THREE ESSAYS ON EFFECTS OF PREDICTIVE ANALYTICS ON PLANNING AND DECISION-MAKING

Dissertation

to obtain the academic degree Doctor rerum politicarum (Dr. rer. pol.)
submitted at the

Faculty of Business and Economics, TU Dortmund University

by

Christian Ertel

from Mühlacker

Dortmund, 2025

1. Reviewer: Prof. Dr. Andreas Hoffjan
2. Reviewer: Prof. Dr. Maik Lachmann
3. Member of the Commission: Prof. Dr. Manuel Wiesche

Table of Contents

Summary (English)

This cumulative dissertation examines how predictive analytics affects planning and decision-making in organizations. In particular, the dissertation focusses on factors fostering or hindering managers' adoption of predictive analytics tools and on actionable strategies for management accounting to steer the respective use. Three studies, combining theoretical reasoning, qualitative interviews and behavioral experiments, address these overarching questions aims:

- **Study 1** explores the individual adoption of a human-resource analytics tool through interviews with potential users. On the one hand, it characterizes salient beliefs that determine perceived behavioral control, attitudes and social norms regarding the examined tool, which in turn influence the later use of the tool. On the other hand, it qualitatively links these beliefs to machine-learning characteristics such as visualizations and explainable AI, the degree of automation and fairness.
- **Study 2** introduces a theoretical framework for "algorithm aversion," a tendency to discount algorithmic advice even when it is superior. Drawing on the Theory of Planned Behavior, it identifies two causes of algorithm aversion: an unfavorable attitude towards algorithmic judgment compared to human judgment and lower perceived behavioral control (PBC) in using algorithms than over human judgment. The study therefore recommends mitigation strategies that focus on the ratio of attitudes towards and PBC over algorithmic and human judgment.
- **Study 3** investigates how managers assess the credibility of algorithmically generated sales forecasts. A laboratory experiment compares two information cues: a capability-based metric signaling technical forecasting skill and a predictability-based metric indicating the stability of the underlying data. Communicating capability-based metrics always increases perceived credibility, even when unwarranted. This result highlights that proposed mitigation strategies for algorithm aversion by prior literature might have unintended consequences. Communicating predictability-based metrics appears superior, because respective metrics help calibrating credibility in line with actual forecast accuracy.

The general discussion synthesizes these insights and explains how the individual studies contribute the overarching research questions. Furthermore, it discusses limitations of the included studies and characterizes avenues for further research.

Summary (German)

Diese kumulative Dissertation untersucht, wie sich Predictive Analytics auf Planung und Entscheidungsfindung in Organisationen auswirken. Im Mittelpunkt der Arbeit steht, welche Faktoren die Nutzung solcher Werkzeuge durch Manager beeinflussen und welche Handlungsmöglichkeiten das Controlling hat, um deren Einsatz zu steuern. Drei Studien, die theoretische Modellierung, qualitative Interviews und Verhaltensexperimente kombinieren, adressieren diese übergeordneten Fragestellungen:

- **Studie 1** beleuchtet die individuelle Einführung eines Human-Ressource-Analytics-Werkzeugs anhand von Interviews mit potenziellen Nutzern. Zum einen werden wesentliche Überzeugungen charakterisiert, die die wahrgenommene Verhaltenskontrolle, die Einstellungen und die sozialen Normen in Bezug auf das untersuchte Tool bestimmen und damit die spätere Nutzung beeinflussen. Zum anderen werden diese Überzeugungen qualitativ mit Maschine Learning-Eigenschaften wie Visualisierungen und erklärbarer KI, dem Automatisierungsgrad und Fairness verknüpft.
- **Studie 2** präsentiert einen theoretischen Rahmen für die sogenannte „Algorithmus-Aversion“ - die Tendenz, algorithmische Empfehlungen selbst dann abzulehnen, wenn sie überlegen sind. Ausgehend von der Theory of Planned Behavior werden zwei Ursachen identifiziert: eine ungünstige Einstellung gegenüber algorithmischen Urteilen im Vergleich zu menschlichen Urteilen und eine geringere wahrgenommene Verhaltenskontrolle (PBC) bei der Verwendung von Algorithmen als bei menschlichen Urteilen. Die Studie empfiehlt daher Gegenmaßnahmen, die sich auf das Verhältnis von Einstellungen und PBC gegenüber algorithmischen und menschlichen Urteilen konzentrieren.
- **Studie 3** untersucht, wie Manager die Glaubwürdigkeit algorithmisch erzeugter Verkaufsprognosen einschätzen. In einem Laborexperiment wird die Wirkung zweier Kennzahlen verglichen: ein fähigkeitsbasiertes Maß, das technische Prognosekompetenz signalisiert, und ein zeitreihenbasiertes Maß, das die Stabilität der zugrunde liegenden Daten anzeigt. Die Kommunikation fähigkeitsbasierter Kennzahlen erhöht die wahrgenommene Glaubwürdigkeit stets, selbst wenn dies nicht gerechtfertigt ist. Diese Ergebnisse zeigen, dass vorgeschlagene Maßnahmen zur Verringerung der Algorithmus-Aversion unbeabsichtigte Nebenwirkungen haben können. Zeitreihen-

basierte Kennzahlen erscheinen überlegen, da sie helfen, die Glaubwürdigkeit entsprechend der tatsächlichen Prognosegenauigkeit zu kalibrieren.

Die allgemeine Diskussion fasst die einzelnen Erkenntnisse zusammen und erläutert, wie die Ergebnisse der einzelnen Studien zur Beantwortung der übergreifenden Forschungsfragen beitragen. Außerdem werden die Grenzen der enthaltenen Studien diskutiert und Ansatzpunkte für künftige Forschung aufgezeigt.

Figures

Tables

Abbreviations and symbols

AAA	American Accounting Association
ACMAR	Annual Conference for Management Accounting Research
AI	Artificial Intelligence
AIS	Accounting Information System
ARIMA	Autoregressive Integrated Moving Average
BI	Business Intelligence
DFG	Deutsche Forschungsgemeinschaft
EAA	European Accounting Association
ENEAR	European Network for Experimental Accounting Research
ERP	Enterprise Resource Planning
EU	European Union
GDPR	General Data Protection Regulation
GenAI	Generative Artificial Intelligence
HR	Human Resources
HRA	Human Resource Analytics
HRM	Human Resources Management
ISAIS	International Symposium of Accounting Information Systems
ISS	Internet Skills Scale
MAS	Management Accounting Section
ML	Machine Learning
PBC	Perceived Behavioral Control
PEQ	Post-Experimental Questionnaire
RPA	Robotic Process Automation
SARIMA	Seasonal Autoregressive Integrated Moving Average
TPB	Theory of Planned Behavior
TRA	Theory of Reasoned Action
TTD	Theory of Technology Dominance
UTAUT	Unified Theory of Acceptance and Use of Technology
XAI	Explainable AI / Explainable Artificial Intelligence

1. Introduction

1.1 Motivation and research aim

"Control cannot be studied apart from technology and context because one will never get to understand the underlying "infrastructure" - the meeting point of many technologies and many types of controls." (cf. Dechow & Mouritsen 2005, p. 691)

The domain of management accounting and control is inextricably linked to the use of various information systems (e.g., Granlund 2011, Dechow & Mouritsen 2005). In particular, machine learning (ML) and other technologies grouped under the terms "data analytics" and "predictive analytics" promise to revolutionize planning and performance measurement (e.g., Nielsen 2022, Nielsen 2018, Appelbaum et al. 2017). The following section provides an overview over the potentials of predictive analytics in management accounting. Moreover, it describes limitations of prior research, which motivate this dissertation. Furthermore, it verbalizes the overarching research aim of this dissertation.

1.1.1 Background

In general, management accountants come in contact with several information systems. First, enterprise resource planning (ERP) systems provide a platform for the aggregation of finance, human-resources (HR), production, logistics, purchasing, sales and marketing information (e.g., Grabski et al. 2011, Rom & Rohde 2007, Granlund & Malmi 2002). Second, business intelligence (BI) solutions leverage the internal reporting processes and allow to build real-time and user-oriented dashboards (e.g., Martins et al. 2025, Rikhardsson & Yigitbasioglu 2018, Peters et al. 2016). Third, spreadsheet software allows for the fast and flexible analysis and aggregation of information from different data sources (e.g., Church et al. 2022, Schmidt et al. 2020b, Goretzki et al. 2018). Fourth, automated bots centered around the term robotic process automation (RPA) can take over repetitive and rule-based tasks of the management accounting department (e.g., Korhonen et al. 2021, Plattfaut & Borghoff 2022), particularly when powered by generative artificial intelligence (GenAI) (e.g., Abbas 2025). Fifth, process mining techniques discover unknown relationships and processes in an organization which can often be optimized (e.g., Zerbino et al. 2021, Werner 2017). Sixth, blockchain-based approaches can replace management controls by enabling the rule-based coordination between departments and organizations (Kostić & Sedej 2022).

One of the most promising set of technologies for management accounting and control are data analytics and particularly predictive analytics. In general, data analytics is a collective term for a comprehensive set of mathematical and statistical techniques, which allow to analyze historical datasets to generate business relevant insights (Rikhardsson & Yigitbasioglu 2018, Holsapple et al. 2014). Predictive analytics are a subset of these techniques, which rely on statistical modeling to predict future developments from past data (Spraakman et al. 2021, Appelbaum et al. 2017).[1] Examples for predictive analytics can range from simple and transparent regression approaches such as Autoregressive Integrated Moving Average (ARIMA) models (e.g., Kinney 1978) to machine learning based approaches such as random-forest algorithms or hierarchical clustering algorithms (e.g., Nielsen 2022). While the application of all types of data analytics requires sound mathematical and coding knowledge (e.g., Petropoulos et al. 2022), there is an increasing number of software companies offering respective services in the form comprehensive IT-tools. For instance, IBM offers its solution IBM Cognos Analytics (IBM 2025), Microsoft offers data analytics functionalities in its Microsoft Fabric Platform (Microsoft 2025) and the German-based software developer SAP offers a solution called SAP Analytics Cloud (SAP 2025).

From the perspective of management accounting and control, predictive analytics tools can support organizational planning and performance measurement activities. On the one hand, organizational planning and budgeting processes have long deemed as time consuming and inaccurate (e.g., Hebeler 2005, Hansen et al. 2003). Prior research identifies several reasons for this ineffectiveness, ranging from a to short-term oriented focus to the promotion of internal gambling behavior (Neely et al. 2003). For the given context, predictive analytics tools are able to provide reasonable and neutral estimations for the future organizational performance and the resulting budget needs of the different departments (Appelbaum et al. 2017). On the other hand, there is an ongoing discussion around actionable targets serving as the basis for performance measurement (e.g., Pfister & Lukka 2019, Sprinkle et al. 2008). Conceptually, targets represent the expectation of management for future individual or group performance and serve as an anchor to evaluate actual performance (Otley 1999). Predictive analytics tools provide reliable forecasts of future

[1] Note that this dissertation treats the terms “data analytics” and “business analytics” as synonyms. While the term “business analytics” is sometimes used to emphasize the use of data analytics in a business context (Holsapple et al. 2014), this context already follows from the topic of the dissertation and therefore appears unnecessary.

performance. These can serve as either an expected target or a minimum target to be exceeded (Appelbaum et al. 2017). On a more fundamental level, predictive analytics tools coincide with a period where management accounting is increasingly expected to leverage the existing data to provide forward-looking and valuable decision-support to management (e.g., Möller et al. 2020, Bhimani & Willcocks 2014).

1.1.2 Limitations of prior research

Prior accounting-based research on predictive analytics shows several limitations, which motivate the studies included in this dissertation. First, there are extremely few publications on data analytics and other accounting information systems (AIS) in major accounting journals (e.g., Jans et al. 2023, Coyne et al. 2010, Efendi et al. 2006).[2] For instance, Jans et al. (2023) analyze the major accounting research between 2000 and 2018 and identify 174 AIS-related publications. With a total of over published 8.000 papers in the examined journals between 2000 and 2018, AIS-related publications only account for 2 percent of the total accounting publications.[3]. This is a stark contrast to the current and historical importance of respective IT-tools in accounting practice (e.g., Zorn et al. 2025) and this gap between research and practice has been highlighted for decades (e.g., Granlund 2011, Sutton 2006). Addressing this research-practice-gap is urgent, because recent developments in artificial intelligence (AI) have the potential to revolutionize the profession and to render established knowledge and concepts largely irrelevant (Boritz & Stratopoulos 2023). Furthermore, understanding the current challenges of accounting practice enables to provide effective accounting education to current and future students (Tucker & Lowe 2014, Tucker & Parker 2014).

Second, drawing on the sparse existing research, there is not much knowledge on how managers actually perceive the information generated by predictive analytics and whether they are able to use it effectively (Casas-Arce et al. 2022). However, decades of research on human-computer-interaction suggests profound biases and subsequent challenges when managers interact with computers. Examples range from well-known cognitive biases such as information overload and overconfidence (e.g., Brüggen et al. 2021, Hogarth

[2] In this context, the term "major accounting journals" describes d the topic-specific journals usually seen as "A+" or "A" journal and therefore for example "The Accounting Review", "Journal of Accounting Research", "Journal of Accounting and Economics", "Review of Accounting Studies", "Accounting, Organizations and Society" and "Contemporary Accounting Research" among others.
[3] Although Jans et al. (2023) do not state this total number of publications in their paper, it can be quickly identified with the help of an online search on the online database SCOPUS.

& Makridakis 1981) to an irrational rejection of computer-generated advice by decision-makers (e.g., Kawaguchi 2021, Dietvorst et al. 2018, Dietvorst et al. 2015, Önkal et al. 2009, Dawes 1979, Meehl 1954) and to loss of core competencies due to an irrational trust in expert systems (Sutton et al. 2023, Arnold et al. 2004, Arnold & Sutton 1998) among others. Recent research in management accounting finds similar effects, for instance a reluctance of decision-makers to rely on algorithmically generated advice which predicts the breaking of a sales trend (Fehrenbacher et al. 2023, Chen et al. 2022a). Better understanding these managerial biases is of high interest, because fulfilling the business partner role of management accounting requires the guidance and control of managerial decision-making (van Slooten et al. 2024, Järvenpää 2007), for instance through a selective and thought-out communication of metrics (Casas-Arce et al. 2017).

Third, there is little research on the individual tools and system design alternatives of predictive analytics in management accounting (e.g., Sutton et al. 2021, Geerts 2011). Most of prior studies treat data analytics as one contingency factor and examine its effect on management accountings roles and tasks. Findings include the change to a more business-partner orientated and explainer role (e.g., van Slooten et al. 2024, Kokina et al. 2021, Al-Htaybat & Alberti-Alhtaybat 2017), the necessity to acquire competencies in data science and statistics (e.g., Oesterreich & Teuteberg 2019) and the observation of a complicated change process (e.g., Korhonen et al. 2021) among others. In particular German-based research has long avoided to examine specific IT-systems and has instead often relied on more general concepts as “digitalization” and “digital transformation” (e.g., Hiebl & Pielsticker 2020, Möller et al. 2020, Schäffer & Weber 2016). This research paradigm is problematic, because prior research in information systems observes significant differences in the use of predictive analytics tools depending on their presentation and conveyed expertise (e.g., Berger et al. 2021, Hou & Jung 2021). Therefore, aggregating data analytics into one contingency factor likely simplifies the underlying theoretical complexity and suggests a false level of generalizability of the empirical findings (Mauldin & Ruchala 1999).

1.1.3 Overarching research aim

This cumulative dissertation addresses the described limitations of prior accounting-related research on predictive analytics. The included studies aim to investigate managers individual perception and decision-making in planning tasks when confronted with a forecast generated by predictive analytics. These studies hereby provide a refined and extended understanding of the salient attitudes and opinions that convince or deter managers from incorporating respective forecasts into their work. Moreover, these studies characterize the necessary decision-support from management accounting that is required to calibrate and control managers' effective use of these forecasts. In general, the included studies in this dissertation therefore answer two overarching research questions:

- What are relevant factors that encourage or deter managers from incorporating predictive analytics tools into planning activities?
- How can management accounting influence managers' use of predictive analytics in these tasks?

The general discussion chapter further characterizes the included studies contributions to these research questions and characterizes avenues for further research.

1.2 Focus of the included studies

Following the practice of consecutive research and due to the absence of suitable accounting-related research, the included studies in this dissertation focus on managers' use of predictive analytics in human resource (HR) as well as in sales and operations planning and the underlying psychological mechanisms, in particular a well-known and counterintuitive behavior called "algorithm aversion". The following section characterizes the chosen areas of research and justifies their contribution for the overall research aim of this cumulative dissertation.

