Decision Making: AI advice in Performance Evaluation

Maria Pontífice Sousa

Dissertation written under the supervision of Professor Cristina Mendonça

Dissertation submitted in partial fulfilment of requirements for the MSc in Business, at the Universidade Católica Portuguesa, 28/05/2026.

Abstract

Title: Decision Making: AI advice in Performance Evaluation

Author: Maria Pontífice Sousa

Artificial intelligence (AI) is increasingly being used to support decision-making in organizational contexts. However, understanding how individuals respond to AI-generated advice and which psychological factors shape this process remains an important empirical question. This study aims to examine how individuals integrate AI-generated advice into performance evaluation decisions, focusing on the role of perceived usefulness, perceived ease of use, and trust in shaping advice-taking behavior. A scenario-based, between-subjects design was employed, in which participants evaluated a performance scenario before and after receiving AI-generated advice. Advice-taking was measured using the Weight of Advice (WOA). The results indicate that participants adjust their evaluations in response to AI-generated advice, although to a limited extent. Furthermore, perceived usefulness and perceived ease of use were positively associated with trust in AI, and trust was positively related to individuals' willingness to follow AI-generated recommendations. These findings contribute to the literature by integrating insights from the Technology Acceptance Model and research on algorithmic advice-taking, highlighting trust as a key mechanism linking user perceptions to behavioral responses. From a practical perspective, the results suggest that organizations should enhance the usability, perceived usefulness, and transparency of AI systems to foster trust and facilitate their effective adoption. Overall, this study provides relevant insights into human–AI interaction in performance evaluation contexts and offers a foundation for future research on AI-assisted decision-making in organizations.

Keywords: Artificial Intelligence, Advice-Taking, Trust in AI, Technology Acceptance Model (TAM), Performance Evaluation, Decision-Making

Resumo

Título: Tomada de Decisão: Conselho de IA em Avaliação de Desempenho

Autor: Maria Pontífice Sousa

A Inteligência Artificial (IA) tem vindo a ser cada vez mais utilizada para apoiar a tomada de decisão em contextos organizacionais. No entanto, compreender como os indivíduos respondem a conselhos gerados por IA e quais os fatores psicológicos que influenciam este processo permanece uma questão relevante. Este estudo analisa de que forma os indivíduos integram conselhos gerados por IA em decisões de avaliação de desempenho, com foco na utilidade percebida, facilidade de utilização e confiança. Foi utilizado um design entre sujeitos baseado em cenários, no qual os participantes avaliaram uma situação antes e depois de receberem aconselhamento gerado por IA. A aceitação do conselho foi medida através da Weight of Advice (WOA). Os resultados mostram que os participantes ajustam as suas avaliações em resposta ao conselho da IA, ainda que de forma limitada. A utilidade percebida e a facilidade de utilização apresentam associação positiva com a confiança, sendo esta positivamente relacionada com a disposição para seguir recomendações da IA. Estes resultados contribuem para a literatura ao integrar perspetivas do TAM com investigação sobre advice-taking, destacando a confiança como mecanismo central. Do ponto de vista prático, sugerem a importância de sistemas de IA percecionados como úteis, fáceis de utilizar e transparentes. Em suma, este estudo fornece contributos para a compreensão da interação humano–IA em contextos de avaliação de desempenho, abrindo caminho para investigação futura.

Palavras-chave: Inteligência Artificial (IA), Advice-Taking, Confiança na IA, Technology Acceptance Model (TAM), Avaliação de Desempenho, Tomada de Decisão

Table of Contents

Acknowledgements

Submitting this thesis means much more than “simply” earning an academic degree, because the truth is that it was a very challenging process, not only academically, but also because it took place at a critical time in my personal life. I am very happy, because this achievement marks the end of a very important chapter in my life – my academic life. The journey to this point has not always been easy, but it has enabled me to acquire many skills and has helped shape me into the person I am today.

First, I would like to thank Professor Cristina Mendonça for all her guidance, but above all, for every kind, supportive and encouraging word. She was always available to help me and answer my questions and was very understanding on the occasions when I fell short during this journey.

Secondly, I would like to thank my parents, my brother and the members of my family who are closest to me, for all the support and encouragement they have given me over these past few months. I thank them for listening to my frustrations and for lending me a helping hand when I needed it. I must also thank my closest friends, who were often the reason I kept going and never left me feeling alone.

Finally, I dedicate this achievement to my grandfather José Eurico, who, sadly, did not live to see me reach this day, but who asked me every day how my thesis was coming along and told me how enthusiastic he was to see this day arrive. He was a man who was utterly dedicated to his professional life, having achieved a level of success that many can only dream of. He was always a model of hard work and dedication. As a doctor, he devoted himself to saving and changing lives, always with a sense of purpose that he never lost. As a child, he taught me to swim, to ride a bike without training wheels and to ride a motorbike, and, as I grew older, he always encouraged me to be the best student and professional I could be. Today, I am grateful for the legacy of values he left me and I hope he is proud of this achievement of mine.

I acknowledge the support from FCT – Portuguese Foundation of Science and Technology for the project UID/GES/00407/2020 and for the project 2024.07497.IACDC, supported by “RE-C05-i08.M04 – “Apoiar o lançamento de um programa de projetos de I&D orientado para o desenvolvimento e implementação de sistemas avançados de cibersegurança, inteligência artificial e ciência de dados na administração pública, bem como de um programa de capacitação científica”, do Plano de Recuperação e Resiliência – PRR, framed within the financing agreement entered into between the Estrutura de Missão Recuperar Portugal (EMRP) and FCT, as intermediate beneficiary.

List of Figures

List of Abbreviations

AI	Artificial Intelligence
HR	Human Resources
HRM	Human Resources Management
TAM	Technology Acceptance Model
WOA	Weight of Advice

1. Introduction

1.1. Topic presentation

Over the past decades, Artificial Intelligence (AI) has become one of the most transformative forces shaping organizational processes, and the field of Human Resource Management (HRM) is no exception. AI-driven systems are increasingly embedded in organizational contexts, supporting human decision-making by providing insights and recommendations, a process often referred to as decision augmentation (Burton et al., 2020; Araujo et al., 2020). At the same time, these systems can enable decision automation, whereby AI systems make decisions independently without human intervention (Araujo et al., 2020).

Performance evaluation is one of the core HR activities and is increasingly impacted by technological developments, particularly AI-driven decision-support systems (Marler & Boudreau, 2017; Meijerink et al., 2021). As a process that influences promotions, rewards, career development opportunities, and perceptions of organizational justice, performance evaluation plays a critical role in organizational effectiveness and employee outcomes (Aguinis et al., 2013; DeNisi & Murphy, 2017). Importantly, traditional performance evaluation processes are often subject to biases, such as leniency bias and subjectivity, which may compromise decision quality (Aguinis et al., 2013). In this regard, AI systems have the potential to enhance objectivity and improve decision accuracy (Shrestha et al., 2019).

Therefore, the integration of AI in this domain raises critical questions about how HR managers interact with automated recommendations and how these shape their final decisions. As referred in Montealegre-López (2025), prior research suggests that AI can influence decision-making by providing faster, more accurate, and cost-efficient outputs, while also enhancing decision-makers' sensing and analytical capabilities, ultimately supporting more rational and evidence-based decisions (Shrestha et al., 2019). Importantly, the effectiveness of AI-generated recommendations depends heavily on human acceptance and the willingness to rely on algorithmic outputs (Logg et al., 2019; Glikson & Woolley, 2020)

In this context, a key question concerns the extent to which decision-makers accept and rely on AI-generated advice. Advice taking refers to the degree to which individuals adjust their initial judgments after receiving external recommendations (Bonaccio & Dalal, 2006). Prior research on advice taking suggests that decision-makers compare external recommendations with their initial judgments and adjust their final decisions accordingly, although this process

is often influenced by individual and contextual factors, such as expertise, confidence, and task characteristics (Bonaccio & Dalal, 2006; Mesbah et al., 2021).

The Technology Acceptance Model (TAM) provides a relevant theoretical framework to understand how individuals accept and use new technologies. TAM emphasizes perceived usefulness and perceived ease of use as key determinants of technological adoption (Davis, 1989). Additionally, TAM was initially developed to assess the impact of internal variables on actual use, referring to users' effective engagement with a given technology (Almeida et al., 2025).

Trust is a central factor in human–AI interaction, defined as the willingness to rely on an agent under conditions of uncertainty and vulnerability (Lee & See, 2004). Research shows that trust strongly influences the acceptance, adoption, and overall effectiveness of AI-driven decision-making (Castelo et al., 2019; Chiou & Lee, 2023). Moreover, trust is shaped by factors such as perceived competence, transparency, and task characteristics, being particularly lower in subjective tasks where individuals may doubt the ability of algorithms to capture contextual nuances (Castelo et al., 2019). As such, trust plays a key role in determining whether individuals rely on AI-generated advice.

