +/- SD from the mean.

NOTE: The following variables were mean centered prior to analysis:
          CONF     TRUST

------ END MATRIX -----
```

Appendix 11: Correlation matrix

Correlations

		Gender	Age	Country	Educ	Employm	UnderAI	Condition coded	WOA	CONF	TRUST	PREKNOW	DIST
Gender	Pearson Correlation	1	,083	-,014	,005	-,048	,080	-,048	-,019	,036	-,116	,228**	,057
	Sig. (2-tailed)		,236	,840	,940	,500	,255	,497	,784	,609	,098	<,001	,419
	N	205	204	203	205	204	205	205	205	205	205	205	205
Age	Pearson Correlation	,083	1	-,159*	,159*	-,213**	,044	-,029	-,045	,051	-,107	,090	,141*
	Sig. (2-tailed)	,236		,023	,023	,002	,529	,681	,527	,468	,129	,202	,045
	N	204	204	203	204	203	204	204	204	204	204	204	204
Country	Pearson Correlation	-,014	-,159*	1	-,226**	-,035	-,051	-,001	,023	,093	-,013	,059	-,009
	Sig. (2-tailed)	,840	,023		,001	,620	,470	,989	,746	,186	,858	,400	,893
	N	203	203	203	203	202	203	203	203	203	203	203	203
Educ	Pearson Correlation	,005	,159*	-,226**	1	-,137	,189**	-,020	,029	,023	,049	,133	-,035
	Sig. (2-tailed)	,940	,023	,001		,050	,007	,780	,679	,742	,484	,057	,615
	N	205	204	203	205	204	205	205	205	205	205	205	205
Employm	Pearson Correlation	-,048	-,213**	-,035	-,137	1	-,074	-,152*	-,164*	-,046	,056	-,144*	-,193**
	Sig. (2-tailed)	,500	,002	,620	,050		,294	,030	,019	,512	,429	,040	,006
	N	204	203	202	204	204	204	204	204	204	204	204	204
UnderAI	Pearson Correlation	,080	,044	-,051	,189**	-,074	1	,052	-,066	,135	,045	,405**	-,075
	Sig. (2-tailed)	,255	,529	,470	,007	,294		,463	,351	,055	,523	<,001	,288
	N	205	204	203	205	204	205	205	205	205	205	205	205
Condition coded	Pearson Correlation	-,048	-,029	-,001	-,020	-,152*	,052	1	,212**	-,011	,290**	,030	,094
	Sig. (2-tailed)	,497	,681	,989	,780	,030	,463		,002	,875	<,001	,674	,182
	N	205	204	203	205	204	205	205	205	205	205	205	205
WOA	Pearson Correlation	-,019	-,045	,023	,029	-,164*	-,066	,212**	1	-,328**	,546**	-,041	,268**
	Sig. (2-tailed)	,784	,527	,746	,679	,019	,351	,002		<,001	<,001	,555	<,001
	N	205	204	203	205	204	205	205	205	205	205	205	205
CONF	Pearson Correlation	,036	,051	,093	,023	-,046	,135	-,011	-,328**	1	-,179*	,442**	-,184**
	Sig. (2-tailed)	,609	,468	,186	,742	,512	,055	,875	<,001		,010	<,001	,008
	N	205	204	203	205	204	205	205	205	205	205	205	205
TRUST	Pearson Correlation	-,116	-,107	-,013	,049	,056	,045	,290**	,546**	-,179*	1	-,065	-,067
	Sig. (2-tailed)	,098	,129	,858	,484	,429	,523	<,001	<,001	,010		,353	,339
	N	205	204	203	205	204	205	205	205	205	205	205	205
PREKNOW	Pearson Correlation	,228**	,090	,059	,133	-,144*	,405**	,030	-,041	,442**	-,065	1	-,116
	Sig. (2-tailed)	<,001	,202	,400	,057	,040	<,001	,674	,555	<,001	,353		,097
	N	205	204	203	205	204	205	205	205	205	205	205	205
DIST	Pearson Correlation	,057	,141*	-,009	-,035	-,193**	-,075	,094	,268**	-,184**	-,067	-,116	1
	Sig. (2-tailed)	,419	,045	,893	,615	,006	,288	,182	<,001	,008	,339	,097	
	N	205	204	203	205	204	205	205	205	205	205	205	217

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

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

Appendix 12: WOA Analysis

Although people should average their own estimates with that of others to maximize accuracy, resulting in a WOA of 0.5, results show that people underweighted advice. Thereby, the results show similar effects as previous research on advice taking (*M* = 0.26, *SD* = 0.30 for human advice, *M* = 0.40, *SD* = 0.37 for AI advice).

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
WOAAI	100	,00	1,00	,4020	,37354
WOAHuman	105	-,19	1,00	,2558	,29980
Valid N (listwise)	0				

Data Analysis Study 2

Appendix 13: Population Statistics

	VALUES	FREQUENCY	PERCENT	CUMULATIVE PERCENT
GENDER	female	110	53,7	53,7
	male	95	46,3	100,0
	Total	205	100,0	
AGE	19	1	,5	,5
	20	2	1,0	1,5
	21	30	14,6	16,1
	22	31	15,1	31,2
	23	50	24,4	55,6
	24	37	18,0	73,7
	25	30	14,6	88,3
	26	13	6,3	94,6
	27	4	2,0	96,6
	28	5	2,4	99,0
	29	1	,5	99,5
	30	1	,5	100,0
	Total	205	100,0	

Population Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Understanding of Artificial Intelligence	205	1	7	3,82	1,546

Appendix 14: Comparison of the two tasks (Preference)

I tested whether the effect of the manipulation (self vs. others) differed between the two decision-making scenarios. Thereby, a significant difference between the two scenarios was found regarding the effect of the manipulation on the preference of advice. Thus, the two scenarios will *not* be aggregated but will be treated separately.