1.2.1 Turnover prediction

The first project included in this dissertation addresses the use of HR Analytics for employee capacity planning tasks. In general, quitting of employees can lead to significant cost for organizations, for instance because of the required application process for new employees and the loss of knowledge (e.g., Perryer et al. 2010, Carmeli & Weisberg 2006). Therefore, managers require predictions for potential employees turnover, so that they can convince them to stay or to begin searching for a suitable replacement at an early stage (Morrell et al. 2001).[4] The reasons for employee turnover have been examined extensively by prior research in HR and the literature describes a diverse bundle of

[4] Note that this dissertation draws on HR for the definition of the term "turnover". In particular, the used term in this dissertation does not represent the gross revenue of a corporation.

antecedents with non-linear effects (e.g., Rubenstein et al. 2018, Holtom et al. 2008, Carmeli & Weisberg 2006, Gray & Phillips 1994). Consequently, there is a growing interest in using advanced predictive analytics to analyze the relevant data and to generate respective turnover predictions (e.g., Chowdhury et al. 2022, Choudhury et al. 2021, Rombaut & Guerry 2018). Advanced approaches, for example based on machine learning, reach a prediction accuracy for turnover of around 80 percent (Park et al. 2024).

There are three main reasons that motivate the first study included in this dissertation. From a theoretical point of view, the given HR context provides an interesting opportunity to examine the fundamental research questions of this dissertation from a unique perspective. Note that coordinating personnel management and organizational planning processes is considered a key responsibility of management accounting in the German-speaking accounting literature (Küpper et al. 2024). This task particularly requires the coordination of managements personnel plans with the actual personnel needs of an organization (Küpper et al. 2024), which evidently depend on the expected turnover of employees. Due to the recent technological advancements in machine learning regarding explainable artificial intelligence (XAI), it appears likely that managers will increasingly rely on the turnover predictions from these tools (e.g., Heidemann et al. 2024). Effective coordination of managements personnel-related decisions therefore requires a deeper understanding of their reasons to adopt respective HR analytics tools.

From a pragmatic point of view, there was a unique opportunity to examine the implementation of a real HR analytics tool in an organization. One of the co-authors acquired the possibility for an in-depth case study in a German federal agency, which was implementing a respective analytics tool for employee turnover prediction. This context included the possibility of twelve interviews with managers across the agency, in which we were able to discuss their impressions of the tool and their underlying reasoning. From a publication-oriented point of view, the given context hereby addressed a noticeable gap in prior research on HR analytics. While there is prior research on the adoption of respective HR analytics tools on an organization-wide level (e.g., Coolen et al. 2023, Neumann et al. 2022b, Di Vaio et al. 2022), there was only one study addressing the individual adoption of HR analytics tools by managers at the time the research project was started (Vargas et al. 2018). Moreover, as this prior study draws on a survey research method, it failed to provide an in-depth understanding of managers reasoning during the adoption process (Vargas et al. 2018).

1.2.2 Algorithm aversion

The second project included in this dissertation addresses the observed "algorithm aversion" of decision-makers in many forecasting tasks, where these discount or disregard objectively valuable algorithmic advice (Jussupow et al. 2024, Mahmud et al. 2022, Dietvorst et al. 2018, Dietvorst et al. 2015).[5] [6] While this interdisciplinary stream of literature originates from psychology and has examined diverse decisions and decision-makers (e.g., Dietvorst et al. 2018, Dietvorst et al. 2015), it has expanded to customers' use of algorithmic services (e.g., Reich et al. 2023, Castelo et al. 2019), physicians' use of medical decision-aids (e.g., Pezzo et al. 2021, Longoni et al. 2019), judges' use of criminal risk assessment tools (e.g., Fine et al. 2023, Bonezzi & Ostinelli 2021), investors' use of algorithmic investment assistance (Downen et al. 2024, Himmelstein & Budescu 2023) and managers' use of forecasting algorithms (e.g., Greiner et al. 2025, Kawaguchi 2021) among others. Due to a growing research interest in the last decade, the stream of literature on algorithm aversion includes well over 100 peer-reviewed papers and over 50 identified antecedents that influence the observed resistance behaviors (e.g., Jussupow et al. 2024, Mahmud et al. 2022).

The second study is motivated by two reasons. From a theoretical point of view, the applied research paradigm of prior research on algorithm aversion matches the described decision problems in corporate accounting practice and the stream of literature therefore provides interesting insights for the given context. Recall that the aim of this dissertation is to examine how managers incorporate the forecasts generated by predictive analytics tools into their decision-making. As operating respective tools requires time and expert knowledge (e.g., Appelbaum et al. 2017), it is likely that managers are only provided with the final forecast of a predictive analytics tool. This context is very similar to the applied judge-advisor-system research paradigm of the algorithm aversion literature (Jussupow et al. 2024, Bonaccio & Dalal 2006). Here, algorithms are treated as an advisor and prior research only examines how decision-makers react to this algorithmically generated

5 Note that there is no unified definition of "algorithm aversion". Definition approaches range from "the reluctance of human decision-makers to use superior but imperfect algorithms" (cf. Burton et al. 2020, p. 220) to "a behavior of neglecting algorithmic decisions in favor of one's own decisions or other's decisions, either consciously or unconsciously" (cf. Mahmud et al. 2022, p. 1). However, all definitions include a decision-maker who irrationally discounts or rejects advice from a well-performing algorithm (Jussupow et al. 2024)

6 Further note that this stream of research follows a very broad understanding of an „algorithm ". Definitions include "a statistical model put together by careful analysis" (cf. Dietvorst et al. 2015, p. 4) to "a proprietary AI system called the Amadeus System" (cf. Commerford et al. 2022, p. 184) among others.

advice. In line with a manager receiving a final forecast, these algorithmic advisors are mostly described verbally to control for potential effects of the interface design of a respective algorithm (e.g., Dietvorst et al. 2015).

From a publication-oriented point of view, focusing to the ongoing research on algorithm aversion answers several general calls for further research. First, recent accounting-related research on the use of predictive analytics recommends to further examine algorithm aversion in forecasting tasks and to focus on effective mitigation strategies (e.g., Downen et al. 2024, Chen et al. 2022a, Commerford et al. 2022). Second, as the empirical findings in regards to algorithm aversion are mostly in line with older research on the use of AIS by professional decision-makers, more research on the this behavior also contributes to these streams of research (Sutton et al. 2023). Third, prior research on algorithm aversion is limited by the specific and often ambiguous characteristics of the examined forecasting tasks and seeks for more diverse decision-problems to validate the relevance of the observed effect (e.g., De-Arteaga et al. 2020, Dietvorst et al. 2018).[7]

1.2.3 Demand forecasting

The third project included in this dissertation addresses the use of demand forecasts for sales and operations planning. Respective forecasts are seen as an important control mechanism of an organization, because they allow managers to coordinate an organization's product assortment with actual market demand (e.g., Brüggen et al. 2021, Jordan & Messner 2020). It is hereby necessary that these forecasts are somewhat accurate, because errors can lead to costly overproduction as well as supply shortages that negatively affect service quality (Oliva & Watson 2009). Furthermore, unrealistic targets can demotivate the salespersons of an organization (Jordan & Messner 2020). As some markets are volatile, demand forecasting is a challenging and error-prone task and demand forecasters rely on advanced forecasting algorithms for decades (e.g., Makridakis et al. 2020b, Hyndman & Koehler 2006, Makridakis et al. 1987). Novel forecasting algorithms, for example based on machine learning techniques, have leveraged their effectiveness compared to traditional approaches for time series analysis (e.g., Feizabadi 2022, Petropoulos et al. 2022, Sohrabpour et al. 2021, Kilimci et al. 2019).

[7] For instance, Dietvorst & Bharti (2020) rely on fictional tasks, for example the forecasting of crops harvests and mining outputs on an alien planet.

Again, the third research project is motivated by three reasons. From a theoretical point of view, sales and operations planning is an area where managers require assistance when evaluating algorithmically generated forecasts. In general, respective predictive analytics tools often exceed the capabilities of human judgment in demand forecasting (e.g., Brüggen et al. 2021, Cui et al. 2018, Andreassen & Kraus 1990) and their forecasts should be incorporated into organizational planning. However, due to unpredictable events such as COVID19 (e.g., Ioannidis et al. 2022) and methodical limitations (e.g., Khashei & Bijari 2011), these tools will always remain error-prone and erroneous forecasts should be disregarded. While managers are able to account for and adjust some of the described errors (e.g., Fildes et al. 2009), they nevertheless have difficulties in identifying respective erroneous forecasts (e.g., Goodwin & Fildes 1999). Thus, the context of sales and operations planning allows to examine novel approaches for decision assistance to calibrate managers' use of predictive analytics tools and addresses particularly the second research question of this dissertation.

From a publication-oriented point of view, the error-proneness of predictive analytics tools is a blind-spot in the aforementioned literature on algorithm aversion. Prior literature often argues that using a superior but error-prone algorithm on average improves forecasting accuracy and planning quality (e.g., Greiner et al. 2025, Dietvorst et al. 2015). This argumentation disregards that identifying erroneous forecasts would improve the respective accuracy and quality even further (e.g., De-Arteaga et al. 2020). Moreover, this blind-spot questions the effectiveness of many derived mitigation strategies for algorithm aversion (e.g., Reich et al. 2023, Berger et al. 2021), because they likely make decision-makers less attentive for erroneous forecasts. From a pragmatic point of view, there was the possibility to confront people with realistic forecasting tasks and to examine their reactions to a real predictive analytics tool designed for these tasks. One of the co-authors of the third project got access to a high volume of real time series of sales from a German manufacturer and is developing a predictive analytics tools based on these time series for her dissertation. To refine the tool, she was very interested in examining how potential users perceive the tool and to identify influencing factors.

1.3 Research framework

To examine the overarching topic of this dissertation from different perspectives, the included studies follow conceptual, case-study and experimental research methods. Although, there was no formal research plan across all studies beforehand, the studies are nevertheless related to each other. In particular, these studies can be seen as parts of the design cycle for management accounting approaches that calibrates managers' use of predictive analytics tools in planning tasks (Geerts et al. 2013, Hevner et al. 2004). The following section characterizes the design cycle for a respective artifact and places the three studies inside the framework.

1.3.1 The relevance and rigor cycle

The practical relevance of research in the field of accounting, including management accounting, has been discussed and questioned for decades (e.g., Marton 2025, Moser 2012, Choudhury 1986). Common points of critique are a narrow focus on publishable but mostly academic topics such as earnings management and a focus on established and mostly deductive research methods such as archival research (McCarthy 2012, Moser 2012). Addressing these points of critique, Geerts et al. (2013) draw on the design science research paradigm (e.g., Hevner et al. 2004) and propose a research framework that combines academic rigor and practical relevance. The fundamental idea is to understand accounting as a set of tools and methods that were created "to solve a problem in a specific environment", which are called artifacts (cf. Geerts et al. 2013, p. 815). Artifacts include costing methods such as activity based costing, control frameworks such as the COSO framework and AIS-tools such as predictive analytics among others (Geerts et al. 2013). In general, accounting research can be seen as a dynamic cycle of developing new artifacts from theory and refining theory based on existing artifacts that uses the diversity of all available research methods (Geerts et al. 2013).

In particular, the framework of Geerts et al. (2013) consists of a relevance and rigor cycle for research.[8] Research following the relevance cycle identifies urgent business challenges from corporate practice that need to be solved. Respective research then defines the aims of a respective solution and develops an artifact addressing the identified challenge. Lastly, the developed artifact is evaluated regarding its fulfillment of the set aims (e.g., Geerts 2011). Research following the rigor cycle identifies changes in the environment, for example the emergence of a novel artifact, and theorizes why and how this artifact should work or actually works. Respective research then proposes testable research propositions and hypotheses for empirical examination. Lastly, the developed

[8] Note that Geerts et al. (2013) build their framework on the similar framework of Hevner et al. (2004). While there are great similarities, this dissertation mostly draws from the newer and more specialized framework that is Geerts et al. (2013).

hypotheses are justified with a respective empirical examination (e.g., Sutton et al. 2021). The relevance and rigor cycles are notably interconnected. Note that the development of artifacts requires theoretical and methodological guidance to effectively address technological, organizational and behavioral challenges.[9] Furthermore, the different empirical research methods from the rigor cycle assist in the evaluation of developed artifacts. Moreover, the application of existing theories via artifacts helps to evaluate their usefulness and informs potential needs for refinement (Geerts et al. 2013). Figure 1-1 illustrates the framework of Geerts et al. (2013).

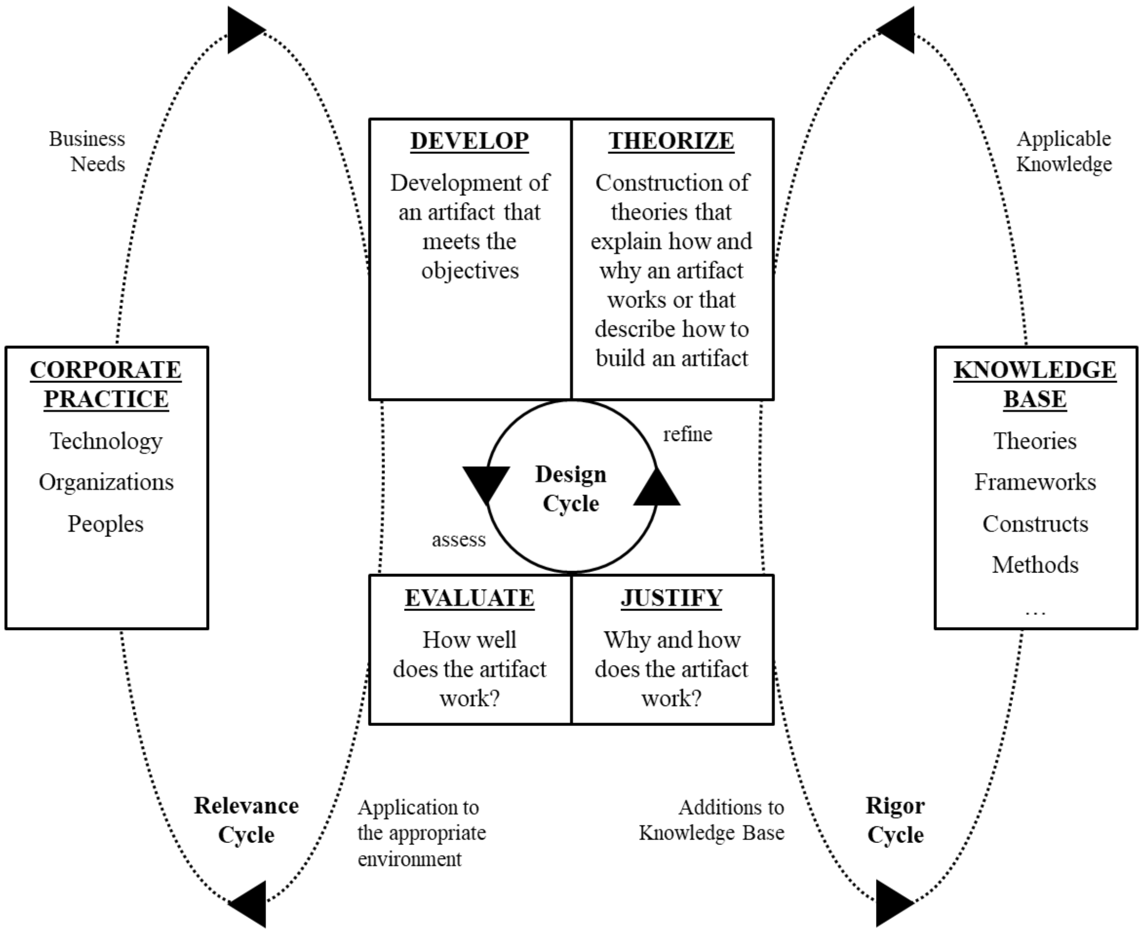

Figure 1-1: The relevance and rigor cycles for accounting research (adapted from Geerts et al. 2013).

[9] While Geerts et al. (2013) cite Hunton (2002) for characterizing the environment of accounting, Prof. James E. Hunton has since been convicted of systematic fraud and at least 37 of his papers have been retracted (Retraction Watch 2025). This dissertation therefore uses the characterization of Hevner et al. (2004) for corporate practice.

1.3.2 Methodological characterization of the dissertation

Recall that one aim of this dissertation is to deduce actionable strategies for management accounting to influence managers' use of predictive analytics tools for planning activities. The studies included in this dissertation therefore assess three different artifacts for respective decision support. Study 1 examines a strategy of providing XAI methods, in particular ALE-plots and SHAP-values (Heidemann et al. 2024) to managers in HR planning tasks. Respective strategies are relevant for predictive analytics tools based on ML, because they improve the interpretability of these complex and ambiguous algorithms (e.g., Dwivedi et al. 2023, Arrieta et al. 2020). For instance, they allow to detect preexisting biases in the dataset that could lead to harmful decisions for employees (Arrieta et al. 2020). Note that the first study is a follow-up to Heidemann et al. (2024), who already examine the technical effectiveness and efficiency of XAI methods in HR planning tasks ("evaluate"). There are also multiple publications on the theoretical use cases of XAI across various domains ("theorize", e.g., Haque et al. 2023, Zhang et al. 2022). Therefore, the first study therefore aims to justify and refine these conceptual and theoretical assumptions with the help of focused ethnographic interviews, which is a qualitative research approach ("justify", Merton & Kendall 1946).