Despite the potential benefits of AI, prior research suggests that individuals often exhibit a preference for human over algorithmic advice, a phenomenon commonly referred to as algorithm aversion (Dietvorst et al., 2015). This tendency may be particularly salient in subjective decision-making contexts, such as performance evaluation, where individuals may perceive algorithms as less capable of handling nuanced or complex judgments (Castelo et al., 2019). Consequently, this reluctance to rely on AI may limit the effective use of such systems and prevent organizations from fully leveraging their potential benefits. Thus, the increasing role of AI in performance evaluation demands a deeper understanding of how AI-generated advice affects human decision-making.

1.2. Problem statement & list of research questions

Although AI has become increasingly common in HRM processes, little is known about how individuals adjust their evaluations after receiving AI-generated advice in this context. According to prior research on human–human and human–AI advice taking, decision-makers compare external recommendations with their initial judgments and adjust their final decisions accordingly. This process is influenced by factors such as trust in the advisor, perceived expertise, and confidence-related cues (Bonaccio & Dalal, 2006; Mesbah et al., 2021).

Nevertheless, this dynamic in performance evaluation contexts has not received much empirical focus. Also, despite the fact that trust in automation has been extensively investigated through models like Mayer et al. (1995) and the Socio-Cognitive Model of Trust (Castelfranchi & Falcone, 2010), little is known about how trust specifically influences the integration of AI-generated advice in evaluative decisions.

Considering these gaps, the present thesis addresses the following questions:

RQ1: To what extent does AI-generated advice impact human decision-making in performance evaluation contexts?

RQ2: How are perceptions of usefulness and ease of use of AI systems related to trust in their recommendations?

RQ3: Does trust in AI systems explain the impact of AI-generated advice on human decision-making in performance evaluations?

1.3. Managerial & Academic relevance

The present thesis aims to contribute to three important streams in the current literature. First, it contributes to research on human–AI collaboration by providing empirical evidence on advice taking in a performance evaluation context. Second, it advances the literature on trust in automation by examining how trust operates and influences decision-makers' behavior in this context. Third, it expands Technology Acceptance Model (TAM) research by moving beyond intention-to-use and exploring the actual influence of AI on judgment tasks. Moreover, this thesis responds to the need for further research on the integration of AI into HR decision-making processes, particularly in the domain of performance evaluation.

For organizations, the results of the study of the present thesis will help to understand how evaluators respond to AI-generated advice, and this will be crucial to designing trustworthy and user-friendly AI systems. In turn, this may contribute to improving fairness and consistency in performance evaluations, as well as enhancing decision quality—outcomes that AI promises to support by reducing biases and increasing objectivity in evaluation processes. Furthermore, these insights can inform the development of more effective implementation strategies for AI systems in HR decision-making. As organizations increasingly integrate AI into their HR departments, such insights become essential for ensuring responsible and effective human–AI collaboration.

1.4. Short overview of the thesis' structure

This dissertation is organized into six chapters: Chapter 1, with the introduction, presents the overall research context, formulates the problem statement, defines the research questions, and discusses the academic and managerial relevance of the study.

Chapter 2 presents the literature review, beginning with a broad overview of Human Resource Management (HRM) and core HR processes, with particular emphasis on Performance Evaluation, given its central relevance to this research. It then deepens the discussion by examining Performance Management. The chapter proceeds to the field of AI, first clarifying its definition and organizational role, before narrowing the focus to AI in HR, highlighting its application in decision-support mechanisms and evaluation practices. It then examines the literature on AI in Decision Making and Advice Taking, analysing how human decision-makers adjust their judgments when exposed to algorithmic recommendations. This is followed by an in-depth discussion of trust in AI, including theoretical models and key determinants shaping trust in automated systems. Finally, the chapter presents the Technology Acceptance Model (TAM), with particular emphasis on perceived usefulness and perceived ease of use.

Chapter 3 describes the methodology in detail, including the between-subjects design (where participants were exposed to different AI advice conditions, but the study primarily focuses on testing correlational relationships between key variables), data collection procedures, measurement instruments, and analytic strategies employed, while Chapter 4 reports the empirical results and interprets them in relation to the proposed hypotheses.

Chapter 5 provides an integrated discussion of the findings, outlines their theoretical and practical implications, acknowledges the study's limitations, and proposes avenues for future research. Finally, Chapter 6 contains the conclusion, with a summary of the dissertation's overall contribution.

2. Literature Review

2.1. Human Resources Management (HRM)

Organizational objectives are closely associated with human resources management (HRM), which can be understood as both a body of knowledge and an organizational strategy. It encompasses a set of interconnected practices designed to improve employee performance, attitudes, and behaviours, thereby enhancing organizational effectiveness. From a systemic perspective, HRM practices operate in complementary "bundles", whose joint implementation has been shown to improve both individual and organizational outcomes (Ferreira et al., 2015).

Additionally, HRM is not exclusively the responsibility of HR specialists, but involves managers at all organizational levels and across different functional areas. Managers in departments such as operations, finance, and sales play a key role in implementing HR practices to foster motivation and alignment with organizational goals. Long-term performance is therefore influenced by employees' attitudes and behaviours, which are shaped through this shared managerial responsibility (Rego et al., 2008).

2.1.1. Human Resources Processes

2.1.2. Performance Evaluation

Performance management has become increasingly important in organizational contexts characterized by growing competitiveness and complexity. It is typically understood as the alignment between organizational outcomes and employees' contributions, with performance management and evaluation systems functioning as integrated tools that support strategic decision-making and organizational effectiveness (Ferreira et al., 2015).

Over time, performance management has evolved from a primarily administrative function to a more strategic and developmental process, emphasizing continuous feedback, goal alignment, and employee engagement. Effective performance evaluation systems contribute to organizational success by linking individual performance to strategic objectives and promoting fairness and transparency in decision-making processes (Murphy & Cleveland, 1995; DeNisi & Murphy, 2017).

Performance can be conceptualized from two complementary perspectives: behaviors and results (Ferreira et al., 2015). While the behavioral perspective focuses on how employees perform their tasks, the results perspective emphasizes the outcomes achieved. These approaches are not mutually exclusive, as effective performance management considers both dimensions depending on the nature of the role (Ferreira et al., 2015). Moreover, performance management systems serve both organizational and individual purposes: At the organizational level, they support decisions related to rewards and workforce management, while at the individual level they facilitate feedback, development, and career progression (Ferreira et al., 2015).

2.1.3. Biases in Performance Evaluation

Despite its importance, performance evaluation is widely recognized as a fundamentally imperfect process, largely due to its reliance on human judgment (Murphy & Cleveland, 1995;

Aguinis, 2013). The subjective nature of evaluation introduces a range of cognitive, social, and structural biases that can distort ratings and undermine decision quality (Murphy & Cleveland, 1995; Aguinis, 2013).

One of the most widely documented biases is the halo effect, whereby an evaluator's general impression of an employee—typically based on a salient positive attribute—extends to other unrelated performance dimensions (Murphy & Cleveland, 1995). Conversely, the horn effect occurs when a negative attribute disproportionately influences the overall evaluation, leading to overly negative assessments (Kromrei, 2015). Both biases reduce the accuracy of performance ratings by overshadowing a more nuanced assessment of individual performance dimensions (Murphy & Cleveland, 1995).

Additionally, evaluators often exhibit systematic tendencies toward leniency or severity, rating employees consistently higher or lower than warranted (Aguinis, 2013; Kromrei, 2015). Similarly, central tendency bias occurs when ratings cluster around the midpoint of the scale regardless of actual performance differences, limiting the ability of evaluation systems to discriminate effectively between levels of performance (Murphy & Cleveland, 1995; Kromrei, 2015).

Beyond these common biases, research highlights the role of temporal and cognitive biases in shaping performance evaluations. The primacy effect leads evaluators to form initial impressions early in the process and maintain them even when faced with contradictory evidence (Kromrei, 2015). In contrast, the recency effect results in disproportionate weighting of recent behaviors, while earlier performance information is discounted (Kromrei, 2015). In addition, confirmation bias may lead evaluators to selectively attend to information that supports their preconceived beliefs about an employee, reinforcing existing judgments (Kromrei, 2015).

Performance evaluations may also be influenced by comparative and social biases. The contrast effect occurs when employees are evaluated relative to others rather than against objective standards (Kromrei, 2015). The similar-to-me effect reflects a tendency to favor individuals who share characteristics with the evaluator, which may introduce systematic favoritism (Kromrei, 2015). More critically, biases related to illegal discrimination, whether conscious or unconscious, may affect evaluations based on characteristics such as gender, age, or ethnicity, raising serious concerns regarding fairness and organizational justice (Aguinis, 2013; Kromrei, 2015).