Descriptive Statistics

	Condition Coded	Mean	Std. Deviation	N
T1	0	3,427	1,7467	103
	1	3,723	1,9190	101
	Total	3,574	1,8355	204
T2	0	5,184	1,6730	103
	1	5,307	1,4747	101
	Total	5,245	1,5753	204

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Noncent. Parameter	Observed Power[c]
Task	Pillai's Trace	,372	119,623[b]	1,000	202,000	<,001	119,623	1,000
	Wilks' Lambda	,628	119,623[b]	1,000	202,000	<,001	119,623	1,000
	Hotelling's Trace	,592	119,623[b]	1,000	202,000	<,001	119,623	1,000
	Roy's Largest Root	,592	119,623[b]	1,000	202,000	<,001	119,623	1,000
Task * COND	Pillai's Trace	,002	,321[b]	1,000	202,000	,572	,321	,087
	Wilks' Lambda	,998	,321[b]	1,000	202,000	,572	,321	,087
	Hotelling's Trace	,002	,321[b]	1,000	202,000	,572	,321	,087
	Roy's Largest Root	,002	,321[b]	1,000	202,000	,572	,321	,087

a. Design: Intercept + COND
Within Subjects Design: Task
b. Exact statistic
c. Computed alpha = ,05

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Two-Sided p
Pair 1	T1 & T2	204	,191	,003	,006

Paired Samples Test

		Paired Differences Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Significance One-Sided p	Two-Sided p
Pair 1	T1 - T2	-1,6716	2,1780	,1525	-1,9722	-1,3709	-10,962	203	<,001	<,001

Appendix 15: Ordering effect on Dependent Variable (Preference)

No significant differences between the two order groups could be identified in any of the two tasks. Thus, the remaining analysis did not distinguish between different orders

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Preference Task 1	Between Groups	1,784	1	1,784	,363	,548
	Within Groups	998,168	203	4,917		
	Total	999,951	204			
Preference Task 2	Between Groups	,697	1	,697	,181	,671
	Within Groups	778,990	202	3,856		
	Total	779,686	203			
Preference Total	Between Groups	,049	1	,049	,018	,892
	Within Groups	534,696	202	2,647		
	Total	534,745	203			

Appendix 16: Test for Self-Other Differences (T1, T2, Combined)

Group Statistics

	Condition Coded	N	Mean	Std. Deviation	Std. Error Mean
T1	1	102	3,716	1,9108	,1892
	0	103	3,427	1,7467	,1721
T2	1	101	5,307	1,4747	,1467
	0	103	5,184	1,6730	,1648
TOT	1	101	4,515	1,3294	,1323
	0	103	4,306	1,3065	,1287

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
T1	Equal variances assumed	,622	,431	1,128	203	,130	,260	,2885	,2557	-,2156	,7926
	Equal variances not assumed			1,128	201,018	,130	,261	,2885	,2558	-,2158	,7928
T2	Equal variances assumed	,583	,446	,554	202	,290	,580	,1225	,2210	-,3132	,5582
	Equal variances not assumed			,555	199,759	,290	,580	,1225	,2207	-,3127	,5577
TOT	Equal variances assumed	,549	,459	1,133	202	,129	,259	,2090	,1846	-,1549	,5729
	Equal variances not assumed			1,132	201,723	,129	,259	,2090	,1846	-,1549	,5730

Appendix 17: Test for Preferences (T1, T2, Combined)

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
T1	205	3,571	1,8314	,1279
T2	204	5,245	1,5753	,1103
TOT	204	4,409	1,3188	,0923

One-Sample Test

Test Value = 4

			Significance			95% Confidence Interval of the Difference	
	t	df	One-Sided p	Two-Sided p	Mean Difference	Lower	Upper
T1	-3,356	204	<,001	<,001	-,4293	-,681	-,177
T2	11,289	203	<,001	<,001	1,2451	1,028	1,463
TOT	4,433	203	<,001	<,001	,4093	,227	,591

Appendix 18: Qualitative Analysis

Main reasons

Reasons for choosing the AI advice	Reasons for choosing the human advice
General trust towards algorithms • Assuming human error • Gut-feeling	General trust towards humans • Assuming algorithmic error • Gut-feeling
Performance • Reliability • Accuracy • Consistency • Precision • Predictability • Machine Intelligence • Error rate (vs. Human error) • Past performance Objectivity and rationality • Impartiality • Rationality (vs. Human irrationality) • Unbiasedness (vs. Human bias) Transparency regarding the process • Human preparation • Training set • Data input • Development status • Transparency Efficiency • Effectiveness • Speed • Cost-saving • Room for better resource allocation Future oriented mindset Assessment of quantifiable data • Numbers driven • Data driven	Emotional intelligence of the advisor • Empathy • Emotions • Feelings • Affective component • Human touch Direct point of contact • Interaction with the human advisor • Responsibility • Tangibility No previous knowledge of AI Fear of job replacement • Job replacement • No need for AI • Need for human control • Tradition • Fairness Assessment of non-quantifiable data • Behaviour • Interpersonal relations • Sentiments • Emotional experience • Communication skills • Attitude • Dedication • Assessment of social, personal, and interpersonal skills
Domain • Numbers, Mathematic question, Data driven, Calculations, Logic problems, Prediction tasks Type of data: quantitative	Domain • People, Personal question, Interpersonal skills, Employee evaluations Type of data: qualitative