Study 2 examines a strategy of solely communicating an algorithmically generated forecast to management. Respective strategies are relevant in context where managers access decision support without an involvement of the management accounting department, for example because of self-service tools (Alpar & Schulz 2016) or in small enterprises with less developed management accounting practices (Lavia López & Hiebl 2015). The stream of literature on algorithm aversion already shows that respective decision-support does not work well and leads to an irrational rejection of credible forecasts by decision-makers ("evaluate", see section 1.2.2). Moreover, prior research has already identified over 50 antecedents that explain why the strategy for decision-support does not work well ("justify", see section 1.2.2). However, there is a notable lack of theory-building explaining the effects of these antecedents ("theorize", e.g., Mahmud et al. 2022). The second study tackles this gap in prior research with the help of a developed theoretical framework and literature review.

Study 3 examines the use of capability-based and predictability-based forecast metrics to calibrate managers' use of algorithmically generated forecasts in sales and operations planning. While capability-based metrics indicate the actual capabilities of a predictive analytics tool (e.g., Reich et al. 2023, Berger et al. 2021), predictability-based metrics indicate the difficulty of a forecasting task (e.g., Flood & Grimm 2021, Salvino et al. 1995). Respective metrics for the steering of managers are interesting because their application is simple and independent from the respective algorithm underlying a predictive analytics tool. Note that capability-based approaches are a key idea to mitigate algorithm aversion (Reich et al. 2023, Berger et al. 2021, Castelo et al. 2019, Dietvorst et al. 2015). Based on prior research in information systems (e.g., Madhavan & Wiegmann 2007, Lee & See 2004), the third study hypothesizes that these strategies can lead to an overuse of predictive analytics tools in risky forecasting tasks ("theorize"). Moreover, it provides evidence for the hypotheses with the help of a laboratory experiment ("evaluate" / "justify"). Drawing on the developed rationale, the third study proposes ("develop") and validates ("evaluate" / "justify") predictability-based metrics as an effective alternative.

1.4 Outline

The following three chapters of this dissertation are the included studies. While the original drafts of studies follow different style guides and types of English, this dissertation presents them in a harmonized style and always uses American English.[10] The reference lists of the three papers have been merged with the reference list of the other sections and a full reference list is provided at the end of the text. Moreover, there is also only one list of abbreviations, one list of tables and one list of figures. All of the three chapters start with a small subsection in which the project timeline, conferences talks and target journals / actual journals are briefly described. After these three chapters follows a general discussion of the research findings and contributions in relation to the developed general research questions (see section 1.1.3). In addition, the discussion chapter also includes a general discussion of limitations of the included studies and characterizes avenues for further research. The last section is a brief conclusion of the text.

[10] Further note that some papers have extensive captions that do not fit into the overall style of this dissertation. These captions have been moved into footnotes instead.

2. Study 1: Exploring the Individual Adoption of Human Resource Analytics - Behavioral Beliefs and the Role of Machine Learning Characteristics

2.1 Description of the research project

This research project is a collaboration with Svenja M. Hülter, who is currently a PhD candidate at the Chair of Management Accounting and Control of the Department of Business and Economics at TU Dortmund University and Ansgar Heidemann, who is currently an external PhD candidate at the same chair (at July 2025). The paper has been published in the Journal "Technological Forecasting and Social Change" (VHB-Jourqual 4: B, Impact Factor (2024): 13.3) and has so far been cited eight times in other peer-reviewed papers (July 2025). The paper is the follow-up to Heidemann et al. (2024), where the other co-authors examined the same ML-based analytics tools from a technical perspective. The author of this dissertation has conceptualized the present follow-up study, and has deduced the applied research method of "focused-ethnographic interviews" (Merton & Kendall 1946), the applied coding approach (Gioia et al. 2013) and the existing framework for the adoption of HR analytics (Vargas et al. 2018) from literature. He was also an active co-interviewer in every interview and worked on the coding of the collected data in collaboration with Svenja Hülter. Lastly, he wrote the theoretical and conceptual parts of the final paper and critically revised the other sections.[11] Before publication, the project was presented at the 2nd European Institute for Advanced Studies in Management (EIASM) Workshop on People Analytics & Algorithmic Management.

2.2 Abstract

The technological capabilities of Human Resource Analytics (HRA), enhanced by recent innovations in Machine Learning (ML), offer exciting opportunities. However, organizations often fail to realize these potentials because of a limited understanding of why individuals choose to adopt or disregard respective tools. Prior research on innovation adoption offers preliminary insights but fails to aggregate the determinants of individual adoption into actionable suggestions for decisions in the ML adoption process. Our study applies focused interviews to examine non-ML experts' reasoning for using a specific tool

[11] Note that the authors signed a so-called "CRedit authorship contribution statement", which clarifies and officially documents the individual contributions of the authors to the paper (cf. Hülter et al. 2024, p. 11).

tailored to a public sector organization, which corresponds to the usual end-user perspective of ML-based HRA adoption. By drawing from the HRA adoption framework, provided by Vargas et al. (2018), we contribute to the literature by identifying relevant beliefs and experiences influencing one's intention to adopt ML-based HRA and by qualitatively linking these beliefs to ML characteristics such as transparency, automation and fairness. For practitioners, we provide actionable guidance emphasizing the need to ensure fairness proactively, as interviewees do not consider this aspect when deciding to adopt ML-based HRA.

2.3 Introduction

The diffusion of analytics into Human Resources Management (HRM) processes, including talent management, performance evaluation and workforce planning, presents a promising opportunity. Human Resources Analytics (HRA), as it is referred to in this context, is classified as diffusing innovation and describes "a practice enabled by information technology that uses descriptive, visual, and statistical analyses of data related to Human Resources (HR) processes, human capital, organizational performance, and external economic benchmarks to establish business impact and enable data-driven decision-making" (Marler & Boudreau 2017, p. 15). Modern-day technological advances help HR professionals asserting their value when defending against possible displacement by finance or data science departments (Angrave et al. 2016), and they provide support for a wide range of different HR functions (Prikshat et al. 2023a). In recruitment, for example, HRA can be used to streamline processes and achieve greater speed and efficiency (Hunkenschroer & Luetge 2022), whilst in HR development, it helps to identify the link between employee engagement and performance metrics and thus positively influences them (Davenport et al. 2010). Although traditional HRA encompasses several statistical approaches and methodologies, Machine Learning (ML), such as deep-learning algorithms and Artificial Intelligence, is expected to drive the greatest change in HRM practice. For example, online work platforms such as Uber, Upwork and Deliveroo automate extensive core business processes, ranging from HRM decision-making to execution in the form of selection, compensation and task assignment, all of which is done through ML (Meijerink et al. 2021, p. 2551). In addition, ML-based HRA tools for predicting voluntary employee turnover allow companies to derive retention strategies that not only reduce costly replacements in the short term, but also retain expertise within the organization, thereby securing a competitive advantage (Chowdhury et al. 2022).

Prior research identifies individual resistance to ML-based HRA which could hinder its success in corporate practice. In contrast to other HRA technologies, sophisticated ML algorithms, for instance, have the disadvantage of being too complex to interpret easily, which subsequently leads to opacity (Kellogg et al. 2020; Langer & König 2021). As more data and multifaceted algorithms become available, a computer learns more complex patterns and "consequently builds its own representation of a classification decision, [which it does] without regard for human comprehension" (cf. Burrell 2016, p. 10). Therefore, algorithms exceed human abilities to understand the system and can generate severe trust issues (Arrieta et al. 2020). Furthermore, prior research observes attempts to manipulate and exploit these advanced ML-based systems, known as "algoactivism" (Kellogg et al. 2020; Meijerink & Bondarouk 2023), and a more general aversion to advanced algorithms, called "algorithm aversion" (Mahmud et al. 2022). Consequently, the successful leverage of the described potential of ML-based HRA critically depends on the ability to convince the individuals of an organization to use these systems (Di Vaio et al. 2022).

However, as most academic HR literature aims to understand the factors determining the adoption of HRA on an organization-wide level (e.g., Margherita 2022), and irrespective of the specific tool (e.g., Vargas et al. 2018), there is very little knowledge on the successful individual adoption of HRA – and especially ML-based HRA. Coming from the apparent need for a better understanding of the individual adoption process for ML-based HRA, as well as the ambiguous effect of ML characteristics, we ask the following two research questions:

RQ1: What beliefs and experiences influence the individual's intention to adopt ML-based HRA?

RQ2: How do the characteristics of ML engender these behavioral beliefs?

To answer these research questions, we examine the individual opinions and thoughts of employees of a public sector organization about a specific ML-based HRA tool for predicting voluntary turnover, the implementation of which the organization is currently evaluating. Drawing from the focused interviews method provided by Merton & Kendall (1946), we discuss the performance of the predictive HRA tool, as well as several explanatory figures, with employees in interviews and analyses their personal perspectives, experiences and spontaneous reactions to these different approaches. Following Vargas et al. (2018), we then interpret our empirical results with the help of a conceptual framework derived from the Theory of Planned Behavior (TPB) by Ajzen (1991). On the one hand, our results show that the perceived (self-) efficacy of interviewees also highly depends on the design of the HRA tool and the entered dataset, in addition to perceived skills and competencies. On the other hand, the attitude of the interviewed employees is not only formed by their personal enjoyment or concerns in terms of working with the tool, but also by the way in which they perceive it assists them in their daily work. We additionally identify that several ML characteristics (perceived self-learning capabilities, degree of automation, transparency and trialability) influence behavioral beliefs and in turn effect the adoption of the tool in HRM processes.

Our study makes three main contributions to the literature. First, it contributes to the ongoing debate about the relevant factors driving the decision to adopt HRA (Coolen et al. 2023). By examining this decision from an individual instead of an organization-wide perspective, we provide deeper insights into the different behavioral beliefs determining the decision to adopt ML-based HRA. Based on our findings, we propose several ML-related extensions and adjustments to the more general adoption framework of Vargas et al. (2018). Second, our study contributes to the current literature on ML design approaches and their effect on HRA adoption (Marler & Boudreau 2017, Langer & König 2021, Haque et al. 2023), and third, it contributes to research on ML transparency, suggesting that appropriate visualization influences end-user adoption (Haque et al. 2023). However, in contrast with Haque et al. (2023), our results demonstrate a lack of ethical reflection, as fairness plays no role in individual decisions to adopt ML-based HRA, albeit protected group differences were made apparent in the interviews.

The paper is organized as follows. The second section reviews the literature on the (individual) adoption of HRA, highlights the related limitations and derives the conceptual framework of our study. The third section summarizes the research method, empirical environment as well as the research object and data analysis process. The fourth section presents the results. A refined model at the end of this section summarizes the factors that influence individual intentions to adopt ML-based HRA and the impact of ML characteristics. Finally, in the fifth section, the results are discussed and propositions made before a conclusion is drawn.

2.4 Related research and theoretical framework

In the following section, a conceptual framework for the present study is derived by summarizing and discussing the state of knowledge on the (individual) adoption of HRA.

2.4.1 Prior research regarding the adoption of HR Analytics

The factors that drive or hinder the adoption of HRA have been almost exclusively explored from an organization-wide perspective (e.g., Margherita 2022; Böhmer & Schinnenburg 2023; Coolen et al. 2023). Prior research draws from the TOE framework (e.g., Pumplun et al. 2019; Chatterjee et al. 2021; Neumann et al. 2022b), with the underlying idea that the adoption of HRA from an organization-wide perspective is mainly driven by technological, organizational and environmental contexts. Technological contexts include, for example, the existing IT infrastructure of an organization (Neumann et al. 2022b), while the environmental contexts can be, for example, competitive pressure or customer readiness (Neumann et al. 2022b). The organizational context includes cultural aspects (such as the culture of innovation or change management) as well as resources (e.g., budgets or human capital) (Neumann et al. 2022b). Prior research concludes that the employees themselves – aligned with their skills and knowledge – play a major role in the adoption of HRA in corporations (Coolen et al. 2023; Di Vaio et al. 2022). Furthermore, work ethics (Basu et al. 2023) or supervisor support (Prikshat et al. 2023b) have been identified as additional major drivers for organization-wide adoption.

To the best of the authors' knowledge, only Vargas et al. (2018) have examined the individual adoption of HRA and proposed a comprehensive framework in this regard. Drawing from the Theory of Planned Behavior by Ajzen (1991) and the Innovation Diffusion Theory posited by (Rogers 2003), the authors explain the actual level of adoption of HRA through an individual's perceived self-efficacy, attitude and social influence regarding its use as well as trialability. Self-efficacy represents an individual's beliefs about their abilities to reach a behavioral goal (Bandura 1977), which translates to their evaluation of the technological and quantitative skills they deem necessary to adopt HRA. One's attitude towards a specific behavior is derived from the expected consequences of this behavior (Fishbein & Ajzen 2010). As the perceived consequences of using HRA partly depend on an individual's self-efficacy regarding the use of HRA, the latter will influence their attitude, among several other beliefs for the given context. Social influence represents the perceived norms in favor of or against HRA, and trialability encompasses beliefs about the degree to which HRA can be tested before adoption. Vargas et al. (2018) distinguish the three different decision-making steps of knowledge-gathering, persuasion and decision, whereby perceived self-efficacy is formed during the knowledge-gathering step, and attitude, social influence and trialability are derived during the persuasion step. The conducted survey empirically supports the proposed causal relationships as well as the effect of technology self-efficacy.

2.4.2 Limitations of the HR Analytics conceptual framework

While the derived conceptual framework provided by Vargas et al. (2018) extends the fundamental understanding of individual HRA adoption, it does have some limitations. First, it only includes trialability as a potential technological factor to distinguish between different HRA technologies. The proliferation of ML questions the reality, as the framework does not distinguish between the different characteristics of the HRA tool. Furthermore, as HRA includes many different algorithms, systems and methods (Meijerink et al. 2021), and prior research in information systems finds significant effects of an IT system's design on its subsequent use (Haque et al. 2023), there is clearly the need to further characterize and differentiate the proposed model from this perspective. Especially in the context of ML, research has emerged in the HR (Langer & König 2021), management (Glikson & Woolley 2020) and information systems (Arrieta et al. 2020) literature arguing that transparency must be another fundamental determinant of individual ML adoption. In contrast to traditional statistical methods in HRA, transparency is not always present in ML, because (a) predictors are not understandable, (b) relationships between predictors and predictions are hidden and (c) no explanation for a specific prediction is given (Arrieta et al. 2020; Burrell 2016; Langer & König 2021). This is problematic, because a prediction without clear explanations, or at least justification for the rationale behind the prediction, can lead to trust issues (Glikson & Woolley 2020; Langer & König 2021). Park et al., for instance, illustrate that only with sufficient transparency can various user burdens (emotional, mental, biases, etc.) be overcome during ML adoption. Transparency is also closely related to another fundamental determinant of HRA adoption, namely fairness (e.g., no discrimination against minorities), which can only be tested when professionals use their expertise and experience to determine the level of fairness of individual ML predictions through intuitive thinking (Chowdhury et al. 2022, p. 18). To achieve sufficient ML transparency, Explainable Artificial Intelligence (XAI) offers a rapidly evolving interdisciplinary research area with multiple technical solutions (Arrieta et al. 2020). Finally, the ability to automate decisions fully is an ML characteristic that represents a major shift for traditional HRA technologies (Meijerink et al. 2021, pp. 2545–2546). When algorithms are used for automated scenarios, they must also be accountable for the decisions they make (Busuioc 2021, p. 826). In summary, and in line with Lee & Cha 2023, we suggest that transparency, fairness and accountability (in terms of automated use) determine the adoption of ML-based HRA. Besides technical ML characteristics, decisions made during roll-out also affect the adoption of ML. In this regard, some studies have found that the ability to try (trialability) has a positive impact on adoption (Omrani et al. 2022). However, research is still inconclusive in terms of exactly how these ML characteristics influence an individual's adoption of HRA.

Second, the notion of self-efficacy and attitude in the proposed framework of Vargas et al. (2018) is relatively narrow, and it might exclude potentially relevant beliefs. Compared to assumed self-efficacy, the perceived behavioral control (PBC) factor from the original TPB is a wider concept that includes beliefs about factors beyond one's individual control (Ajzen 1991). It can be defined as the perceived ease or difficulty of performing a behavior (Ajzen 2002). For instance, the individual adoption of an HRA tool likely depends on the tool's suitability for a task and not only on one's perceived skills to use it. Furthermore, Vargas et al. 2018 examine self-efficacy regarding technology and mathematics in general, which are sufficient to estimate the average intention to adopt HRA but fall short when comparing the adoption of different HRA systems. However, the TPB is built upon the principle of compatibility, which states that the underlying factors must always refer to the underlying behaviour (Fishbein & Ajzen 2010). For the given context, one would therefore expect a notion of self-efficacy that is more directly connected to the individual adoption of a specific HRA tool or system. Furthermore, the attitude of a survey participant is derived from four beliefs solely centred around the personal enjoyment of using HRA (Vargas et al. 2018). This notion contrasts with the Technology Acceptance Model that connects the attitude towards a technology to the beliefs about the perceived usefulness and perceived ease of use of a technology (Davis 1989) and the Unified Theory of Acceptance and Use of Technology that connects the respective attitude to a performance and effort expectancy (Ajzen 2002; Venkatesh et al. 2003).