Finally, structural limitations further complicate the evaluation process. Evaluators may engage in truth avoidance, deliberately avoiding honest assessments to prevent conflict or discomfort (Kromrei, 2015). Additionally, opportunity bias arises when evaluators fail to account for contextual factors that influence performance, attributing outcomes solely to individual ability rather than situational constraints (Kromrei, 2015).

Taken together, these biases highlight the inherent limitations of traditional performance evaluation systems, which are susceptible to inconsistencies, subjectivity, and contextual distortions (Murphy & Cleveland, 1995; Aguinis, 2013). As a result, organizations have begun exploring alternative approaches, including the integration of data-driven and algorithmic tools, to enhance the accuracy and fairness of performance assessments (Marler & Boudreau, 2017).

2.2. Artificial Intelligence (AI)

Artificial Intelligence (AI) has emerged as one of the most transformative technologies in contemporary organizations, fundamentally reshaping how information is processed, decisions are made, and organizational processes are structured. AI refers to computational systems capable of performing tasks that typically require human intelligence, including learning, reasoning, and decision-making, through advanced algorithms and data-driven processes (Imad-dine Bazine, 2025).

In organizational contexts, AI is primarily used as a decision-support tool, enabling individuals to process information more efficiently and make more informed, data-driven decisions (Logg et al., 2019). Compared to human decision-makers, AI systems are often associated with higher consistency, scalability, and analytical capacity, particularly in environments characterized by complexity and uncertainty (Logg et al., 2019; Shrestha et al., 2019).

Rather than replacing human decision-makers, AI is commonly integrated into human–AI collaborative systems, where algorithmic outputs complement human judgment (Duarte, 2023). This reflects a shift from automation toward augmentation, in which AI enhances human capabilities rather than substituting them (Duarte, 2023).

2.2.1. AI in HR

The adoption of AI in HRM has grown significantly in recent years, transforming traditional HR processes and enabling organizations to adopt more efficient and data-driven practices. (Marler & Boudreau, 2017; Meijerink et al., 2021). AI technologies are increasingly

used in areas such as recruitment, performance management, workforce planning, and employee engagement, allowing organizations to process large volumes of data and generate actionable insights for decision-making processes (Logg et al., 2019).

Additionally, transparency and accountability remain critical concerns, as employees may question the fairness of decisions made by algorithmic systems (Marler & Boudreau, 2017). Consequently, the integration of AI in HR requires careful consideration of both technological capabilities and human factors, emphasizing the importance of a human-centered approach where AI supports, rather than replaces, human decision-making (Meijerink et al., 2021). As a result, the successful implementation of AI in HR depends not only on its technical capabilities but also on users' acceptance and trust in these systems (Leicht-Deobald et al., 2019).

The integration of AI into Human Resource Management extends beyond efficiency gains and introduces fundamental changes in how HR processes are designed and experienced. In the context of performance management, AI systems are increasingly used to support evaluation processes by aggregating performance data, identifying patterns, and providing standardized assessments (Meijerink et al., 2021; Marler & Boudreau, 2017). These systems have the potential to enhance consistency and reduce individual biases, thereby improving decision quality.

However, the use of AI in performance evaluation also raises important concerns related to fairness, transparency, and accountability. Employees and managers may question how algorithmic decisions are generated, particularly when the underlying models are complex and difficult to interpret (Leicht-Deobald et al., 2019). This opacity can lead to perceptions of procedural injustice, even when the outcomes are objectively accurate.

Additionally, the introduction of AI in evaluative processes may alter the social dynamics of HR decision-making. Performance evaluation is traditionally a relational process involving feedback, communication, and interpersonal judgment. The involvement of AI may be perceived as depersonalizing this process, potentially reducing perceived empathy and increasing resistance to algorithmic input (Meijerink et al., 2021; Leicht-Deobald et al., 2019).

Moreover, concerns about algorithmic bias remain central. While AI is often promoted as a tool to reduce human bias, it may inadvertently reproduce or amplify existing biases present

in training data (Leicht-Deobald et al., 2019). This creates a paradox in which AI is both a potential solution and a potential risk in ensuring fair performance evaluations.

As a result, the successful implementation of AI in HR processes depends not only on its technical capabilities but also on how it is perceived and accepted by users (Leicht-Deobald et al., 2019; Meijerink et al., 2021). This reinforces the importance of understanding psychological factors such as trust and technology acceptance, which ultimately shape how AI-generated advice is interpreted and used in decision-making (Glikson & Woolley, 2020; Araujo et al., 2020).

In this regard, understanding how individuals respond to AI-generated advice in evaluative contexts becomes particularly important, as it reflects how these systems are likely to be used in real HR decision-making situations.

2.2.2. AI in Decision-making and Advice taking

A key aspect of human–AI interaction lies in how individuals incorporate algorithmic recommendations into their decision-making processes. Research on advice taking suggests that individuals integrate external inputs by adjusting their initial judgments, a process commonly measured through the Weight of Advice (WOA) (Yaniv, 2004; Bailey et al., 2023).

The WOA reflects the extent to which individuals revise their initial decisions in response to external advice, ranging from complete disregard of the advice to full adoption of the recommendation. In organizational contexts, advice taking is particularly relevant when decisions involve uncertainty, complexity, or subjective judgment, as is often the case in performance evaluations (Bonaccio & Dalal, 2006; Yaniv, 2004).

When advice is generated by AI systems, decision-makers may evaluate it differently compared to human advice, depending on their perceptions of the system's competence, reliability, and legitimacy. These perceptions play a crucial role in determining whether AI-generated recommendations are accepted or rejected (Logg et al., 2019).

Given the literature presented, it is expected that individuals will incorporate AI-generated recommendations into their performance evaluations, leading to observable adjustments in their judgments - **H1:** Participants will significantly adjust their performance evaluations after receiving AI-generated advice.

2.2.3. Trust in AI

Trust is widely recognized as a key determinant of human–AI interaction, as it influences the extent to which individuals are willing to rely on algorithmic recommendations. In organizational contexts, trust can be defined as the willingness to accept vulnerability based on the expectation that a system will perform reliably and act in the user's best interest (Mayer et al., 1995; Glikson & Woolley, 2020). However, the literature emphasizes that the objective is not to maximize trust, but rather to achieve appropriately calibrated trust, ensuring that users rely on AI systems only when warranted by their reliability and accuracy (Duarte, 2023).

In the context of AI, trust plays a crucial role because users often lack full understanding of how algorithms operate. This creates uncertainty, making trust a necessary mechanism for enabling reliance on AI systems. Without sufficient trust, individuals may disregard algorithmic advice, even when it is objectively accurate (Glikson & Woolley, 2020).

An important concept to be noted is trust calibration, which refers to the alignment between users' trust and the actual performance of the system. Miscalibrated trust can take two forms: over-reliance, where users blindly follow incorrect recommendations, and under-reliance, where users reject valid advice. Both outcomes can lead to suboptimal decisions (Jussupow et al., 2024).

Empirical evidence consistently shows that system performance is the strongest predictor of trust, with higher-performing systems generating significantly higher levels of perceived trust, understanding, and perceived quality (Duarte, 2023). Importantly, trust in AI is not determined solely by technical performance but is also shaped by social and psychological perceptions.

Beyond these general determinants, trust in AI can be further understood as a multidimensional construct, drawing on classical models of interpersonal trust. According to Mayer et al. (1995), trust is based on perceptions of competence, benevolence, and integrity. This framework has been adapted to the context of AI, where competence refers to the system's ability to generate accurate and reliable outputs, integrity relates to adherence to consistent and fair principles, and benevolence reflects the extent to which the system is perceived as acting in the user's best interest, even if indirectly through its design (Glikson & Woolley, 2020).

Importantly, these dimensions may not be equally salient in human–AI interactions. While competence tends to be the dominant driver of trust in AI systems, as users prioritize accuracy and performance, benevolence and integrity are more difficult to attribute to

algorithmic agents, given their non-human nature (Glikson & Woolley, 2020). This creates a fundamental asymmetry between trust in humans and trust in AI, where relational and emotional components are less influential, and cognitive evaluations assume greater importance.

Furthermore, the role of trust becomes particularly complex in subjective decision-making contexts, such as performance evaluation. Unlike objective tasks, where correctness can be easily verified, subjective evaluations require interpretation, contextual awareness, and sensitivity to nuanced behavioral cues. In such contexts, individuals may perceive AI systems as less capable of capturing qualitative aspects of performance, which can reduce trust even when the system demonstrates high technical accuracy (Castelo et al., 2019).

This limitation highlights the importance of contextual trust, as trust in AI is not only a function of system characteristics but also of task characteristics (Glikson & Woolley, 2020; Castelo et al., 2019). Research suggests that trust tends to be lower in domains requiring human judgment, emotional intelligence, or moral reasoning (Glikson & Woolley, 2020; Castelo et al., 2019). Consequently, in HR processes such as performance evaluation, building trust requires not only demonstrating technical competence but also addressing users' concerns about fairness, transparency, and contextual understanding.

Trust in AI is influenced by multiple factors, including perceived system performance, transparency, explainability, and fairness (Glikson & Woolley, 2020). For instance, systems that provide clear explanations for their recommendations are generally perceived as more trustworthy, although excessive complexity may reduce understanding and confidence (Glikson & Woolley, 2020).

Taken together, these insights suggest that trust in AI is a dynamic and context-dependent process, shaped by both system-related and task-related factors (Glikson & Woolley, 2020; Castelo et al., 2019). These findings reinforce the idea that trust in AI is a central mechanism shaping whether individuals incorporate AI-generated advice into their judgments, particularly in complex and subjective decision-making contexts. In this sense, examining how individuals respond to AI-generated advice in performance evaluation contexts provides valuable insights into how these systems may be effectively implemented in real organizational settings.

Based on the literature, in this thesis, it is expected that trust in AI will increase individuals' willingness to adjust their evaluations in line with AI-generated advice: **H2:** Higher levels of trust in the AI system will strengthen the influence of AI-generated advice on participants' performance evaluations.

2.2.4. Algortitm aversion

Despite the potential benefits of AI, research consistently shows that individuals may exhibit resistance to algorithmic decision-making, a phenomenon known as algorithm aversion. This refers to the tendency to prefer human judgment over algorithmic recommendations, even when algorithms demonstrate superior performance (Dietvorst et al., 2015; Jussupow et al., 2024).

One key driver of algorithm aversion is the perception that algorithms are less capable of handling complex, subjective, or context-dependent tasks. For example, individuals are less likely to rely on AI in domains requiring interpersonal sensitivity, such as performance evaluation, where human judgment is perceived as more appropriate (Castelo et al., 2019).

Additionally, studies have shown that individuals are particularly sensitive to algorithmic errors. Even minor mistakes may lead to a disproportionate loss of trust in AI systems, resulting in reduced reliance on algorithmic advice (Dietvorst et al., 2015). However, algorithm aversion is not universal, and some research suggests the existence of algorithm appreciation, where individuals prefer algorithmic advice under certain conditions, particularly when accuracy and objectivity are emphasized (Logg et al., 2019). This duality highlights the complexity of human–AI interaction and underscores the importance of understanding the factors that influence AI acceptance.

Although algorithm aversion has been widely documented, more recent research suggests that individuals do not always reject algorithmic advice but may instead exhibit algorithm appreciation under certain conditions (Logg et al., 2019). This duality highlights that individuals' responses to AI are not inherently negative but depend on specific contextual and psychological factors (Logg et al., 2019; Jussupow et al., 2024).

One key factor influencing this dynamic is the degree of control users have over the decision-making process. Studies show that allowing individuals to retain some level of autonomy, such as the possibility to adjust or override algorithmic recommendations, can significantly reduce algorithm aversion and increase reliance on AI (Dietvorst et al., 2015). This

sense of control helps mitigate concerns about loss of agency and increases users' willingness to engage with algorithmic advice.

Another important determinant is explainability. When AI systems provide clear and understandable explanations for their recommendations, users are more likely to perceive them as transparent and trustworthy, which can shift attitudes from aversion to appreciation (Glikson & Woolley, 2020). However, this relationship is not always linear, as overly complex explanations may reduce comprehension and lead to confusion (Duarte, 2023).

Task characteristics also play a critical role. As previously noted, individuals are more likely to rely on algorithms in objective and data-driven tasks, where accuracy is easily assessed, whereas in subjective domains, such as performance evaluation, they may prefer human judgment (Castelo et al., 2019). This suggests that algorithm aversion is not a general resistance to AI, but rather a context-dependent response shaped by perceived task suitability.

In the context of this study, performance evaluation represents a particularly relevant setting, as it involves subjective judgment and evaluative discretion (Murphy & Cleveland, 1995; Aguinis, 2013). As such, individuals may experience tension between recognizing the analytical advantages of AI and questioning its ability to fully capture human performance (Castelo et al., 2019; Glikson & Woolley, 2020). This tension helps explain why individuals may partially rely on AI advice, rather than fully adopting or rejecting it, a pattern that is consistent with prior findings on algorithm aversion and appreciation in decision-making contexts (Dietvorst et al., 2015; Logg et al., 2019).

2.3. Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) is one of the most influential theoretical frameworks for understanding how individuals adopt and use new technologies. Originally developed by Davis (1989), TAM posits that two core perceptions—perceived usefulness and perceived ease of use—are the primary determinants of users' attitudes toward a technology, which in turn influence their behavioral intention to use it.

TAM has been widely applied across various organizational contexts, including information systems, e-commerce, and, more recently, AI-driven technologies. (Venkatesh & Davis, 2000). Its simplicity and strong empirical support have made it a dominant framework in the study of technology adoption (Venkatesh & Davis, 2000).

Furthermore, researchers have argued that TAM needs to be extended to account for additional factors, such as trust, perceived fairness, and transparency, which play a critical role in shaping users' responses to AI systems (Almeida et al., 2025). These factors suggest that TAM may need to be adapted to account for the unique characteristics of AI systems and their impact on user behavior (Tuffaha, 2023).

While the Technology Acceptance Model (TAM) provides a robust framework for understanding technology adoption, its application to AI systems requires further theoretical extension. Traditional TAM focuses on behavioral intention to use technology, whereas AI-driven systems often operate as decision-support tools that directly influence judgments and decisions (Venkatesh & Davis, 2000; Shrestha et al., 2019; Baines et al., 2024).

In this context, trust emerges as a critical mechanism linking technology perceptions to behavioral outcomes. Several studies argue that perceived usefulness and perceived ease of use play a fundamental role in shaping trust in AI systems (Glikson & Woolley, 2020; Almeida et al., 2025). When users perceive a system as useful, meaning that it enhances their performance, they are more likely to view it as competent and reliable. Similarly, systems that are easy to use reduce cognitive effort and increase users' sense of control, which can further strengthen trust.

From this perspective, trust can be understood as an extension of the Technology Acceptance Model, acting as a mediating mechanism between system perceptions and actual reliance on AI-generated recommendations (Glikson & Woolley, 2020; Almeida et al., 2025). This is particularly relevant in decision-making contexts, where the key outcome is not merely usage, but the extent to which individuals incorporate AI advice into their judgments (Logg et al., 2019; Baines et al., 2024).

However, the directionality of these relationships remains subject to debate. While the Technology Acceptance Model suggests that perceived usefulness and perceived ease of use drive trust, it is also plausible that trust influences how users perceive a system's usefulness and usability (Glikson & Woolley, 2020; Venkatesh & Davis, 2000). For instance, individuals who already trust a system may be more inclined to interpret its outputs as useful and easy to understand. This suggests the possibility of bidirectional relationships between these constructs (Ahn et al., 2024; Küper & Krämer, 2024).

Given this ambiguity, the relationships between perceived usefulness, perceived ease of use, and trust should be interpreted as theoretically grounded but not definitively causal. This

highlights the need for further experimental research that explicitly manipulates these variables to clarify their causal ordering and better understand their role in human–AI interaction (Araujo et al., 2020; Montealegre-López, 2025).

2.3.1. Perceived Usefulness and Perceived Ease of Use

Perceived usefulness and perceived ease of use are central determinants of technology adoption in TAM (Davis, 1989). Perceived usefulness refers to the degree to which an individual believes that using a technology enhances their job performance, while perceived ease of use refers to how easy the technology is to understand and use (Tuffaha, 2023).

Empirical research in HRM and organizational contexts shows that AI enhances productivity, improves decision quality, and supports strategic processes, reinforcing perceptions of usefulness (Imad-dine Bazine, 2025).

Perceived usefulness is particularly relevant given the complexity of AI systems, as users are more likely to accept AI systems that improve decision-making accuracy and efficiency (Glikson & Woolley, 2020). Similarly, perceived ease of use influences adoption by reducing cognitive effort and increasing user confidence when interacting with complex systems (Glikson & Woolley, 2020). However, previous research also indicates that explanations do not always produce the intended effects, as they may lead to overconfidence or misinterpretation (Duarte, 2023).

It is expected that more positive perceptions of AI systems will be associated with higher levels of trust in their recommendations – **H3:** Perceived usefulness and perceived ease of use of the AI system will be positively associated with trust in its recommendations.