2.4.3 Underlying conceptual framework

Due to the limitations of the conceptual framework of Vargas et al. (2018) described herein, we aim to scrutinize the framework and extend it to ML-based HRA tools. Our further analyses are based on the assumptions described above, which are founded on the current state of knowledge. We also distinguish between the process steps of knowledge, persuasion and decision (Rogers 2003) in an ML-based HRA tool's adoption process. In the knowledge step, personal beliefs are evaluated regarding the ability to utilize an ML-based HRA tool for a given task and form an expectation about the PBC. In the persuasion phase, personal beliefs are evaluated regarding the consequences of using the provided HRA tool and form a tool and task-specific attitude. In addition, personal beliefs are evaluated regarding the opinions of others regarding the use of the provided HRA tool and form an expectation about the relevant social norm (corporate or national culture). In the decision step, personal beliefs are evaluated regarding the PBC, attitude as well as perceived norm and help decide whether to adopt the provided HRA tool. Furthermore, we expect PBC and attitudes to be influenced by the technical characteristics of the provided ML-based HRA tool, in which case we distinguish between the known characteristics of trialability, transparency, degree of automation and fairness and potential unknown characteristics.

2.5 Research approach

2.5.1 Method

To fill the derived conceptual framework for individual adoption with salient beliefs, it is necessary to dive deep into the line of reasoning employed by end-users. We aim to explore these beliefs by applying the "focused ethnographic interview" methodology proposed by Merton & Kendall (1946). We opted for a qualitative research approach because it can provide new insights into individual adoption in an explanatory manner. In addition, the open-ended nature of the interview questions allows for the collection of a wide range of information, including personal perspectives and experiences. The interview procedure was semi-structured around several pieces of information and nudges, used as potential triggers for spontaneous reactions. During the interview, detailed discussions were held on hypothetical but realistic implementation scenarios for the specific HRA tool. Particular attention was paid to employees' understanding of the presented tool, their ideas about its future use in HRM processes and their perceptions of the risks and benefits of using it in various HR applications throughout the organization. In addition, the interviews provided information about the overall intentions of the interviewees as well as any changes in their intention to adopt the HRA tool when providing various information and explanations. This approach follows the interpretive tradition of explorative methods, in that it seeks a deep understanding of human experience rather than rigid explanations of cause and effect – as in positivist epistemology (Einola & Khoreva 2023, p. 121).

2.5.2 Empirical Environment

This study examines a German federal agency from the social insurance industry with about 20,000 employees in the period between 2022 and 2023. The in-depth public sector study approach provides a context in which high legal requirements for the individual adoption of HRA can be investigated and commercial secrecy is not a concern (Desouza et al. 2020, p. 206; Busuioc 2021, p. 826). While the organization frequently uses descriptive analytics based on advanced dashboarding tools, as well as sporadic diagnostic regression-based analytics, this project is the first to incorporate complex ML models to implement predictive analytics use cases within HR.

Our main objective in selecting the interview population was to obtain a diverse sample of HRA users in terms of personal characteristics (age and gender), seniority and statistical background in order to represent the diverse workforce of the organization as well as the different usage objectives in the different personas. Table 2-1 provides an overview of the 12 interviewees. Team leaders supervising one to 21 employees, and heads of departments with 21 to 50 employees, from HR and operational departments, are the main users of the HRA tool. Half of the employees interviewed would work with the HRA tool in the near future, and half of those interviewed were potential recipients for further applications. Each interview lasted between 58 and 98 minutes.

Table 2-1: Interview population of future HR Analytics users.

#	Organisational section	Department	Position	Sex	Seniority(y)
I1	Corporate development	Employer branding & image	Team lead	w	20 to 25
I2	Internal corporate consultancy	Management of future vacancies	Team lead	m	15 to 20
I3	Insurance claim processing	Operational workforce management	Department administration	w	30+

I4	Insurance claim processing	Operational workforce management	Head of department	m	30+
I5	Human Resources	Organisation design	Organisational consulting	m	30+
I6	Human Resources	Organisation design	Team lead	m	0 to 5
I7	Human Resources	Organisation design	Department administration	w	25 to 30
I8	Human Resources	Personnel planning & controlling	Associate	m	10 to 15
I9	Human Resources	Personnel planning & controlling	Project lead	m	15 to 20
I10	Human Resources	Recruiting, development & diversity	Team lead	w	30+
I11	Human Resources	Strategic workforce planning	Senior data analyst	w	0 to 5
I12	Human Resources	Strategic workforce planning	Analyst	m	20 to 25

2.5.3 Research object

The specific ML-based HRA tool investigated herein predicts individual voluntary turnover (excluding age-related reasons and termination on the part of the employer) probabilities within the next 6 months for each employee, using the random forest algorithm (Breiman 2001). The tool is trained on a fully anonymized dataset with monthly data over a three-year time horizon and includes 30 predictors originating from the same federal agency in which the interviews were conducted. Work-related predictors include commuting distance, sick days, salary, salary increases in recent years, seniority and others. Demographic data such as gender, age, number of children and education level are also included. The ML predictions are evaluated in an out-of-sample test dataset. Instead of treating the ML model as a black box, post-hoc XAI explanations at the local (employee-specific) and global (organization-wide) levels are used to extract the effects of the predictors. The confusion matrix used to assess predictive accuracy, as well as some visualizations of the XAI results at the local and global level, were used as nudges during the interviews (see Figure 2-1). The visualization of organization-wide explanations describes how a single predictor influences the employee turnover prediction (strength, positive/negative contribution) on average, considering all employees in that local interval (Apley & Zhu 2020). Additionally, the visualization of employee-specific explanations breaks down the probability of voluntary turnover for each employee and quantifies it in terms of increasing or decreasing effects. The mean value represents the average employee turnover risk of all employees predicted in the model.

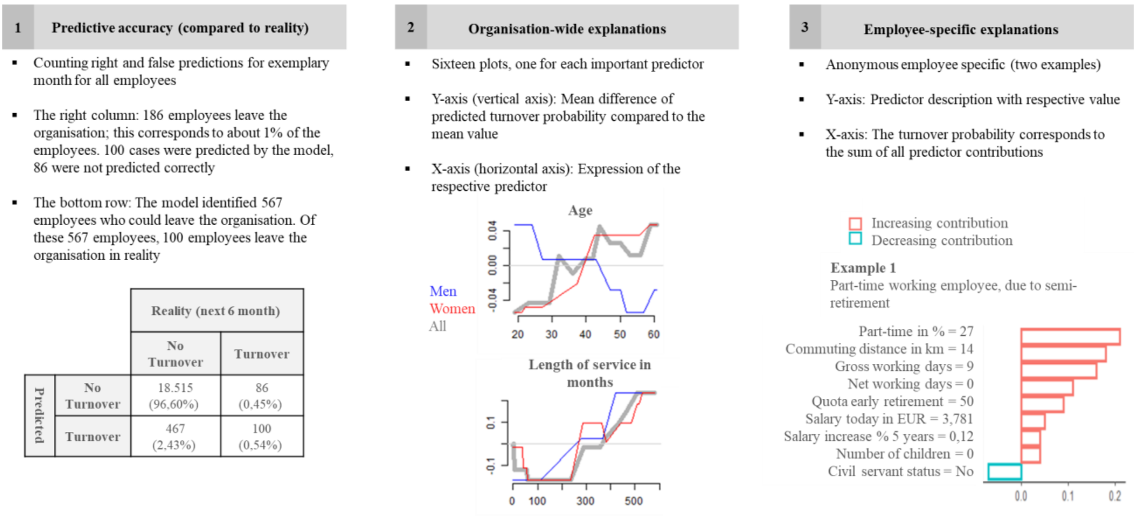

Figure 2-1: Information provided as nudges during the interviews: Predictive accuracy report, predictor effect explanations on the organization wide and employee specific level.

2.5.4 Data analysis

Given the inductive nature of the study, we coded the transcripts manually, following the methodology proposed by Gioia et al. (2013), which has demonstrated its validity in numerous renowned publications over the last decade (e.g., Friedman & Ormiston 2022; Schuessler et al. 2023; Mula et al. 2024). All interview transcripts were coded independently by the first and second authors using MAXQDA15 software. A total of 392 codes resulted. In the initial coding phase, we strictly adhered to the terms, phrases and descriptions of the interviewees, so that many first-order categories emerged. After eight interviews, both authors re-checked their coding to improve the reliability of the process and increase its rigor and authenticity. After coding all transcripts, first-order categories were compared against each other. Disagreements regarding interpretation, and thus coding, if any, were resolved through discussion. In the next step, for the second coding (axial coding), the first and second authors looked for similarities between and differences among the many first-order categories, in order to summarize and condense them. To this end, we went through each interview transcript as well as the first-order categories again. Subsequently, we discussed each passage and then reconciled different interpretations and conclusions to generate suitable second-order categories (Gioia et al. 2013).

Based on the TPB (Ajzen 1991), each of the second-order categories was independently assigned to PBC, attitude or norm (aggregated dimensions) by the first and second author (as determinants of an individual intention to adopt HRA) and then discussed. Subsequently, all second-order categories were critically reflected in correspondence with the framework provided by Vargas et al. (2018). The third author, who participated directly in project meetings and reviewed relevant project documents, critically reflected on the results in the final analysis step. Additionally, the coding of the entire interview material was repeated to verify validity. The re-coding of the first author, 11 months after the first coding, resulted in an overlap of 90.4% (intra-coder reliability). Coding by a person not previously involved in the research process resulted in an appropriate accordance of 79.0% (inter-coder reliability) (Miles et al. 2014). The interview coding process is summarized in Figure 2-2. The second-order categories and the aggregated dimensions formed the basis for the framework developed for the present study, and the first- and second-order categories, as well as the aggregate dimensions, became the basis for building the data structure (see Figure 2-3).

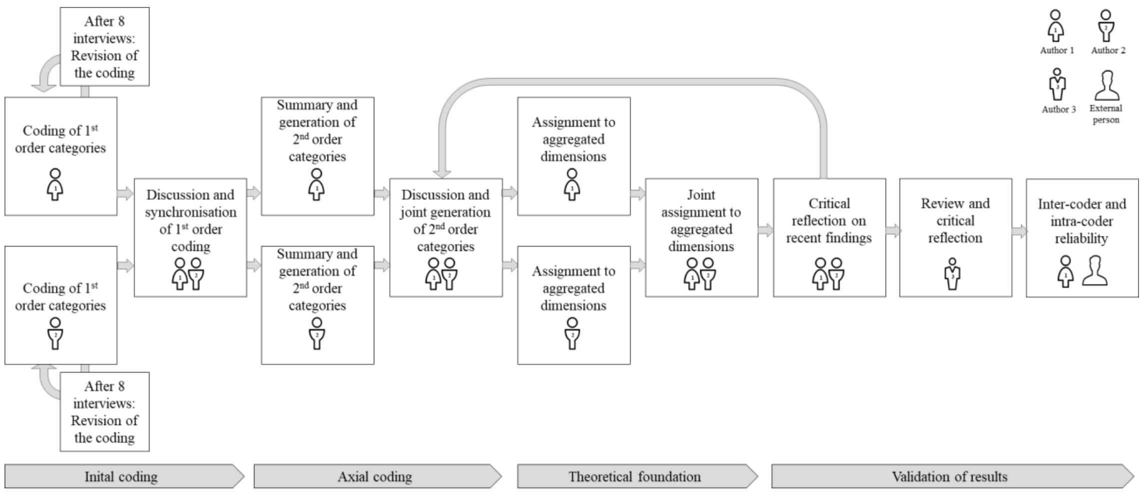

Figure 2-2: Interview coding process, including critical reflection steps to ensure reliability.

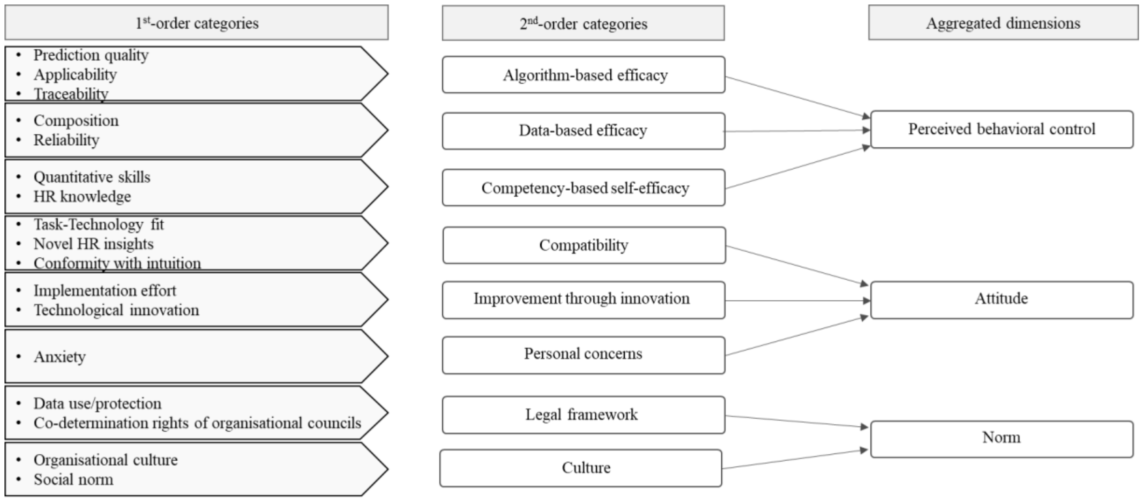

Figure 2-3: Data structure (first-order categories summarized by key topic).

2.6 Results

2.6.1 Factors influencing the individual intention to adopt ML-based HR Analytics

The results regarding the beliefs and experiences that influence individuals' intention to adopt ML-based HRA (RQ1) are presented below. First, the findings on PBC are illustrated (see Table 2-2), following which, after describing attitudes towards the adoption of the tool (see Table 2-3), the perceived norms (see, Table 2-4) of the interviewees are presented.

2.6.1.1. PBC

Algorithm-based efficacy describes all the capabilities and characteristics that interviewees attribute to and expect from an ML-based HRA tool, which allows them to utilize it in their daily work. Prediction quality captures the prediction accuracy of the tool, whereby the interviewees seem to require a sufficient level of prediction accuracy to view it as applicable. For our model, prediction accuracy is rated differently by I1 and I11. Another key factor for algorithm-based efficacy is traceability. The lines of reasoning by I10 reveal that the HRA tool’s feasibility stems from explanations of the ML model and how this helps to optimize the organizational retention of top performers at the individual level. I12 points towards the XAI visualizations provided as an important distinction compared to popular Generative Artificial Intelligence models such as ChatGPT. While skepticism towards these technologies is generally high, trust-building and achieving actionable insights can partially be attributed to traceability. The third driving factor for algorithm-based efficacy is applicability in practice. Among other things, the interviewees consider the extent to which the tool is mature and ready for use, its susceptibility to errors and the basic functions (e.g., the selection of different prediction periods) that it offers.

In addition, data-based efficacy plays an important role when forming the PBC. In our case, the opinions of the interviewees regarding the composition of the dataset differed widely. Some interviewees, like I2, identified missing and crucial predictors of voluntary turnover from their point of view, which had a critical impact on their evaluation of the tool. Others, like I9, were very satisfied with the included predictors. If they identified turnover predictors in the data, which they might consider important, data-based efficiency was considered high. Similarly, if the contribution and importance of the predictors in the XAI visualizations are as expected, then data-based efficiency seems to increase. Additionally, we find that dataset reliability plays an important role in the intention to adopt the HRA tool. For instance, I10 directly attributed the reliability of the department responsible for managing the database to the tool. Others, like I11 questioned the timeliness and quality of the data, its realism or rapid changes in the included turnover predictors.

In our study, competence-based self-efficacy reflects the beliefs an individual holds regarding their ability to use the ML-based HRA tool successfully in their daily work. Some interviewees, like I6, quickly understood the nudges shown, were interested in them and interpreted the information in detail. In their daily work, these people mostly take on analytical tasks and often have a background in statistics, which indicates that they have more pronounced quantitative skills. Others, like I8, were overwhelmed with the interpretation and had no deep interest in the information provided to them. In addition, the interviewees considered it necessary to have a certain level of HR knowledge, to be able to apply the results of the tool in practice and to derive potential application scenarios. The findings on our interviewees' PBC-related adoption of ML-based HRA tools, which are presented in Table 2-2, match interview insights into the dimensions related to the Theory of Planned Behavior.

Table 2-2: Interviewees' PBC related adoption of ML-based HRA tools.