3. Methodology

This thesis' study applied a between-subjects design to examine how AI-generated influences human decision-making in performance evaluation contexts. Although participants were exposed to different AI advice conditions, the study primarily focuses on testing correlational relationships between key variables (perceived usefulness, perceived ease of use, trust in AI.

The design was structured to address the three research questions by (1) manipulating the AI advice condition (negative, neutral, positive), (2) measuring perceived usefulness and perceived ease of use, and (3) assessing participants' trust in AI, alongside their performance

evaluations. It should also be noted that the study consisted of two phases – a pre-AI advice phase and a post-AI advice phase.

Furthermore, perceived usefulness, perceived ease of use, and trust in AI played an important role in this analysis, as they help explain how participants respond to AI-generated advice. Drawing on the Technology Acceptance Model (Davis, 1989), perceived usefulness and perceived ease of use were included as key dimensions capturing participants' evaluations of the AI system.

In turn, perceived usefulness was included as a key variable in the analysis, reflecting the extent to which participants perceived the AI system as enhancing their decision-making, and its potential association with trust in AI. While perceived ease of use may also be related to trust, this relationship is typically weaker (Ahn et al., 2024). Finally, trust in AI is positively associated with advice-taking behaviour, as higher levels of trust are linked to a greater reliance on AI recommendations (Ahn et al., 2024; Bailey et al., 2022). Overall, these variables were included to capture different aspects of participants' responses to AI-generated advice.

A between-subjects approach was chosen due to its ability to support causal inference regarding the effect of AI-generated advice on participants' evaluations. The participants were exposed to different AI advice conditions (negative, neutral and positive), but the study primarily focuses on testing correlational relationships between key variables.

Previous research shows that external informational signals can influence competence judgments even when such signals are not objectively relevant, highlighting individuals' susceptibility to external input during evaluative decision-making (Stellar & Willer, 2018). In this context, AI-generated advice serves as the stimulus that may shape participants' performance evaluations.

In contrast, perceived usefulness, perceived ease of use, trust in AI, and weight of advice (WOA) were not experimentally manipulated but measured as individual responses, and the relationships among these variables are therefore interpreted as correlational.

3.1. Participants

As regards the validated participants, only 217 of the 293 in the database met the inclusion criteria for completing the questionnaire and passing the attention check question. Consequently, the entire analysis relates to the 217 respondents mentioned (see Appendix 1 for the frequency tables regarding participant removal and validation).

In addition, regarding descriptive demographic statistics, the sample was predominantly composed of participants from Portugal (96.8%), the majority of whom reported Portuguese as their first language (96.8%). In terms of gender distribution, 56.2% of participants identified as female and 43.8% as male. Participants were relatively diverse in age, with the largest group aged between 18 and 24 years (37.3%). In terms of education levels, most participants held either a bachelor's degree (39.6%) or a master's degree (38.7%) (see Appendix 2 for the frequency tables of the study's categorical variables).

Participants were volunteers recruited through social channels, including social media platforms such as Instagram, LinkedIn, WhatsApp groups, and other personal networks. Participation was anonymous and uncompensated. Prior to beginning the study, participants read and agreed with an informed consent form, which described the purpose, expected duration, confidentiality, and right to withdraw at any point.

3.2. Materials

The main material consisted of a performance evaluation scenario which centered around a two and half minutes video about the concept of supply and demand. The video included both audio and visual components, consisting of a speaker verbally explaining the concept supported by visual content. Participants had to evaluate the explanation provided in the video along four dimensions (clarity, technical accuracy, structure and overall competence, taken from the study of Harari & Zedeck, 1973), on a scale from 1*(very poor)* to 7 *(very good)*, which was chosen based on input from existing literature that argues that, in behavior-focused performance dimensions, evaluations grounded in behaviorally anchored criteria reduce ambiguity and increase reliability by focusing raters on concrete observable aspects of performance (Harari & Zedeck, 1973).

Furthermore, three AI feedback conditions were developed, using an AI system fictionally named TalentScope: a negative condition with uniformly low scores (more specifically, clarity of explanation: 1/7, technical accuracy: 2/7, structure and organization: 2/7, overall competence: 1/7), a neutral advice condition with moderately positive scores (more specifically, clarity of explanation: 4/7, technical accuracy: 5/7, structure and organization: 3/7, overall competence: 4/7), and a positive advice condition, with uniformly high scores (specifically, clarity of explanation: 7/7, technical accuracy: 6/7, structure and organization: 6/7, overall competence: 7/7) – these feedback conditions constitute the manipulation of AI advice, allowing the examination of whether participants' perceptions of usefulness, ease of

use, and trust vary as a function of the valence of the advice (i.e., whether the feedback was negative, neutral, or positive) provided.

TAM-based scales (Davis, 1989), which ranged from 1 (*strongly disagree*) to 7 (*strongly agree*), were also included in the study, more specifically, the scales of perceived usefulness and perceived ease of use. Also, trust in AI was measured using multiple items assessing participants' confidence in and reliance on the AI system with the same likert scale (where 1 = *strongly disagree* and 7 = *strongly agree*). Some examples of these items include: "TalentScope helped me make a more informed decision" (for perceived usefulness), "It was easy to understand TalentScope's advice" (for perceived ease of use) and "I trust that TalentScope analyzes data fairly" (trust in AI).

In addition, AI familiarity was assessed through self-reported measures of participants' knowledge, familiarity and frequency of use of AI systems, with a scale ranged from one to ten. The questions included in this part were: "How would you rate your knowledge of Artificial Intelligence on a scale of 1 *(extremely poor)* to 10 *(extremely good)*?"; "On a scale of 1 *(not very familiar)* to 10 *(extremely familiar)*, how familiar are you with the use of AI?"; and "How often do you use AI systems on a scale from 1 *(I never used them, or only used to try it)* to 10 *(I use it almost every day)*?"

Also, there were several demographic questions asked to the participants: nationality (with a list of countries), first language (where the options were Portuguese, English, and other, and this last option required the specification on the language), gender (with the options male, female and "prefer not to say"), age (with ages ranged between 18 and above 65), and education level (from less than high school to doctoral degree, with the "other" option, which required specification).

3.3. Procedure

The experiment started with an informed consent, explaining the objectives of the study and giving participants the explicit opportunity to decide whether to participate or withdraw from the study, which they could also do at any time. There were some introductory demographic questions, asking the participants where they were from and which was their first language.

Secondly, the scenario was presented – which consisted of the video with a concept explanation. After that, participants had the opportunity to evaluate the performance of the

instructor using the performance scales (in order to rate clarity, technical accuracy, structure and global competence) – followed by an attention check-question, where the respondents were asked to select "strongly disagree", from the following set of options: Strongly disagree; Disagree; Neither agree nor disagree; Agree; Strongly agree.

After that, participants were confronted with one of the three AI feedback conditions, generated by TalentScope. Participants were then given the chance to change/correct their first performance evaluation using the same performance scales used before seeing the advice. Consequently, the main dependent variable is the weight of advice (WOA). This was calculated using the following formula: (final estimate – initial estimate) / (advice – initial estimate; Bailey et al., 2023; Yaniv, 2003). This formula was applied separately to each of the four evaluation criteria assessed by participants (clarity, technical accuracy, structure, and overall competence), resulting in four distinct WOA scores. To obtain an overall measure of advice-taking, an aggregated WOA variable was then computed as the average of these four scores (see Appendix 3 for detailed descriptive statistics of the WOA variables).

Next in the questionnaire was manipulation check-question, which asked the participant about their own assessment of the AI system, and they were asked to rate it using one of three options: negative, neutral, or positive. Towards the end of the study, participants were asked the TAM-based and AI familiarity scales. Finally, demographic questions were presented at the end of the survey, collecting gender, age range and education level, as well as a final optional feedback text box.

4. Results

The internal consistency of the study measures was assessed using Cronbach's α. All constructs demonstrated good reliability, with values exceeding the recommended threshold of .70 (Bland & Altman, 1997), specifically, WOA (α = .83), while trust in AI (α = .97) and perceived usefulness (α = .96) exhibited a higher internal consistency. The AI-related familiarity scale also showed satisfactory reliability (α = .87), supporting the use of these measures in subsequent analyses (see Appendix 4 for the output of the Cronbach's α analyses).

Bivariate correlations were conducted to examine the relationship between WOA and potential control variables. In this study, gender, country, and age were included as control variables. The results indicated that WOA was not significantly associated with gender (r = −.02, p = .738), country (r = .02, p = .797), or age (r = −.12, p = .091), suggesting that

demographic factors did not meaningfully influence participants' responses (see Appendix 5 for the output of the bivariate correlations and with all control variables).