Example Insights	1st-order Categories	2nd-order Categories	Aggregate Dimension
"*... actually, quite impressive prediction quality.*" (I1, 34)	Prediction quality	Algorithm-based efficacy	PBC
"*... if you look at the absolute numbers, 467 and 100, prediction accuracy doesn't look so great.*" (I11, 102)			
"*I understood that the higher the absenteeism due to illness, the higher the probability that these people will leave the organization, which would allow me as a personnel manager to conclude: How high is my sickness rate? [...] Unfortunately, my department has a very high sickness rate, and it would be exciting for me to see whether this has led to increased turnover – that the sick days were perhaps even the criterion.*" (I10, 104)	Traceability		
"*I do not know how to calculate the predictions. I do not know how the database must be prepared, how the model must be fed and so on. But if I look specifically at the XAI visualizations, I can already work with that. […] You first must deal with it [like a new software program] to be able to use it. [...] A little bit of skepticism is quite healthy, but ChatGPT has now increased our trust somewhat.*" (I12, 171)			
"*I would need 5 or 10 years and not the individual level of an employee, but I would have to look at the entire department, and I would have to look at certain levels, e.g., professional groups.*" (I11, 162)	Applicability		
"*All these flexible working time models with remote working etc. are not integrated into the model. After all, these affect 50% of our employees.*" (I2, 108)	Composition of the dataset	Data-based efficacy	
"*These are very important predictors that I could then use for the future. So that would be very helpful, very helpful.*" (Interview 9, 92)			
"*So you certainly calculated this from the data collected by Mr. […], I assume? From there, the data basis is safe for me – and from there I also trust in the numbers.*" (I10, 242)	Reliability		
"*So, for example, the economic situation, the issue of security, the issue of a personal family situation. The factors change. I hire someone who does not have any children, and then I know, well, maybe in 5 years there will be children. That means I cannot exert any influence. Likewise, what about changing health situations?*" (I11, 128)			
"*I understood the figures shown.*" (I6, 85)	Quantitative skills	Competency-based self-efficacy	
"*But at the moment, it slays me and everything – honestly.*" (I8, 249)			
"*I cannot say anything about that [how to adopt the tool specifically]. My colleagues in the HR department are more closely involved in this issue. I cannot assess the potential.*" (I12, 167)	HR knowledge		

2.6.1.2. Attitude

The interviewees' attitude towards the adoption of HRA was influenced, among other things, by its perceived compatibility. The assessment of the task technology fit was very different (see I3 and I7), with the added value and the concrete integrability of the ML-based HRA tool in everyday work being questioned and analyzed. The interviewees' attitudes also seemed to be influenced by whether the results of the HRA tool revealed novel HR insights. Like I9, almost all interviewees stated that the tool could help identify at least some novel factors for employee voluntary turnover – and thus provide a basis for the development of personnel measures. It is notable that the consistency of the tools' predictions with personal intuition is an important determinant of attitude. Provided that the results matched the intuition, this manifested in an improved attitude, and vice versa.

In addition, considerations related to the improvement achieved by the tool seemed to have an impact on attitude. In our study, the extent to which implementation effort and technological innovation were perceived as impactful by the interviewees was important in this context (see Table 2-3). Some, like I4, felt that the innovative nature of the HRA tool enables new approaches to old challenges (such as demographic challenges) and improves previous processes (e.g., the quality of workforce planning). Essentially, innovation brings new perspectives and approaches. Others, like I11, critically questioned implementation efforts in terms of a cost-benefit trade-off.

A few interviewees questioned the personal consequences of adopting the HRA tool and evaluated them accordingly. I9 and I11 were particularly afraid that superiors use the HRA tool inappropriately (e.g., findings led to monitoring by the superior or mobbing), that false predictions led to negative effects (e.g., in the allocation of tasks) or that misunderstandings occurred. Findings relating to the interviewees' attitude to the adoption of ML-based HRA tools are presented in Table 2-3.

Table 2-3: Interviewees' attitude to the adoption of ML-based HRA tools.

Example Insights	1st-order Categories	2nd-order Categories	Aggregate Dimension
"*I just wanted to add, because I think that our organization is so big, that you do not have to look at individual employees. I cannot use that method at all.*" (I3, 109)	Task technology fit	Compatibility	Attitude
"*Through the tool, managers are encouraged to be active in their role.*" (I7, 362)			
"*Then I can look at this and analyzes the most important factors that influence why this employee is*	Novel HR insights		

leaving and use this to initiate optimization." (I9, 120)		
"*Of course, this creates trust when you see that even without algorithms.*" (I3, 64)	Consistency with intuition	
"*No, there must be a mistake. It says that the probability of turnover is higher for civil servants.*" (I5, 115)		
"*I see this as a great support, and it goes much further than what we could do in the past. [...] You can draw insights from the data that give the organization a positive kick in any case.*" (I4, 283)	Technological Innovation	Improvement through innovation
"*I think it is great that such an approach has been found at all.*" (I6, 160)		
"*It is always nice to try something out, but of course, the question is then always cost and benefit. Does it bring us anything?*" (I11, 162)	Implementation Effort	
"*I am afraid of the surveillance now that the supervisor monitors me like this: do I go or not?*" (I11, 206)	Anxiety	Personal concerns
"*On the negative side, my responsible tasks could be taken away from me, because there would be a risk that I would leave the organization.*" (I9, 212)		

2.6.1.3. Norm

The adoption of (ML-based) HRA tools is also limited by the ‘legal framework’. The interviewees stated that possible applications of the tool were severely limited or not possible due to legal conditions and the strict interpretation of data protection regulations in the public sector. Interestingly, some interviewees mentioned experiences with – from their point of view – overly strict data protection rules for historical organizational initiatives. For example, I8 stated that sensitive personal data should also be included in the tool and used for individual decision-making. Organizational councils have far-reaching co-determination rights that go beyond the law and enable employee representatives to object to various decisions affecting the entire organization, thus obstructing adaptation. The interviewees perceived those initiatives based on employee data were – in principle – prevented. Among other things, organizational councils receive evaluations of all HR reports requested in the IT system, and they strictly ensure that each employee processes only the amount of information needed to complete tasks.

Culture is an important factor in the adoption of HRA (Vargas et al. 2018), where we distinguish between the social norm and organizational culture. During our interviews, some interviewees critically evaluated their social norm's compliance with the adoption and use of the HRA tool. For example, in terms of the individual employee’s privacy, they questioned whether the analysis of personal data was acceptable from their point of view, or with whom the responsibility for ensuring appropriate use lay. Others, like I11, saw no threat to (personal) privacy. From the interviewees' responses, we were also able to find indications of the anchored organizational culture, which in our case tends to have a hindering effect on the HRA tool. The interviewees stated that there were many people with reservations and sceptics who viewed changes to previous processes or systems as negative; in addition, decision-making processes within the authority were often perceived as not rational and were very time-consuming. Moreover, the adoption of the HRA tool was a complicated undertaking because employees had difficulty dealing with predictions and uncertain expectations. Interview insights on subjective norms regarding adoption of ML-based HRA tools are presented in Table 2-4.

Table 2-4: Interviewees' norms regarding the adoption of ML-based HRA tools.

Example Insights	**1st-order Categories**	**2nd-order Categories**	**Aggregate Dimension**
"*Data protection is a very important topic in our organization. [...] Because we run analyses here that can be evaluated on a personal basis, and conclusions can be drawn about a person*" (I6, 180)	Data use/protection	Legal Framework	Norm
"*This is very sensitive data with which the tool works – highly explosive in terms of data protection. Therefore, it cannot be implemented in this form.*" (I7, 326)			
"*And when it comes to data protection […] I think [it] tends to protect those who have something to hide rather than benefit others. […] Instead of excluding variables, you might have to take other variables in addition*." (I8, 351-359)			
"*The problem is: The implementation of tools like this in-house must be approved by the organizational councils. From my work as an organizational consultant, I also know that software like this is not simply approved*" (I5, 169)	Co-determination rights of organizational councils		
"*That would just be too much intrusion into my personal life for my supervisor to have that information to hand.*" (I9, 258)	Social norm	Culture	
"*My boss has all my data at his disposal. He knows how many children I have, and he also knows where I live. He also knows when I'm sick and how much I earn.*" (I11, 214)			
"*We have many doubters – there is not only the political thing in the house.*" (I7, 366)	Organizational culture		
"*The management always tries to be supportive, of course, but decisions are not made as quickly as in the private sector.*" (I9, 240)			

2.6.2 Impact of ML characteristics

The results for how the characteristics of ML affect behavioral beliefs (RQ2) are presented below. Our findings are successively illustrated in terms of trialability, transparency, automation, self-learning capabilities and fairness, as well as their effect on beliefs and experiences.

2.6.2.1. Trialability has an impact on attitude

Overall, we observe a positive effect of trialability on the intention to adopt the provided HRA tool. Similar to the findings of Vargas et al. (2018), our interviewees believe that it is important to try out the ML model before it is implemented in the organization, in order to gain experience of using it. They argue that a high level of trial and error makes it easier to assess the accompanying consequences of actually applying the ML model, which in turn could lead to an increase or a decrease in one's attitude regarding the tool. On the one hand, trialability helps to assess whether the ML model provides a presumed improvement through innovation (technological innovation):

> *“I like to try something like this out in practice [...]. ML does not really help here yet. I think we always have to make our own experiences with applications. [...] They have to prove themselves in practice somewhere. And if they do not, then I have to analyze that. Where is the problem, or where does it not bring the benefit that I had hoped for? And, if necessary, I have to adapt it.” (I4, 283)*

On the other hand, trialability helps to mitigate any potential personal concerns of employees:

> *“For matters that are more critical, it is wise to first try things out, test them, see where adjustments can be made, involve the people and initially test it in a small area to then see how it is received [...] But having these sceptics around all the time makes everything a bit more difficult.” (I11, 222)*

2.6.2.2. ML transparency has an impact on both attitude and PBC

We observe positive and negative effects of transparency on the intention to adopt the provided tool. At the beginning of the interviews, we asked the interviewees about their intention to adopt the employee turnover predictions in their daily work. Interestingly, most initially saw little to no application in the tool's predictions when it came to pure predictive accuracy without understanding the effects of the predictors:

> *"Unfortunately, I am not able to determine the value added because I have not performed any [proving] calculations [...]. Therefore, I could honestly plan better for the future based on historical data." (I9, 98)*

However, the more transparency provided by the presentation of multiple predictor effects, the more diverse and extensive the applications identified by the interviewees (in their areas of responsibility), improving their attitude via the perception of compatibility of the tool and especially the personal task-technology fit. Besides reflecting on how the predictions could be used (e.g., for workforce planning and identifying future staff shortages), the interviewees also recognized that the tool provides explanations for turnover. Thus, it offers opportunities to either mitigate turnover at an individual level or derive strategic and organization-wide initiatives that address employee wellbeing (e.g., increasing remote working opportunities) and employer attractiveness (e.g., increasing childcare offerings). The discussions in all interviewees about possible applications of the tool in other HRM processes, made possible by transparency, indicate a higher algorithm-based efficacy.

In addition, providing more transparency can have a positive or a negative effect on one's attitude when the derived predictor effects contradict personal intuition (compatibility). On the one hand, our interviewees found contradicting evidence useful in questioning their personal intuition:

> *"But it definitely brings insights that straighten out the picture and probably bring it closer to reality. Yes, I would use it if I had to decide for my hotdog stand." (I1, 135)*

On the other hand, a few interviewees questioned the functionality of the provided tool when identifying evidence that contradicted their own intuition (conformity with intuition, Table 2-3). Furthermore, our results suggest that these interviewees demanded a high degree of traceability to help them understand the underlying calculations of the ML model (algorithm-based efficacy):

> *"I [...] want to understand what is happening behind the system, [...] In Excel, you can see how the calculation is done and what the result will be. With machine learning, you probably won't be able to see it that way. The machine learns based*

> *on the data and then outputs something. So, I always need a certain level of traceability for each step." (I9, 294)*

To summarize, we find that transparency influences attitudes via the perception of compatibility in two ways (personal task-technology fit and conformity with intuition), as well as PBC via algorithm-based efficacy, also in two ways (applicability and traceability).

2.6.2.3. Degree of automation through ML decision-making influences attitude and PBC

We mostly observe a negative effect of the degree of automation on the intention to adopt the provided ML model. All interviewees agreed that decisions should only be augmented with the help of the tool and that a fully automated decision-making process should not be implemented. Several reasons regarding attitude, especially personal concerns, were given for this, such as the fact that the interpersonal component must not be lost, especially when decisions are made on an individual basis:

> *"At the top level, they want numbers, and there's also the risk that when they see those numbers, they do not want to deviate from them [...]. However, the human factor, and the perspective and the focus on the individual employees, is simply lost as a result. The decision-makers who normally have management responsibility, who actually manage people, have to look at the results." (I1, 52)*

I4 pointed out that automation is only useful if the model does not make a single mistake. This in turn is reflected in the expected accuracy of the tool (algorithm-based efficacy):

> *"[For automation], the probability of correct predictions is not yet high enough, not until the hundred per cent mark is reached. Until then, decisions are up to personnel analysis by management – instead of letting the machines think completely." (I4, 275)*

The interviewees believed that the responsibility and rational towards decisions lies with humans (an ML-augmented decision process) and questioned whether the provided ML model is suitable for drawing the right conclusions and deriving appropriate actions from a prediction. This translates to a low perceived applicability (algorithm-based efficacy):

> *"If you were to go only by the machine: a woman has a salary of 3,000. We will just raise it to 4,000 – but a man does not get that raise. [...] I would see it critically*

> *in the first instance. In any case, it does not replace the interpersonal connection. Well, I do not work together with the machine, especially not in a subordinate relationship. Ultimately, such a decision must be made by a manager." (I7, 402)*

2.6.2.4. The self-learning capabilities of ML affect PBC

Unexpectedly, we identified the perceived learning capabilities of the provided ML model as a further relevant ML characteristic influencing the algorithm-based efficacy (PBC). A few interviewees associated continuous learning with ML and expected continuously increasing accuracy due to future learning iterations with more data or feedback loops:

> *"[With] ML and Artificial Intelligence work – as far as I have now generally heard – the more you feed, let's say, the machine with information, the better it becomes. And that's exactly the direction it should go if you use it more often and feed it with more and more data. It will get better and better, and that will also reduce the error rate, in my opinion." (I9, 182)*

Interestingly, some of the interviewees translated the automated self-learning characteristics they were aware of from a reinforcement learning ML model in another context to this specific ML model, without knowing whether these feedback loops were actually implemented:

> *"[ML]… is a self-learning system, and the more often I run it, the better my predictions become. In this respect, if I have understood correctly, we are still at the start. And the more data is fed in and compared with real things, the more accurate the predictions will be – at least that's what I would expect." (I2, 85)*

2.6.2.5. (Un-)Fairness does not affect the intention to adopt

As noted in the literature review, increased ML transparency can also affect perceptions of ML fairness. In the interviews, we specifically asked about the fairness perception in hypothetical scenarios (e.g., What requirements do you have for the model in terms of fairness or equal opportunity? Would you remove certain data from the dataset, for example, for reasons of fairness or equal opportunity? From the position of your supervisor making decisions about you, do you have any reservations or concerns about using the model?). Interestingly, none of the interviewees mentioned significant caveats regarding fairness aspects. For example, excluding protected group variables from the dataset was not suggested by any interviewee. We were able to ascertain this even after providing a list of predictors used in the ML model as well as XAI visualizations that (1) listed protected group variables such as gender or age and other data that require extensive protection, such as health-related information, (2) clearly documented differences between protected group variables and their impact and (3) were considered the basis for local (employee-specific) or global (organization-wide) decision-making – and thus varying degrees of impact on individual employees. Several arguments were made by the interviewees as to why the HRA tool in its current state is fair and no adjustments are needed. First, I9 referred to the objectivity of the data and the responsibility for any consequences:

> *"You cannot influence the fact that our organization is over 70% women, so you cannot do anything by saying that we now have to hire only men. That would be discriminatory. That is why I do not think it is so bad. Those are the facts, that is the database, you cannot change that. Or that older workers are less likely to quit. As a human being, as a decision-maker, you also have the information, and you have to interpret it accordingly." (I9, 280)*

Second, I1 made a similar argument, pointing out the possible lower prediction accuracy of the tool when predictors are eliminated. The interviewee argued that differences are not unfair if they are based on differences between protected groups that can be explained by real facts, which he exemplified in terms of intergenerational differences:

> *"My father was in the same company for 30 years, my mother worked in the same company for over 40 years. So that is the thing, for me, that would be unthinkable. I would not exclude predictors like age or gender […]. I can understand that you want to leave such factors out so as not to discriminate against anyone […], but it would also just be out of touch with reality. […]. I would claim that there are significant differences, and that tells me that it must not be left out at all." (I1, 151)*

Overall, we found no evidence that the interviewees had changed their intention to adopt the HRA tool due to the different treatment of women or men, or younger and older employees. This can be explained by the above statement that potentially unfair discriminatory decisions should be corrected by human judgment in an augmented (non-automated) decision-making process.

2.6.3 Refined model: Individual intention to adopt HR Analytics

The results of our study are summarized in a qualitative model, as illustrated in Figure 2-4. Vargas et al. (2018) framework forms the basic structure on which the various factors influencing PBC, attitude and norms are concretized. As a further addition, the influences of ML characteristics emerge. Please note that ML characteristics have effects on zero (fairness), one (self-learning capabilities) or two (transparency, automation and trialability) constructs of behavioral beliefs.

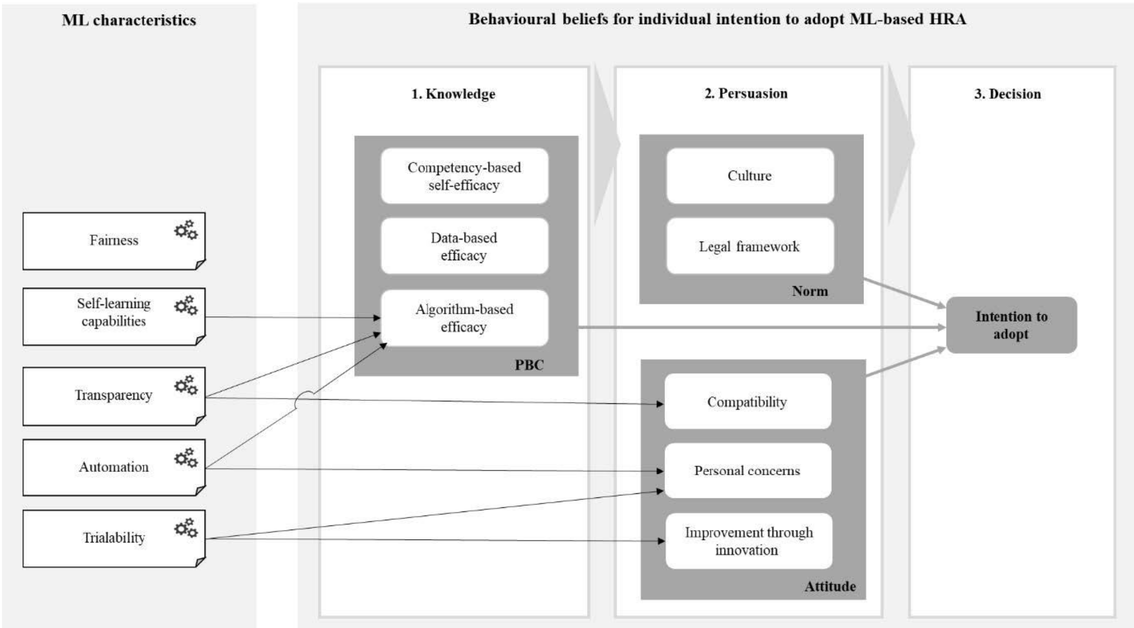

Figure 2-4: Proposed qualitative model for individual intention to adopt HRA based on ML characteristics.

2.7 Discussion and implications

Notwithstanding the asserted importance of HRA, research examining its impact on organizational performance remains underdeveloped (Marler & Boudreau 2017, p. 15). While recent studies find evidence for the effects of HRA and organizational performance, this link is mediated by an organizational shift to more evidence-based management practice. Moreover, it is argued that HRA can only provide a benefit for an organization when predictions and estimations are incorporated into neutral and evidence-driven decision-making (McCartney & Fu 2022, p. 38). Evidently, a respective shift requires a congruent employee mindset and therefore a strong intention to incorporate HRA into their daily work. Our study contributes to this ongoing debate by extending and contextualizing current knowledge on the adoption of ML-based HRA for a specific use case.

2.7.1 PBC, attitude and norms influence an individual's intention to adopt HRA

Regarding salient beliefs and experiences, Vargas et al. (2018) find five factors determining the individual adoption of HRA: technology self-efficacy (PBC), quantitative self-efficacy (PBC), attitude towards using HRA, social influence and tool trialability. Vargas (2016) investigates a larger number of possible factors influencing the user level of adoption, whereby general self-efficacy and data availability, which were queried in the survey with narrowly specified items, showed no significant influence. Therefore, we investigate the personal perspectives, experiences and spontaneous reactions of knowledgeable employees in a specific use case with field data. This particularly provides a lens for a deeper understanding of the individual adoption of ML-based HRA. For example, we find evidence that interviewees with higher quantitative skills, as well as overall competency-based self-efficacy (e.g., HR knowledge), were more open to incorporating the displayed HRA tool in their work. In addition to technological aspects, the input data on which the ML-based HRA tool is trained was highly important to the interviewees. We therefore suggest that data-based efficacy, as an aggregation of the composition and reliability of the dataset, is another important dimension of a potential user's efficacy. The findings published by Omrani et al. (2022) suggest similar relationships in terms of data, albeit their study argues that concerns about discrimination in the use of Artificial Intelligence reduce trust. In addition, we find evidence that attitudes toward the adoption of HRA are made up of a wide variety of aspects. These can be divided into three categories, namely the assessment of the tool's compatibility in daily work, an assessment of the potential for improving the tool in organizational processes and personal concerns. Venkatesh et al. (2003) find similar results in their model of technology use. In their work, performance expectancy, as the "extent to which an individual believes that using the system will help them improve their job performance", represents a relevant factor similar to the perceived usefulness of the Technology Acceptance Model (Davis 1989). Regarding the norm, the interview data also divides this factor into legally established norms and (organizationally) culturally determined aspects.

Proposition 1: For ML-based HRA, an understanding of related studies is not sufficient to explain individual adoption. Instead, we propose additional important determinants for PBC (data-based and algorithm-based efficacy), attitude (compatibility of tools and tasks, personal concerns, improving practices through innovation) and norm (organizational culture and legal framework).

2.7.2 Most ML characteristics have an influence on behavioral beliefs

Furthermore, our results suggest that several additional ML characteristics drive perceived algorithm-based efficacy as well as attitudes to the use of the displayed ML tool. First, to make ML-based HRA useful, predictors' diagnostic results – provided by the XAI visualizations – are an important facilitator in terms of coalescing employee turnover predictions (Chowdhury et al. 2022). Identifying additional uses of the HRA tool when understanding the causes of employee turnover suggests a more evidence-based management practice, which is theorized as an important enabler of HRA in an organization (McCartney & Fu 2022, p. 29). The studies by Kim et al. (2023b) and Haque et al. (2023) reveal that XAI visualizations in particular contribute to user understanding and adoption, when appropriately designed. To summarize, in most cases, higher transparency leads to a higher attitude and PBC. We thus provide an empirical example demonstrating that with sufficient ML transparency, the various burdens on users (emotional, mental, prejudices, etc.) (Park et al.) can be overcome when introducing ML-based HRA. However, in line with other research (Schmidt et al. 2020a), we also find that these effects can be reversed when the rational explanations of the model and the reasoning of experts contradict each other.

Second, our interviewees were reluctant to automate an entire decision, for example a promotion, by delegating it to the HRA tool, as most of them did not expect the respective tool to possess the necessary skills to solve the task adequately on its own. This finding is similar to the results by Dietvorst et al. (2018), who identified a significant aversion to fully automated predictive analytics tools that vanishes when participants get at least some degree of control over the underlying decision. Lee & Cha (2023) confirm this notion by showing that choosing augmentation over automation is one of the two key factors in adopting Artificial Intelligence recruitment systems.

Third, we observe a difference in the beliefs of the interviewees, who expect the displayed HRA tool to have self-learning capabilities or not. In sum, our empirical data indicates that interviewees are more forgiving of an error-prone prediction in the first case, as maybe because they expect the HRA tool to subsequently improve upon its past mistakes and thus increase algorithm-based efficacy. Note that this finding is in line with (Reich et al. 2023) and Berger et al. (2021), both of whom identify self-learning capabilities as an important factor in mitigating algorithm aversion. Our study extends these interesting points by the fact that self-learning ability may be taken for granted by using the ML term, while such abilities (e.g., reinforcement learning) are not even implemented.

Fourth, as proposed by Vargas et al. (2018), our interviewees were interested in trying out the displayed HRA tool to assess its capabilities and then form an attitude on it. In addition, indicators of initial anxiety about the ML-based HRA tool's capabilities were apparent (see Table 2-2), which can be addressed with trialability. This finding is in line with the Innovation Diffusion Theory (Rogers 2003). In summary, we therefore propose:

Proposition 2: Several ML characteristics influence attitude and PBC in relation to the intention to adopt ML-based HRA: (a) the degree of transparency created, (b) the choice of automated usage, (c) the implementation of self-learning capabilities and (d) the enabling of trialability.

Consistent with Neumann et al. (2022b), we find no reference in the interviews to ethical considerations regarding the adoption of the ML-based HRA tool in our empirical setting from public sector. Our results show that fairness and non-discrimination were not critically questioned, even when potential biases were highlighted by XAI visualizations and the interviewees were explicitly asked about them. This outcome is alarming, as HRA, and especially ML, can foster discrimination and create various risks for employees and the organization (Tursunbayeva et al. 2022). For example, according to the General Data Protection Regulation of the European Union, the use of an ML model that includes protected class variables for individual decision-making can be considered a legal case of discrimination, referred to as “disparate treatment” (Goodman & Flaxman 2017). Lee & Cha (2023) confirm that solving the fairness problem remains complex, even if this complexity mitigates discovering an unfair decision basis.

Proposition 3: Fairness does not matter for the individual deciding whether to adopt ML-based HRA and must be ensured by other appropriate measures.

2.7.3 Implications for practice

In practice, organizations seeking to leverage the potential of efficiency gains through ML-based HRA might try to increase adoption at the individual level. Our results reveal that various adjustable modifiers exist during adoption, in particular the degree of automation in algorithm-based decision-making and provided transparency. Most importantly, technical measures (such as XAI) can positively or negatively influence both PBC and attitude, because providing understandable visualization allows the user to compare the included predictors, as well as their effects with their intuition.

However, in trying to increase the intention to adopt ML-based HRA on an individual level, organizations should be careful to avoid pitfalls. For instance, negative examples have already demonstrated that biases can lead to unfair decisions based on ML (e.g., Alon-Barkat & Busuioc 2023). With the risk of resulting high social and economic damage, the consideration of ethical challenges is necessary in HRA projects (Langer et al. 2023; Edwards et al. 2022, p. 5). As legislation for the responsible use of ML comes into force soon (High-Level Expert Group AI (EU) 2019), organizations need to address such potential unfairness proactively. Thus, our finding, namely that ethical considerations and fairness of HRA in the early adoption stages were not challenged by the interviewees, is alarming and should therefore be paid careful attention in practice. Neumann et al. (2022b) note that specifically the early adoption phases of ML applications are characterized by (1) a focus on positive business cases, (2) reliance on external partners, (3) change management processes to increase acceptance and (4) little to no real recognition of ethical considerations such as algorithm accountability and fairness. Thus, we specifically advise organizations to ensure proactively the inclusion of ethical considerations in the early stages of adoption and to implement internal policies and approval procedures with the help of internal or external expertise.

Ultimately, the responsible use of ML-based HRA can only be achieved when HR professionals have the knowledge necessary to evaluate ML models critically, based on the transparency provided by technical measures such as XAI visualizations and internal guidelines (Langer & König 2021). However, Vargas et al. (2018) note that HR professionals have low levels of quantitative self-efficacy (fear of math/statistics, lack of quantitative training, low awareness of analytics, lack of resources and organizational support to promote analytics and its tools). Our results extend these findings, which suggest investing in training initiatives that demonstrate the importance of achieving ML transparency and in turn encourage the acquisition of skills specifically to interpret performance statistics of ML algorithms or XAI visualizations.

2.7.4 Limitations and further research

There are two main points that limit the findings of this study. First, our qualitative approach is based on the manual coding of interview transcripts; however, we took several measures to ensure the validity of our findings during the coding process and after the final analysis (see Figure 2-2). For example, the credibility of our findings was established by independent coding by two of the authors in three coding steps. In addition, our results were critically reflected on the basis of existing evidence (Vargas et al. 2018) and by the third author. The findings were verified by inter-coder and intra-coder reliability (Miles & Huberman 1994).

Second, while the methodological choice of a single case study has a solid foundation in HRM, and recent studies using this method advance the field considerably (e.g., Ellmer & Reichel 2021; van den Broek et al. 2021; Remneland Wikhamn et al. 2023), this methodological choice limits the transferability of our findings (Flyvbjerg 2006). Nonetheless, it offers the advantages of an in-depth investigation of HRM practices with a heterogeneous interviewee population (e.g., diverse backgrounds and experience) and an examination of deep cause-effect relationships (from ML characteristics to HRA adoption) that are overlooked in broader studies. In our particular case, the results thus pave the way for future quantitative studies that can explain the individual adoption of HRA more holistically and further develop the previous framework by Vargas et al. (2018), which can only explain about 35% of the observed variance. To achieve this, the exploratory and qualitative nature of our study leaves the following concrete possibilities for future research. First, future studies should examine multiple organizations to further validate the transferability of the three propositions for individual HRA adoption. Second, we invite future research to formulate and quantitatively test hypotheses based on our proposed qualitative model. Especially, the effects of the automated usage of ML predictions and ML transparency provide interesting opportunities in this regard, as they both affect PBC as well as attitude.

In addition, our study is not able to provide insights into the effect strength of the assumed causal relationships between ML characteristics and the intention to adopt. Third, we focus on the first implementation of an ML-based HRA tool, which means that the key beliefs and experiences identified, as well as underlying ML characteristics, may not apply to a more mature stage of ML adoption. Therefore, given that information systems research has found a considerable number of factors influencing the intention to use HRA (e.g., Mahmud et al. 2022), future research could investigate whether the effects and significance of certain factors change over the course of the implementation and utilization phase. For example, does the importance of ML transparency decline as users of ML-based HRA gain experience over time and learn that the system provides (in-)accurate results? Fourth, we agree with the widespread view that HRM systems need to be tailored to the individual case (e.g., Remneland Wikhamn et al. 2023), which is why we also call for more qualitative research examining the individual adoption of technological advances in ML-based HRA.

2.8 Conclusion

In contrast to existing technological tools, ML-based HRA generates unique challenges, most notably the potential opacity of the rationality used by models to formulate predictions, as well as the potential to automate HRM decision-making fully. This study provides deeper insights into behavioral beliefs determining the decision to adopt ML-based HRA from an individual perspective and sheds light on how ML characteristics affect it. Based on the focused interview methodology, we introduce novel propositions and an extended qualitative framework with new constructs of important factors from the perspective of end-users of individual HRA adoption. Investigating the lines of reasoning also reveals that potential ML model users do not include fairness considerations in their decision to neglect or adopt the tool. We hope our findings help to guide both the interdisciplinary research on HRA and organizations to a successful path in their mission to achieve the responsible proliferation of ML-based HRA.

3. Study 2: A Comprehensive Framework for Algorithm Aversion in Business Contexts

3.1 Description of the research project

This research project is a single-author project of the author of this dissertation. The idea for this study arose from an in-depth literature search of the author on the topic of "algorithm aversion", Originally, the identified literature was intended to serve as the basis for hypotheses development and the design of a behavioral experiment. However, the identified literature revealed that there were already dozens of published experimental papers on algorithm aversion identifying different antecedents of the behavior following diverse research paradigms.[12] Instead, prior research on algorithm aversion called for more theorizing on the underlying reasons why this behavior occurs (e.g., Jussupow et al. 2020). To address this call for further research, to differentiate the project from the existing literature reviews (Burton et al. 2020, Jussupow et al. 2020) and to account for non-linear empirical findings (Dietvorst & Bharti 2020), the project first applied the research method of analytical modeling (e.g., Dikolli et al. 2013, Heinle et al. 2012). However, due to feedback from several conferences, the method was changed to conceptual theory building. The present study was accepted for presentation at the Annual Conference for Management Accounting Research (ACMAR) 2024 and the Annual Meeting of the European Accounting Association (EAA) 2024. At both conferences, the paper was accepted for presentation in "parallel sessions" suggesting a high level of quality and advanced level of maturity.[13] Furthermore, the present study was accepted for poster exhibition at the Management Accounting Section (MAS) midyear meeting of the American Accounting Association (AAA) 2024.

Positioning this research project in the established streams of literature has proven to be challenging, First, prior AIS-related literature has heavily criticized the relevance of algorithm aversion for the profession and has called the empirical findings "business press folklore" and "old wine in new bottles" (Sutton et al. 2023). Second, the paper was rejected for presentation at the MAS midyear meeting because of doubts regarding its management accounting-related contributions. While the anonymous reviewer suggested the

[12] For example, see the literature reviews of Burton et al. (2020) and Jussupow et al. (2020), which were already published at the start of the project.

[13] For example, see the information on the EAA Annual Meeting 2024 Review Process (EAA Congress 2024)

renowned journal “Academy of Management Annals” (VHB-Jourqual4: A+ across several disciplines, Impact Factor (2024): 16.5) as an outlet for the paper, a submission was deemed to challenging and time-consuming for a single-author and PhD candidate. Third, the recent publication of Jussupow et al. (2024) in the journal “Management Information Systems Quarterly” (VHB-Jourqual4: A+ or A across several disciplines, Impact Factor (2023): 7.0), provides a fundamental framework for further research on algorithm aversion from an information systems perspective. Publication in an information systems-based journal would therefore require another round of substantial rewrites to acknowledge Jussupow et al. (2024) and to differentiate the paper form their findings.

The current target journal for this research project is the “Journal of Business Economics” (VHB-Jourqual 4: B across several disciplines, Impact Factor (2024): 3.03), which is well-known for its interest in conceptual and theoretical studies. Furthermore, the final revision of the paper will be done by Aaron A. Engelbertz, who is currently also an PhD candidate at the Chair of Management Accounting and Control. Therefore, he will be included as a co-author in future versions of the manuscript.

3.2 Abstract

This study proposes a comprehensive theoretical framework for "algorithm aversion “, which is a tendency to favor a human judgment over a superior but imperfect algorithmic one. The paper addresses the conceptual challenges of this rapidly growing literature, including a lack of a clear definition and a framework to organize its many identified factors. Drawing on the Theory of Planned Behavior (TPB) of Ajzen (1991), the framework models the decision-maker's choice between human and algorithmic judgment. It identifies two main, non-mutually exclusive causes for algorithm aversion: one stemming from a human-favoring ratio of attitudes and the other from a human-favoring ratio of Perceived Behavioral Control (PBC). This theoretical approach provides a new way to understand algorithm aversion, clarifying its novelty compared to older research on technology dominance and serving as a basis for interpreting prior findings and guiding future research.