After that, in order to test **H1**, which proposed that participants would significantly adjust their performance evaluations after receiving AI-generated advice, a one-sample *t*-test was conducted, with 0 as the test value. The results revealed a statistically significant effect, $t(216) = 2.46$, $p = .015$, with a mean difference of 0.14 and a 95% confidence interval ranging from 0.03 to 0.25. The effect size was small (Cohen's $d = 0.17$). These findings indicate that participants exhibited a small but significant tendency to adjust their evaluations in line with AI-generated advice, thereby providing support for **H1** (see Appendix 6 for the output of the t-test analysis - One Sample T-test: H1).

To test **H3**, which predicted that perceived usefulness and perceived ease of use would be positively associated with trust in the AI system, a multiple linear regression analysis was conducted with trust in AI as the dependent variable and perceived usefulness and perceived ease of use as predictors. The model was statistically significant, $F(2, 214) = 709.09$, $p < .001$, explaining 86.9% of the variance in trust in AI. Both perceived usefulness ($\beta = .72$, $p < .001$, see Figure 1) and perceived ease of use ($\beta = .25$, $p < .001$, see Figure 2) were significant positive predictors of trust in AI. This result provides strong support for **H3**, indicating that both perceived usefulness and perceived ease of use are positively associated with trust in AI recommendations (eee Appendix 7 for the Multiple Linear Regression analysis for H3).

Finally, to test **H2**, which suggested that higher levels of trust in the AI system would strengthen the influence of AI-generated advice on performance evaluations, a linear regression analysis was conducted with trust in AI as a predictor of willingness to accept AI advice. The model was statistically significant, $F(1, 215) = 4.77$, $p = .030$, although it explained a small proportion of the variance ($R^2 = .022$). Trust in AI was a significant positive predictor of willingness to accept AI advice ($\beta = .150$, $p = .030$). These findings suggest that higher levels of trust in the AI system are associated with a greater likelihood of adjusting performance evaluations in line with AI advice, providing support for **H2**, despite the small effect size (see Appendix 8 for the Linear regression analysis for H2). These relationships are further illustrated in Figures 1–3, presented below.

Figure 1: Relationship between perceived usefulness and trust in AI

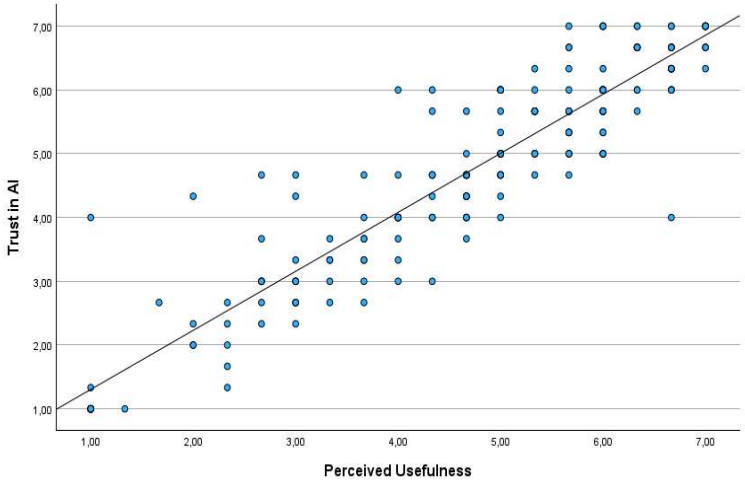


Figure 2: Relationship between perceived ease of use and trust in AI

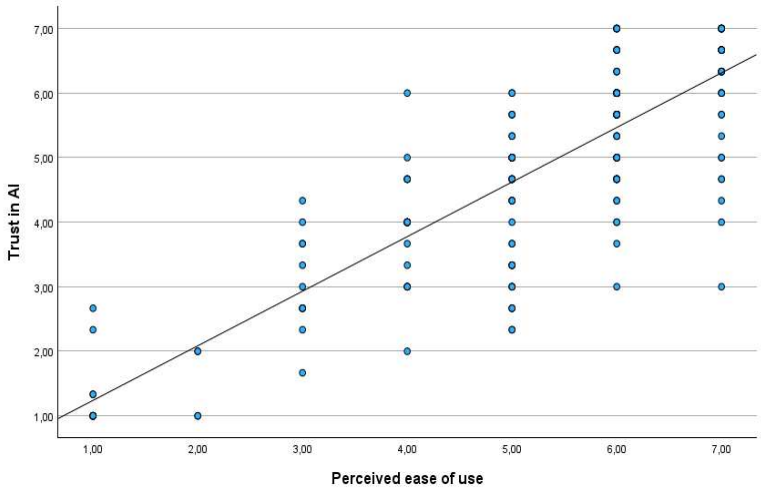
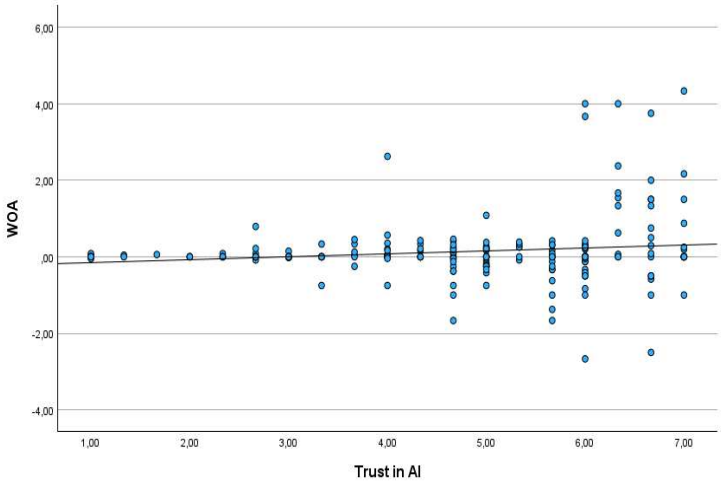


Figure 3: Relationship between trust in AI and WOA



5. Discussion

The present study aimed to examine how individuals integrate AI-generated advice into decision-making processes, focusing on the role of perceived usefulness, perceived ease of use, and trust in shaping willingness to accept AI recommendations in a performance evaluation context.

Overall, the findings provide consistent support for the proposed model and contribute to a growing body of research on human–AI interaction in organizational decision-making. First, the results showed that participants exhibited a small but significant tendency to adjust their performance evaluations after receiving AI-generated advice, supporting H1. This finding suggests that individuals are willing to incorporate algorithmic input into their judgments.

This result is in line with prior research on algorithmic advice-taking (Dietvorst et al., 2015; Logg et al., 2019), which shows that individuals tend to adjust their judgments in response to algorithmic recommendations, althought to a limited extent. The relatively small effect size observed in this study further aligns with the algorithm aversion/appreciation literature, reinforcing the idea that individuals incorporate AI input into their decisions without fully deferring to it. In the context of HR decision-making, this partial reliance may reflect the perceived responsibility and consequences associated with evaluating individuals' performance.

Regarding H3, the findings showed strong support for the proposed relationships between perceived usefulness, perceived ease of use, and trust in AI. Specifically, both perceived usefulness and perceived ease of use were significant positive predictors of trust in the AI system. These results are highly consistent with the Technology Acceptance Model (TAM; Davis, 1989), reinforcing the idea that users' perceptions of a system's usability and utility are fundamental drivers of trust in AI-based decision support tools.

Finally, the results provided support for H2, indicating that higher levels of trust in AI were associated with a greater willingness to adjust performance evaluations in line with AI advice. Although the effect size was relatively small, this finding highlights the role of trust as a key psychological mechanism underlying AI advice-taking (Mayer et al., 1995). In line with prior literature (Glikson & Woolley, 2020) trust appears to function as an enabling factor that increases individuals' openness to algorithmic input, thereby strengthening the influence of AI recommendations on decision-making outcomes.

5.1. Implications

From a theoretical perspective, this research contributes to the literature on AI in HR by integrating insights from the Technology Acceptance Model (Davis, 1989) with research on algorithmic advice-taking. In particular, it highlights trust as a central mechanism linking system perceptions (i.e., usefulness and ease of use) to behavioral outcomes, such as the acceptance of AI advice. This is consistent with prior research emphasizing trust as a key determinant of individuals' willingness to rely on automated systems (Mayer et al., 1995; Glikson & Woolley, 2020). This contributes to a more comprehensive understanding of how individuals interact with AI systems in organizational contexts.

From a practical standpoint, the results suggest that organizations aiming to implement AI-based decision support systems should prioritize not only the technical performance of these systems but also how they are perceived by users. Ensuring that AI tools are easy to use and perceived as useful can significantly enhance trust, which in turn increases the likelihood that employees will follow AI-generated recommendations.