3.3 Introduction

This conceptual study is motivated by the rapidly growing and interdisciplinary stream of literature on “algorithm aversion”, which is a seemingly universal tendency of decision-makers to discount the advice of superior but imperfect intelligent decision aids

(Burton et al. 2020). Since the influential study of Dietvorst et al. (2015), the topic has inspired a substantial number of articles, ranging from practitioner press (e.g., Frick 2015) to well over one hundred peer reviewed papers (e.g., Mahmud et al. 2022). As decision support is an important function of accounting information systems (AIS) (Mauldin & Ruchala 1999), it is of no wonder that the topic of algorithm aversion draws more and more attention in accounting research too (e.g., Commerford et al. 2024; Downen et al. 2024; Fehrenbacher et al. 2023; Commerford et al. 2022; Chen et al. 2022a; Jung & Seiter 2021). However, some scholars have noted similarities with older AIS research and have dismissed the topic as “popular culture”, “business press folklore” and “old wine in new bottles” (cf. Sutton et al. 2023, pp. 7–8).

Despite the recent attention, the body of research on algorithm aversion is still developing and the underlying causes of this phenomenon are not understood very well (Jussupow et al. 2020). In particular, the literature is hampered by the lack of a clear definition, the rapid growth of empirically identified factors and the absence of a comprehensive theoretical framework to organize these interrelated factors. (Mahmud et al. 2022). These characteristics make it difficult to evaluate the general importance of algorithm aversion, to assess the transferability of the effect into different contexts and to derive actionable mitigation strategies for accounting and auditing practice. The purpose of this study is to resolve the described problems by developing a comprehensive theoretical framework for algorithm aversion for business context. Drawing on the Theory of Planned Behavior (TPB) of Ajzen (1991), the present study proposes a framework that models the choice of decision-maker between human and algorithmic judgment. From this point of view, algorithm aversion is defined as a relatively higher behavioral intention to rely on human judgment than to rely on algorithmic judgment. The proposed framework distinguishes between two fundamental causes of algorithm aversion: algorithm aversion from attitudes and algorithm aversion from Perceived Behavioral Control (PBC). On the one hand, algorithm aversion can occur because of a human-favoring ratio of attitudes. On the other hand, algorithm aversion can occur because of a human-favoring ratio of PBC. These causes are not mutually exclusive and can reinforce or negate each other. A comprehensive literature identifies beliefs and experiences that contribute to these differences in PBC and attitudes and highlights avenues for further research.

First, this study contributes to the general literature on algorithm aversion. While prior studies identify several important antecedents of algorithm aversion, such as a desire to stay in control (Commerford et al. 2024) or doubts about source credibility (Chen et al. 2022), the present study is the first to provide a theoretical framework that integrates these antecedents and explains their effects. In particular, the present study shows that algorithm aversion can be explained by two distinct yet related causes that result from conscious considerations. This allocation into two causes is a new concept that exceeds the conceptual depth of inductive frameworks structuring prior research. (e.g., Mahmud et al. 2022, Jussopow et al. 2020). Moreover, positing conscious behavior contrasts with research that treats algorithm aversion as an unconscious cognitive bias. (e.g., Commerford 2022, Dietvorst et al. 2018). The proposed framework can assist future studies that struggle to explain their empirical results from a theoretical perspective.

Second, the present study clarifies the novelty of the algorithm aversion literature in comparison to older research in accounting and auditing that culminated in the reliance model of the Theory of Technology Dominance (TTD) (Sutton et al. 2023, Arnold & Sutton 1998). It shows that algorithm aversion is a relative effect that describes a decision-maker's intention to rely on human rather than algorithmic judgment. Hereby, human judgment can either represent one's personal judgment or the judgment a human specialist. Moreover, the form of occurrence of algorithm aversion depends on the respective task, where exclusive choices between human and algorithmic judgment lead to a choice of human judgment and non-exclusive choices lead to an overweighting of human judgment compared to algorithmic judgment. In contrast, the TTD reliance model explains only the absolute degree of reliance on a decision aid for a given task. (e.g., Arnold and Sutton 1998). Algorithm aversion effects that nevertheless show a high degree of absolute reliance are out of scope for the model (Sutton et al. 2023). Consider, for example, a decision-maker with a high degree of reliance on an algorithmic judgment. Such a decision-maker is still algorithm-averse if they would rely even more on comparable advice from a human specialist.

The paper consists of three main sections. The first section briefly discusses prior research on algorithm aversion in accounting and auditing and motivates the need for a comprehensive theoretical framework. The second section presents the proposed framework and derives the causes of algorithm aversion. The third section demonstrates the framework's applicability and identifies avenues for further research.

3.4 Motivation for the proposed framework

This section outlines the motivation for proposing a theoretical framework for algorithm aversion. It first explains and characterizes the recent interest in algorithm aversion in accounting-related research. It then describes the challenges of comprehending this interdisciplinary literature. Finally, it discusses the difficulties of explaining the observed empirical findings using established theories of individual IT use.

3.4.1 Background

Accounting and auditing tasks often require the prediction of the future. Consider for example, the evaluation of management's complex estimates in audits (e.g., Bratten et al. 2013) or the allocation of budgets with the help of sales, cost and production predictions (e.g., Horngren et al. 2018). These advances enable the implementation of advanced decision aids, collectively referred to as data analytics, to support accountants and auditors with these complex tasks (e.g., Appelbaum et al. 2017). These decision aids exceed human capabilities in forecasting tasks and therefore often provide valuable decision support (e.g., Brynjolfsson et al. 2011). Good examples are algorithm-based earnings forecasts, where predictive analytics systematically outperform the prediction accuracy of human specialists (Chen et al. 2022b). Consequently, many firms are currently investing into data analytics systems to assist accountants and auditors in these forecasting tasks (Austin et al. 2021).

However, the actual success of data analytics in corporate practice remains uncertain (Rikhardsson & Yigitbasioglu 2018). Some studies observe significant performance increases (Brynjolfsson et al. 2011) and a growing importance of accounting practices after the introduction of data analytics (Youssef & Mahama 2021), while others report reluctance to incorporate new data (Spraakman et al. 2021) and general digital anxiety (Firk et al. 2024). Given the benefits of data analytics for accounting and auditing, their ambiguous success appears somewhat counterintuitive. Inspired from psychology, recent studies in accounting and auditing therefore propose an empirical observation called algorithm aversion as a potential explanation (e.g., Commerford et al. 2024; Downen et al. 2024; Fehrenbacher et al. 2023; Commerford et al. 2022; Chen et al. 2022a; Jung & Seiter 2021). Broadly, algorithm aversion describes the tendency of decision-makers to discount the advice of superior but imperfect intelligent decision aids for forecasting tasks, but not the advice of comparable human advisors (Jussupow et al. 2020). Algorithm aversion effects are observed in management science (e.g., Kawaguchi 2021), marketing (e.g., Castelo et al. 2019), information systems (e.g., Berger et al. 2021), medicine (e.g., Longoni et al. 2019) and psychology (e.g., Dietvorst et al. 2015) and in the aforementioned accounting and auditing context (e.g., (Commerford et al. 2024)) among others.

The recent focus on algorithm aversion has sparked a debate in accounting research on the novelty of a respective studies. On the one hand, the novelty and relevance of algorithm aversion for accounting and auditing has been questioned by Sutton et al. (2023), who compare the findings of the research stream to older research in auditing and describe several interconnections (e.g., Ashton 1990). Respective interconnections can be seen, for example, through the description of the observed behaviors. While Dietvorst et al. (2015) describe algorithm aversion as an "[aversion] to algorithmic forecasters after seeing them perform, even when they see them outperform a human forecaster." (cf. Dietvorst et al. 2015. p. 114), older accounting research observes a "non-reliance on [an] aid even though the aid has been shown to produce, on average, better outcomes than unaided judgment." (cf. Mauldin and Ruchala 1999, p. 326). On the other hand, other authors argue such as Commerford et al. (2022) argue that modern accounting information systems differ significantly from the information systems that older accounting-related research has examined and that these older studies therefore do not provide clear implications. For instance, considering the evolution of forecasting algorithms from simple autoregressive models to neuronal networks (e.g., Khashei & Bijari 2011), this rationale is justifiable too.

3.4.2 Conceptual challenges of algorithm aversion

It is somewhat challenging to fully comprehend and assess the recent literature on algorithm aversion. On the one hand, it remains ambiguous of whether all of the referenced studies even examine the same behavior or whether algorithm aversion must be treated as a collective term for several behaviors. For example, Jung & Seiter (2021) among others draw on the research paradigm of Dietvorst et al. (2015) and treat algorithm aversion as a reluctance to rely on algorithms altogether. In particular, these studies create experimental conditions in which participants are required to exclusively decide between handing over a forecasting task to an algorithm or solving this forecasting task by themselves (e.g., see Dietvorst et al. 2015, who provide full access to all their study materials). Other studies such as Commerford et al. (2022) draw on the judge-advisor-system research paradigm (e.g., Bonaccio & Dalal 2006) and treat algorithm aversion as a stronger advice discounting of advice from algorithmic advisors compared to advice from human advisors. Respective studies artificially create an experimental context in which participants first are required to submit a forecast for a forecasting task. After, they are offered advice from either an algorithmic advisor or a human advisor and are allowed to modify their original forecast. These then compare the relative average forecast adjustment depending on the advisor (e.g., see Commerford et al. 2022). Moreover, there are some studies operationalizing algorithm aversion via actual use of a decision aid (e.g., Dietvorst et al. 2018), while other studies examine the perceived trust or confidence without controlling for actual use of a tool (e.g., Castelo et al. 2019).

On the other hand, the interdisciplinary research on algorithm aversion draws on a large number of theories and offers a similar large number of factors leading to algorithm aversion without ever theorizing about their interdependencies. For example, while Dietvorst et al. (2015)) ask their participants about their confidence in a forecasting models forecast, Castelo et al. (2019) ask their participants about their trust in a respective algorithm. Although prior research on trust distinguishes the terms of “trust” and “confidence”, confidence would require not evaluating potential alternatives for a behavior (e.g., Mayer et al. 1995). Consequently, as Dietvorst et al. (2015) create the aforementioned exclusive choice between a forecasting algorithm and oneself, the use of the term “confidence” is misleading. In total, prior research identifies over 50 factors related to algorithm aversion that broadly fall into the categories of high-level factors, individual factors, task-related factors and algorithm-related factors (Mahmud et al. 2022). This is problematic, because future research cannot control for all of these antecedents or even develop mitigation strategies that address all of these factors.[14]

Moreover, note that there have been several attempts in the recent psychology and information systems literature to provide analytical and inductive frameworks for algorithm aversion. However, these frameworks often only focus on specific research paradigms or fail to explain why their identified factors influence the occurrence of algorithm aversion. On the one hand, for instance studies such as Sinclair-Desgagné (2024), and Kumar et al. (2021) provide comprehensive analytical models for algorithm aversion based on

[14] While the studies of Mahmud et al. (2022) and Jussupow et al. (2020) provide structured illustrations of prior research on algorithm aversion that could be interpreted as inductive theory-building, both studies call for the proposition of a more fundamental theoretical framework, which is grounded in established theories.

perceived utility. Sinclair-Desgagné (2024) hereby draws on the example of a personal navigation problem, which addresses the initial example of Dietvorst et al. (2015) but hinders the transferability into a corporate context. Kumar et al. (2021) state that their developed framework is not able to explain empirical findings following the judge-advisor-system research paradigm. On the other hand, studies such as the literature reviews of Mahmud et al. (2022) and Jussupow et al. (2020) provide structured aggregations of prior research on algorithm aversion. While one could see these aggregations as inductive theory-building approaches, both studies acknowledge that their findings do not fully explain why algorithm aversion occurs and call for the proposition of a more fundamental theoretical framework. The proposed framework seeks to build on the described limitations.

3.4.3 Relation to established theories

Adding to the described conceptual ambiguity around algorithm aversion, the empirical findings of the literature stream also exceed many well-known theories and models for the use of IT-systems and technology as a whole. The following section provides three examples that these established theories cannot predict. The technology acceptance literature, for example, proposes the individual acceptance of IT-systems to be a key determinant for later use. In this literature stream, the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al. 2016, Venkatesh et al. 2012; Venkatesh et al. 2003) is seen as a culmination of prior models and as the most general one. The UTAUT model focusses on the performance expectancy of users as a major determinant of IT-system use and predicts an expected performance increase from an IT system to be a strong predictor for later use. However, one central finding of the algorithm aversion literature is that the users of the examined algorithms are consciously aware that they can improve their performance from relying on these tools and still decide against reliance (Burton et al. 2020). Dietvorst et al. (2015), for example, ask their participants about their expected performance regarding the offered algorithm. They find that even after experiencing the error proneness of the offered algorithm, participants ascribe a significantly higher likelihood to the algorithm than to themselves to achieve a perfect prediction.

In contrast the aforementioned literature on technology acceptance, the literature on technology trust proposes trust in IT-systems or technology as a key determinant for subsequent use (e.g., Glikson & Woolley 2020, Hoff & Bashir 2015, Lee & See 2004). Hereby, the development of trust is assumed to depend on one's rating about an IT-systems trustworthiness, which in turn is a function of the perceived performance of an IT-system, the perceived functionality of the underlying algorithm and the actual purpose why one or more developers have created the tool (e.g., Lee & See 2004). In line with this reasoning, the literature on algorithm aversion sometimes observes differences in the perceived trustworthiness of forecasting algorithms and human advisors (e.g., Candrian & Scherer 2024, Castelo et al. 2019). However, other studies using different measures for trust do not observe a respective difference (e.g., Daschner & Obermaier 2022). Furthermore, even in case of an observable difference in trustworthiness between human and algorithmic advisors, this difference does not seem to affect the actual use of an algorithm (e.g., Himmelstein & Budescu 2023, Dietvorst et al. 2015).

Lastly, the aforementioned TTD literature explains the use of intelligent decision aids by auditors and similar professions with the help of task expertise, task complexity, aid familiarity, and cognitive fit or congruence (Sutton et al. 2023, Arnold & Sutton 1998). As stated by Sutton et al. 2023, the literature on algorithm aversion empirically supports the predicted lower degree of use of algorithms by more experienced decision-makers (e.g., Logg et al. 2019) and the predicted higher degree of use after becoming familiar with a well-performing decision aid (e.g., Filiz et al. 2021). However, the TTD interprets intelligent decision aids as "replacement colleagues" and therefore reliance behavior relative to a human expert's opinion is out of scope for the model (cf. Sutton et al. 2023, p. 7). Recall that a respective research paradigm is applied by several prior studies on algorithm aversion (e.g., Commerford et al. 2024, Commerford et al. 2022). Therefore, the transferability of the TTD model to some of the described algorithm aversion settings seems questionable.

3.5 Explanation of the proposed framework

This chapter explains and discusses the proposed theoretical framework for algorithm aversion. First, it provides an overview of the framework and provides an in-depth example for the proposed causal effects. Afterwards, it characterizes the assumed choice problem and explains the assumed causal relationships with the help of the underlying TPB. Moreover, it explains the two causes of algorithm aversion that follow from the proposed theoretical framework. Lastly, this section differentiates the proposed framework form established theories.

3.5.1 Overview

Figure 3-1 illustrates the proposed theoretical framework for algorithm aversion. In particular, this study proposes that on fundamental level, algorithm aversion is always an individual's relative intention to rely on human judgment (H) instead algorithmic judgment (A) in a forecasting task. It hereby assumes that there is always a choice between respective forms of judgment and consequently that none of the two decision options is compulsory. In turn, this intention in favor of human judgment is determined by four factors: One's attitudes towards the two available forms of judgment and one's Perceived Behavioral Control (PBC) regarding the two available forms of judgment. While a positive attitude towards human judgment increases the intention to rely on human judgment, a positive attitude towards algorithmic judgment decreases this intention. Doubts about one's PBC to rely on a form of judgment negatively moderate the described effects. There are two main causes of algorithm aversion, namely algorithm aversion due to an unfavorable ratio of attitudes and algorithm aversion due to an unfavorable ratio of PBC. Note that these causes are not independent from each other and favorable ratios in one pair of factors can offset unfavorable ratios in the other pair of factors.