More specifically, prior research suggests that organizations can enhance perceived usefulness, perceived ease of use, and trust by improving the transparency and explainability of AI systems, providing clear and understandable justifications for algorithmic recommendations (Glikson & Woolley, 2020), and allowing users to retain a certain level of control over the decision-making process, for instance by enabling them to adjust or override AI suggestions (Dietvorst et al., 2015). Additionally, designing user-friendly interfaces and reducing the complexity of interactions with AI systems can further support ease of use and foster more positive user perceptions.

Furthermore, it is also important to note that this study contributes to the integration of AI systems into an underexplored area of HR contexts, namely performance evaluation. This is particularly relevant in HR contexts, where resistance to algorithmic decision-making may arise due to concerns about fairness, accountability, or lack of transparency (Glikson & Woolley, 2020; Leicht-Deobald et al., 2019).

5.2. Limitations and Future Research

Despite its contributions, this study presents several limitations that should be acknowledged. First, the sample was predominantly composed of participants from a single country, which may limit the generalizability of the findings. Cultural factors may influence attitudes toward AI and trust in automated systems, as prior research shows that perceptions of AI vary across social and contextual environments (Glikson & Woolley, 2020; Araujo et al., 2020). Future research should examine more diverse samples.

Second, the study relied on a scenario-based between-subjects design, which, while useful for controlling variables, may not fully capture the complexity of real-world HR decision-making. Participants' responses may differ in actual organizational settings. Future research could address this limitation by conducting studies in real organizational environments, allowing for a more valid assessment of how individuals respond to AI-generated advice in performance evaluation contexts, as prior research highlights the importance of field studies for enhancing external validity (Aguinis & Bradley, 2014).

Third, although trust in AI was found to predict willingness to accept AI advice, the effect size was relatively small. This suggests that other factors not included in the model—such as perceived fairness, transparency, or individual differences—may also play an important role in shaping AI advice-taking behavior (Glikson & Woolley, 2020; Leicht-Deobald et al., 2019). Future research should explore these additional variables to develop a more comprehensive model. Additionally, although the study included different AI advice conditions, the potential effect of advice valence on advice-taking behavior was not explicitly examined. Future research could explore this aspect by directly.

Finally, an additional limitation relates to the correlational nature of the relationships examined in this study, which does not allow for definitive conclusions regarding causality. Although the findings suggest that perceived usefulness and perceived ease of use are associated with trust in AI, and that trust in turn is linked to willingness to accept AI advice, the direction of these relationships cannot be conclusively determined.

While the Technology Acceptance Model (TAM) traditionally assumes that perceived usefulness and perceived ease of use influence user responses, it is also plausible that trust in the system may shape individuals' perceptions of its usefulness and ease of use. For instance, individuals who trust an AI system may be more likely to perceive it as useful and easy to use, just as positive experiences with its usefulness and ease of use may foster greater trust.

These findings highlight the possibility of bidirectional relationships among these variables. Future research should address this limitation by experimentally manipulating trust, perceived usefulness, and perceived ease of use to better understand the causal direction and potential bidirectionality of these relationships.

6. Conclusion

In conclusion, this study provides evidence that individuals are willing to adjust their decisions in response to AI-generated advice, although to a limited extent. It further

demonstrates that trust in AI plays a key role in this process and is strongly influenced by perceived usefulness and ease of use. By integrating perspectives from technology acceptance and decision-making research, this study contributes to a better understanding of how AI systems are adopted and used in HR contexts. As organizations increasingly rely on AI-driven tools, understanding the factors that shape trust and advice-taking will become essential for ensuring effective and responsible implementation.

References

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351–371. https://doi.org/10.1177/1094428114547952

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Avoiding a "me" versus "we" dilemma: Using performance management to turn teams into a source of competitive advantage. *Business Horizons*, *56*(4), 503–512. https://doi.org/10.1016/j.bushor.2013.02.004

Ahn, D., Abdullah Almaatouq, Monisha Gulabani, & Kartik Hosanagar. (2024). *Impact of Model Interpretability and Outcome Feedback on Trust in AI*. https://doi.org/10.1145/3613904.3642780

Almeida, F., Junça Silva, A., Lopes, S. L., & Braz, I. (2025). Understanding Recruiters' Acceptance of Artificial Intelligence: Insights from the Technology Acceptance Model. *Applied Sciences*, *15*(2), 746. https://doi.org/10.3390/app15020746

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). https://doi.org/10.1007/s00146-019-00931-w

Bailey, P. E., Leon, T., Ebner, N. C., Moustafa, A. A., & Weidemann, G. (2023). A meta-analysis of the weight of advice in decision-making. *Current Psychology*, *42*(28), 24516–24541. https://doi.org/10.1007/s12144-022-03573-2

Baines, J. I., Dalal, R. S., Ponce, L. P., & Tsai, H.-C. (2024). Advice from artificial intelligence: a review and practical implications. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1390182

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's Alpha. *BMJ*, *314*(7080), 572–572.

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. https://doi.org/10.1016/j.obhdp.2006.07.001

Burton, J. W., Stein, M., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, *33*(2). https://doi.org/10.1002/bdm.2155

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, *56*(5), 809–825.

Chiou, E. K., & Lee, J. D. (2023). Trusting Automation: Designing for Responsivity and Resilience. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *65*(1), 001872082110099. https://doi.org/10.1177/00187208211009995

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

DeNisi, A. S., & Murphy, K. R. (2017). Performance Appraisal and Performance management: 100 Years of progress? *Journal of Applied Psychology*, *102*(3), 421–433. https://doi.org/10.1037/apl0000085

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.

Duarte, R. (2023). Towards Responsible AI: Developing Explanations to Increase Human-AI Collaboration. *Frontiers in Artificial Intelligence and Applications*. https://doi.org/10.3233/faia230126

Ekaterina Jussupow, Izak Benbasat, & Heinzl, A. (2024). An Integrative Perspective on Algorithm Aversion and Appreciation in Decision-Making. *MIS Quarterly*. https://doi.org/10.25300/misq/2024/18512

Ferreira, A. I., Martinez, L. F., Nunes, F. G., & Duarte, H. (2015). GRH em Portugal: evolução concetual, investigação e aplicação. *GRH para Gestores* (1st ed., pp. 42-45). Editora RH.

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, *14*(2), 627–660.

Harari, O., & Zedeck, S. (1973). Development of behaviorally anchored scales for the evaluation of faculty teaching. *Journal of Applied Psychology*, *58*(2), 261–265. https://doi.org/10.1037/h0035633

Imad-dine Bazine. (2025). Exploring the Development of the Technology Acceptance Model (TAM): A Chronological Overview. *International Journal of Research and Scientific Innovation*, *XII*(VI), 1643–1655.https://doi.org/10.51244/ijrsi.2025.120600138

Kromrei, H. (2015). Enhancing the Annual Performance Appraisal Process: Reducing Biases and Engaging Employees Through Self-Assessment. *Performance Improvement Quarterly*, *28*(2), 53–64. https://doi.org/10.1002/piq.21192

Küper, A., & Krämer, N. (2024). Psychological Traits and Appropriate Reliance: Factors Shaping Trust in AI. *International Journal of Human-Computer Interaction*, 1–17. https://doi.org/10.1080/10447318.2024.2348216

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2019). The Challenges of Algorithm-Based HR Decision-Making for Personal Integrity. *Journal of Business Ethics*, *160*(2), 377–392. https://doi.org/10.1007/s10551-019-04204-w

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*(1), 90–103.

Marler, J. H., & Boudreau, J. W. (2017). An Evidence-Based Review of HR Analytics. *The International Journal of Human Resource Management*, *28*(1), 3–26. https://doi.org/10.1080/09585192.2016.1244699

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

Meijerink, J., Boons, M., Keegan, A., & Marler, J. (2021). Algorithmic human resource management: Synthesizing developments and cross-disciplinary insights on digital HRM. *The International Journal of Human Resource Management*, *32*(12), 1–18. https://doi.org/10.1080/09585192.2021.1925326

Mesbah, N., Christoph Tauchert, & Buxmann, P. (2021). Whose Advice Counts More – Man or Machine? An Experimental Investigation of AI-based Advice Utilization. *Proceedings of the ... Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2021.496

Montealegre-López, N. (2025). Exploring the role of trust in AI-driven decision-making: a systematic literature review. *Management Review Quarterly*. https://doi.org/10.1007/s11301-025-00526-4

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal social, organizational, and goal-based perspectives*. Sage Publications.

Perello Marin, M. R., & Tuffaha, M. (2021). Artificial intelligence definition, applications and adoption in Human Resource Management: a systematic literature review. *International Journal of Business Innovation and Research*, *1*(1), 1. https://doi.org/10.1504/ijbir.2021.10040005

Rego, A., Cunha, M. P., Gomes, J. F. S., Cunha, R. C., Cardoso, C. C., & Marques, C. A. (2008). O que é a gestão de pessoas/recursos humanos?. *Manual de Gestão de Pessoas e do Capital Humano* (3rd ed., pp. 55-57). Sílabo.

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. https://doi.org/10.1177/0008125619862257

Stellar, J. E., & Willer, R. (2018). Unethical and inept? The influence of moral information on perceptions of competence. *Journal of Personality and Social Psychology*, *114*(2), 195–210. https://doi.org/10.1037/pspa0000097

Thakur, A., Hinge, P., & Adhegaonkar, V. (2023). Use of Artificial Intelligence (AI) in Recruitment and Selection. *Advances in Computer Science Research*, 632–640. https://doi.org/10.2991/978-94-6463-136-4_54

Tuffaha, M. (2023). The Impact of Artificial Intelligence Bias on Human Resource Management Functions: Systematic Literature Review and Future Research Directions. *European Journal of Business and Innovation Research*, *11*(4), 35–58. https://doi.org/10.37745/ejbir.2013/vol11n43558

Venkatesh, V. and Davis, F.D. “A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies,” *Management Science* (46:2), 2000, 186-204. https://doi.org/10.1287/mnsc.46.2.186.11926

Wen, Y., Wang, J., & Chen, X. (2025). Trust and AI weight: human-AI collaboration in organizational management decision-making. *Frontiers in Organizational Psychology*, *3*. https://doi.org/10.3389/forgp.2025.1419403

Yaniv, I. (2004). Receiving other people’s advice: Influence and benefit. *Organizational behavior and human decision processes*, *93*(1), 1-13.

Appendix

Appendix 1: Frequency tables – participant removal and validation

Estatísticas

		Terminado	att_check_valid	attcheck_valid = "correct" AND Finished = 1 (FILTER)
N	Válido	217	217	217
	Omisso	0	0	0

attcheck_valid = "correct" AND Finished = 1 (FILTER)

	N	%
Selected	217	100,0%

Appendix 2: Frequency tables – categorical variables

List of Countries

		Frequência	Porcentagem	Porcentagem válida	Porcentagem acumulativa
Válido	Austria	1	,5	,5	,5
	France	1	,5	,5	,9
	Germany	3	1,4	1,4	2,3
	Portugal	210	96,8	96,8	99,1
	United Kingdom of Great Britain and Northern Ireland	2	,9	,9	100,0

Frequency Table 1: List of Countries

What is your first language? - Selected Choice

		Frequência	Porcentagem	Porcentagem válida	Porcentagem acumulativa
Válido	Portuguese	210	96,8	96,8	96,8
	English	2	,9	,9	97,7
	Other	5	2,3	2,3	100,0
	Total	217	100,0	100,0	

Frequency Table 2: First Language

What is your gender?

		Frequência	Porcentagem	Porcentagem válida	Porcentagem acumulativa
Válido	Male	95	43,8	43,8	43,8
	Female	122	56,2	56,2	100,0
	Total	217	100,0	100,0	

Frequency Table 3: Gender

How old are you?

		Frequência	Porcentagem	Porcentagem válida	Porcentagem acumulativa
Válido	Under 18	1	,5	,5	,5
	18-24 years old	81	37,3	37,3	37,8
	25-34 years old	42	19,4	19,4	57,1
	35-44 years old	28	12,9	12,9	70,0
	45-54 years old	27	12,4	12,4	82,5
	55-64 years old	30	13,8	13,8	96,3
	65+ years old	8	3,7	3,7	100,0
	Total	217	100,0	100,0	

Frequency Table 4: Age

What's the highest level of education you achieved? - Selected Choice

		Frequência	Porcentagem	Porcentagem válida	Porcentagem acumulativa
Válido	Less than High School	1	,5	,5	,5
	High School	27	12,4	12,4	12,9
	Bachelor's degree	86	39,6	39,6	52,5
	Master's degree	84	38,7	38,7	91,2
	Doctoral degree	17	7,8	7,8	99,1
	Other	2	,9	,9	100,0
	Total	217	100,0	100,0	

Frequency Table 5: Education level

Appendix 3: Weight of Advice

Estatísticas

		WOA_1	WOA_2	WOA_3	WOA_4	WOA_mean
N	Válido	206	209	210	205	217
	Omisso	11	8	7	12	0

Appendix 4: Cronbach's alpha

Estatísticas de confiabilidade

Alfa de Cronbach	N de itens
,831	4

Cronbach's alpha 1: WOA

Estatísticas de confiabilidade

Alfa de Cronbach	Alfa de Cronbach com base em itens padronizados	N de itens
,966	,966	3

Cronbach's alpha 2: Trust in AI

Estatísticas de confiabilidade

Alfa de Cronbach	Alfa de Cronbach com base em itens padronizados	N de itens
,865	,872	3

Cronbach's alpha 3: AI familiarity scales

Estatísticas de confiabilidade

Alfa de Cronbach	Alfa de Cronbach com base em itens padronizados	N de itens
,959	,959	3

Cronbach's alpha 4: PU

Appendix 5: Bivariate Correlations

Estatísticas

WOA_mean

N	Válido	217
	Omisso	0
Média		,1393
Erro Desvio		,83556

Descriptive Statistics 1: WOA's Mean and SD

Correlações

		WOA_mean	Gender dummy	country dummy	Age
WOA_mean	Correlação de Pearson	1	-,023	,018	-,115
	Sig. (2 extremidades)		,738	,797	,091
	N	217	217	217	217
Gender dummy	Correlação de Pearson	-,023	1	-,003	,070
	Sig. (2 extremidades)	,738		,960	,306
	N	217	217	217	217
country dummy	Correlação de Pearson	,018	-,003	1	,113
	Sig. (2 extremidades)	,797	,960		,097
	N	217	217	217	217
Age	Correlação de Pearson	-,115	,070	,113	1
	Sig. (2 extremidades)	,091	,306	,097	
	N	217	217	217	217

Bivariate Correlations 1

Appendix 6: One Sample T-test for H1

Teste de uma amostra

	Valor de Teste = 0						
			Significância		Diferença média	95% Intervalo de Confiança da Diferença	
	t	df	Unilateral p	Bilateral p		Inferior	Superior
WOA_mean	2,456	216	,007	,015	,13930	,0275	,2511

Tamanhos de efeitos de amostra

		Padronizador[a]		Intervalo de Confiança 95% Inferior	Superior
WOA_mean	d de Cohen	,83556	,167	,033	,301
	Correção de Hedges	,83848	,166	,032	,299

a. O denominador usado na estimativa dos tamanhos dos efeitos.
O d de Cohen usa o desvio padrão de amostra.
A correção de Hedges usa o desvio padrão de amostra, além de um fator de correção.

Appendix 7: Multiple Linear Regression analysis for H3

Variáveis Inseridas/Removidas[a]

Modelo	Variáveis inseridas	Variáveis removidas	Método
1	PEOU_Mean, PU_Mean[b]	.	Inserir

a. Variável Dependente: TrustAI_Mean

b. Todas as variáveis solicitadas inseridas.

Resumo do modelo

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa
1	,932[a]	,869	,868	,58493

a. Preditores: (Constante), PEOU_Mean, PU_Mean

ANOVA[a]

Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	485,217	2	242,609	709,090	<,001[b]
	Resíduo	73,218	214	,342		
	Total	558,435	216			

a. Variável Dependente: TrustAI_Mean

b. Preditores: (Constante), PEOU_Mean, PU_Mean

Coeficientes[a]

Modelo		Coeficientes não padronizados B	Erro Erro	Coeficientes padronizados Beta	t	Sig.
1	(Constante)	,043	,139		,310	,757
	PU_Mean	,725	,043	,722	16,834	<,001
	PEOU_Mean	,248	,043	,245	5,711	<,001

a. Variável Dependente: TrustAI_Mean

Appendix 8: Linear regression analysis for H2

Variáveis Inseridas/Removidas[a]

Modelo	Variáveis inseridas	Variáveis removidas	Método
1	TrustAI_Mean[b]	.	Inserir

a. Variável Dependente: WOA_mean

b. Todas as variáveis solicitadas inseridas.

Resumo do modelo

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa
1	,147[a]	,022	,017	,82837

a. Preditores: (Constante), TrustAI_Mean

ANOVA[a]

Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	3,270	1	3,270	4,766	,030[b]
	Resíduo	147,533	215	,686		
	Total	150,803	216			

a. Variável Dependente: WOA_mean

b. Preditores: (Constante), TrustAI_Mean

Coeficientes[a]

Modelo		Coeficientes não padronizados		Coeficientes padronizados	t	Sig.
		B	Erro Erro	Beta		
1	(Constante)	-,230	,178		-1,289	,199
	TrustAI_Mean	,077	,035	,147	2,183	,030

a. Variável Dependente: WOA_mean