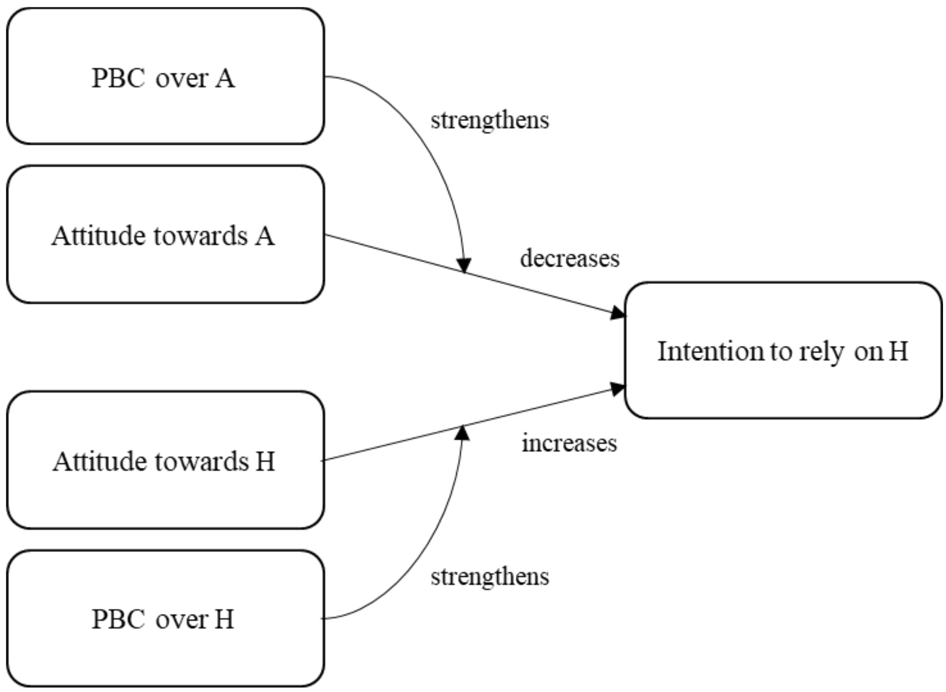

Figure 3-1: The proposed theoretical framework for algorithm aversion.

For further explanation, consider the example of a management accountant who is offered a new data analytics tool to assist with the forecasting of future sales for sales and operations planning (e.g., Brüggen et al. 2021). Further, assume that the management accountant enjoys to prepare respective forecasts with the help personal expert interviews and is not familiar with the performance of the offered data analytics tool. Lastly, also assume that the management accountant has a lot experience with the qualitative interview method but only basic knowledge about the principles and use of data analytics. For this context, the framework predicts that the management accountant will show an intention to rely on the established forecasting method, which is human judgment, and consequently show algorithm aversion. This can be explained with the help of two causes. On the one hand, enjoying the established forecasting method implies a positive attitude towards human judgment and the lack of experience with the offered data analytics tool implies a neutral attitude towards algorithmic judgment. Taken together, there is likely a ratio of attitudes favoring human judgment. On the other hand, possessing knowledge about the established forecasting method but not about the alternative one, implies that the management accountant likely only perceives to possess the necessary skills to rely on human judgment. In turn, this context also implies a PBC ratio favoring human judgment.

To gain a deeper understanding of the interaction of the two causes of algorithm aversion, now assume that there is a PBC ratio favoring algorithmic judgment. This difference in perception could for example arise, because the last forecast raised significant doubts of whether the interview partner possess any knowledge of future sales. Clearly, a respective context creates a dilemma for the management accountant between the type of judgment he or she enjoys and the type of judgment that likely yields reasonable and accurate forecasts. In case these accuracy concerns exceed a certain threshold, the management accountant will lose his or her preference for human judgment and favor algorithmic judgment. To gain a deeper understanding of the moderation effect of PBC, consider the detailed effects of PBC in the described example. It is highly unlikely that the management accountant will choose on type of judgment only because he or she is capable of doing so, which translates to a high PBC. Instead, a respective choice requires at least a somewhat positive attitude towards the type of judgment. For instance, because of expected performance gains.

3.5.2 Characterization of the choice problem

The fundamental idea underlying the proposed framework is to interpret algorithm aversion as a behavioral intention and therefore a state of mind instead of an actual behavior. This assumption allows to integrate research paradigms in which algorithm aversion is treated as a reluctance to rely on algorithms altogether as well as research paradigms in which algorithm aversion is treated as a stronger advice discounting of advice from algorithmic advisors compared to advice from human advisors. On the one hand, this behavioral intention favoring human judgment leads to a choice for human judgment in exclusive choice situations. On the other hand, this intention leads to a relative overweighting of human judgment compared to algorithmic judgment in non-exclusive choice situations. The proposed framework hereby implicitly assumes that there is always a choice between human and algorithmic judgment. This implicit assumption is justified by a broad definition of human judgment, where this factor can either include one’s personal judgment or the judgment of another human expert. While this complicates the interpretation of the type of judgment, it also allows to integrate studies examining the use of algorithms relative to one’s personal judgment (e.g., Dietvorst & Bharti 2020, Dietvorst et al. 2018, Dietvorst et al. 2015) and studies examining the use of algorithms relative to the judgment of another human expert (e.g., Commerford et al. 2024, Commerford et al. 2022, Berger et al. 2021).

Applying the concept of a behavioral intention further implies that algorithm aversion is a conscious decision and not an unconscious one. At first glance, a respective assumption appears to be refuted by prior research on algorithm aversion. On the one hand, it contrasts the literature review of Mahmud et al. (2022), who define algorithm aversion as a discounting of algorithmic advice “either consciously or unconsciously” (cf. Mahmud et al. 2022, p. 2). On the other hand, it contrasts several studies defining algorithm aversion as a cognitive bias that requires no additional explanation (e.g., Commerford et al. 2022, Dietvorst et al. 2018). However, the qualitative and non-experimental studies on algorithm aversion do not support the interpretation of algorithm aversion as an unconscious cognitive bias (e.g., Commerford et al. 2024, Neumann et al. 2023a, Mahmud et al. 2022, De-Arteaga et al. 2020). For instance, Commerford et al. (2024) support their experimental results with in-depth interviews with 17 auditors regarding their reasoning when using a decision aid incorporating AI. These interviews indicate that auditors consciously evaluate respective decision aids and hereby acknowledge their likely superiority in achieved accuracy compared to a human specialist. Furthermore, these interviews also indicate that the interviewed auditors are skeptical whether these decision aids are able to understand the reasoning of a client and therefore require additional information on how a decision aid generates advice.

3.5.3 Underlying theoretical mechanisms

The proposed theoretical framework follows the TPB of Ajzen (1991) and its definition of a behavioral intention. According to this theory, a behavioral intention is the personal willingness or readiness to perform a specific behavior, where a positive intention makes the performance of this behavior more likely (Ajzen 1991). In turn, behavioral intentions are a function of one's attitude and PBC regarding the behavior of interest. An attitude can be broadly defined as an individuals' positive or negative evaluation of performing the behavior. PBC aggregates one's expectations about one's ability to perform a behavior (Ajzen 1991). Attitudes influence a behavioral intention directly and PBC moderates the respective effects (e.g., La Barbera & Ajzen 2020). It is important to note that this moderation effect is assumed to be decreasing, where doubts about ones PBC over a behavior weaken the effect of a positive attitude on a behavioral intention (La Barbera & Ajzen 2020, Yzer & van den Putte 2014).[15]

While the TPB was originally developed to predict the decision to perform a specific behavior, it can also be applied to choices between behavioral options (Ajzen 2020). The given context is a respective choice problem, where one must decide between relying on two different behavioral options, namely relying on human or algorithmic judgment. For respective choice problems, a behavioral intention serves as an indication of preferences across the available behavioral options, where a high intention to show one behavioral option equals a low intention to show the other behavioral options and vice versa (Ajzen & Fishbein 1969). Furthermore, the other behavioral options are also assumed to influence whether one will show a specific behavior. In particular, one's attitude towards the behavior of interest are evaluated relative to the one's attitude and perceived social norm

[15] Note that the complete TPB also includes social norms as a predictor of behavior whose effect is similar to the described mechanisms of attitudes (e.g., Ajzen 1991). In general, social norms can be defined as "perceived social pressure to perform (or not to perform) a given behavior" (Ajzen 1985). However, the moderation effect of PBC on the effect of social norms on a behavioral intention is unclear and still the subject of current research (e.g., La Barbera & Ajzen 2020). Therefore, the present study disregards any potential effects of social norms on behavioral intentions.

towards the alternative behavioral options (Gardner & Abraham 2010; Ajzen and Fishbein 1969).

While the moderation effects of PBC in respective choice situations between available behavioral option have not yet been clarified in a peer reviewed publication, a thought experiment provides a strong indication regarding the underlying causal mechanisms. In total, it follows that the difference in attitudes is likely evaluated after the moderation effect of PBC occurs. To gain a better understanding of this hypothesized relationship, first consider the sole evaluation of human judgment. Assume that one forms a positive attitude towards algorithmic judgment, because he or she expects to enjoy to working with an offered decision aid. Moreover, assume that one is very convinced that he or she is capable of operating the offered decision aid and therefore perceives a lot of behavioral control over the forecasting approach. Drawing on the propositions of the TPB, one would likely rely on the generated forecast of a forecast for a given forecasting task.

Afterwards, consider a choice between human and algorithmic judgment, where one shows a positive attitude towards relying on human and algorithmic judgment and perceives an equally high PBC regarding both behavioral options. For instance, assume that one forms a positive attitude towards human judgment, because he or she expects to enjoy to manually deduce a forecast for a given forecasting task. Furthermore, assume that one is also very convinced to be capable of manual forecasting. As both behavioral options appear to be feasible, there is no negative moderation of PBC on the respective attitudes and the positive attitudes therefore likely cancel themselves out. In turn, the actual behavioral intention to rely on the decision aid decreases. Now consider how these relationships change in case one does not perceive behavioral control over relying on human judgment, for example because he or she has not the necessary skills for manual forecasting. Evidently, relying on human judgment is not a feasible behavioral option despite one's positive attitude towards it. Subsequently, when comparing human and algorithmic judgment, the positive attitude towards relying on human judgment is likely strongly discounted because the behavior is not feasible. Consequently, the actual difference in attitudes remains significantly larger than in the previous case and the actual behavioral intention to rely on the decision aid remains relatively high.

3.5.4 Propositions on algorithm aversion

Recall that following the described rationale of the TPB, algorithm aversion occurs when one forms a higher behavioral intention to rely on human judgment than on algorithmic judgment. However, due to the two moderation effects of PBC, the factor combinations leading to algorithm aversion are not always directly apparent. For instance, consider a case where one has a less positive attitude towards algorithmic judgment than towards human judgment. Additionally assume that one perceives more behavioral control over relying on algorithmic judgment than over relying on human judgment. According to the TPB, the actual occurrence of algorithm aversion depends on the effect sizes of attitudes and PBC.

First, note the assumed difference in attitudes alone clearly favors relying on human judgment. However, whether a respective behavioral intention actually materializes, completely depends on the moderation effects of PBC: While the low PBC regarding human judgment decreases the impact of the positive attitude towards this behavior on the actual behavioral intention, the high PBC regarding algorithmic judgment increases the impact of the less positive attitude towards this alternative behavior. Depending on the effect sizes of the moderation effects, the PBC difference can either decrease, cancel out or reverse the effect of the attitude difference on behavioral intention. The resulting continuum of factor combinations leading to algorithm aversion can be visualized with the help of an indifference curve, which separates factor combinations leading to algorithm aversion from factor combinations not leading to algorithm aversion. Figure 3-2 depicts the indifference curves for two combinations of PBC across the available behavioral options.

This study distinguishes between two main causes of algorithm aversion: algorithm aversion from a human-favoring ratio of attitudes and algorithm aversion from a human-favoring ratio of PBC. Following the proposed rationale, algorithm aversion only occurs when there is a large difference in attitudes or PBC favoring human judgment, which is not nullified or reversed by a difference in the other factor favoring algorithmic judgment. For instance, consider a case where a management accountant has only a marginally lower attitude towards human judgment than towards algorithmic judgment but perceives a lot more control over human judgment than over algorithmic judgment. While the attitude difference could predict the reliance on algorithmic judgment, this effect is likely still dominated from the human-favoring difference in PBC and it is very likely that the management accountant will actually show algorithm aversion. Drawing on the causal mechanisms of the proposed framework (Figure 3-1), one can analytically describe the minimum ratios of factors required to trigger algorithm aversion. On the one hand, algorithm aversion from a human-favoring ratio of attitudes occurs for factor combinations fulfilling the equation:

$$\frac{Attitude\ (H)}{Attitude\ (A)} > \frac{PBC\ (A)}{PBC\ (H)}. \tag{1}$$

On the other hand, algorithm aversion from a human-favoring ratio of PBC occurs for factor combinations fulfilling the equation:

$$\frac{PBC\ (H)}{PBC\ (A)} > \frac{Attitude\ (A)}{Attitude\ (H)} \tag{2}$$

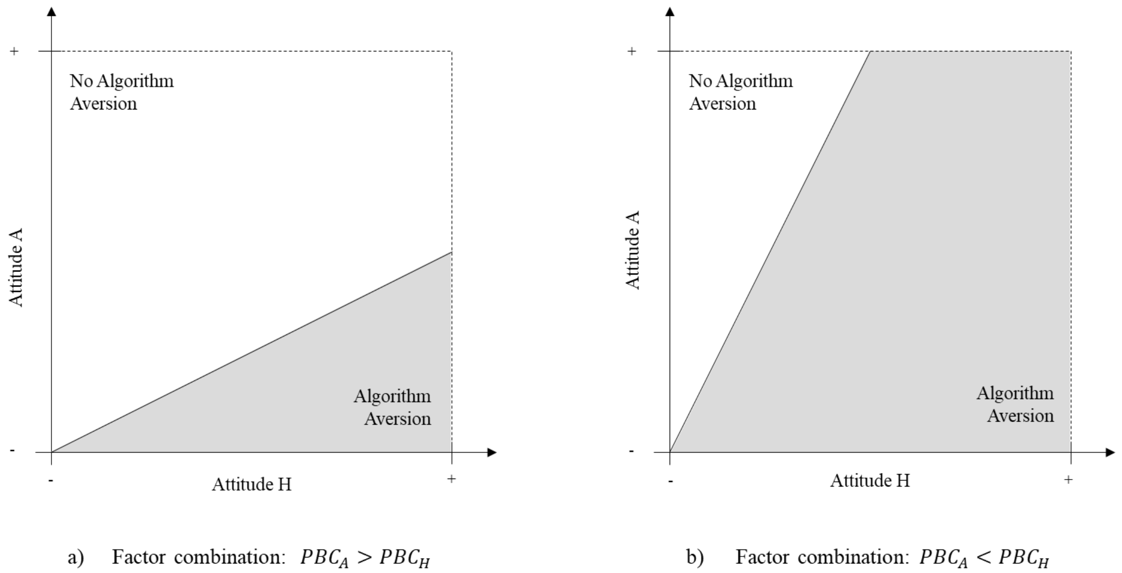

Figure 3-2: The continuum of factor combinations leading to algorithm aversion.

3.5.5 Differentiation of the proposed framework

The question arises on how the proposed framework relates to established theories on the use of IT systems and on the prediction of behavior in general. The following section discusses three important characteristics of the framework and its relations to the TTD, the UTAUT and Prospect Theory. First, the proposed framework assumes a behavioral choice between relying on human judgment and relying on algorithmic judgment. As already discussed in section 3.4.3, this context is out of scope of the TTD in case human judgment includes the consultation of a human specialist (Sutton et al. 2023). The UTAUT partly includes a respective choice through the definition of the attitude-related "performance expectancy" and "effort expectancy", which are measured relative to the status quo of solving a task (Venkatesh et al. 2003). However, while PBC is included in the model through "facilitating conditions" it is measured only for the IT-system to be adopted and in absolute terms (Venkatesh et al. 2003). Lastly, while choice situations are a core concept of Prospect Theory, considerations related to PBC are not part of the theory (Kahneman & Tversky 1979).

Second, the proposed theoretical framework focusses on the occurrence of algorithm aversion, which is defined as a relative effect and the absolute degree of reliance on an IT-system is of lesser importance.[16] Conversely, the TTD defines reliance as "the degree to which the user of a decision aid applies and incorporates the recommendations of the aid during judgment formulation" (cf. Arnold & Sutton 1998, p. 180) and therefore as an absolute factor. Similarly, the UTAUT also focusses on the prediction of the absolute degree of use of an IT-system to be adopted (Venkatesh et al. 2003). Prospect Theory focusses on perceived utility, which is a very fundamental concept that could be translated to the given context (Kahneman & Tversky 1979). Third, the proposed framework is grounded in the established and well-known TPB (Ajzen 1991), that has been cited over 150,000 times according to Elsevier.[17] In contrast, the TTD has been derived inductively from the accounting and auditing-related studies on IT-use at the time of publication (Arnold & Sutton 1998). The UTAUT and the previous TAM model draw on the Theory of Reasoned Action (TRA), which is a predecessor of the TPB (e.g., Fishbein & Ajzen 2010, Fishbein 1963). However, while the TRA excluded considerations regarding PBC, the TAM model has introduced a similar factor of "ease of use" and the literature has therefore departed from its theoretical origin (Ajzen 2020, Davis 1985). The aforementioned utility construct of Prospect Theory and the attitude concept of TPB can be seen as identical (e.g., Fishbein 1963).

3.6 Application of the proposed framework

While the previous chapter describes the likely causal mechanisms of algorithm aversion, the included factors remain rather abstract. In general, the TPB expects seven to ten salient beliefs each regarding the attitude and PBC towards a specific behavior. (Fishbein & Ajzen 2010). The following chapter therefore applies the proposed theoretical framework to prior literature on algorithm aversion. On the one hand, this application follows the

[16] An exception are exclusive choice situations (e.g., Dietvorst et al. 2015), where relative differences in reliance on human or algorithmic judgment are conceptually not possible.
[17] See for further information. https://www.sciencedirect.com/science/article/abs/pii/074959789190020T.

aim of further explaining the included factors and their interpretations. On the other hand, this application also illustrates gaps in prior research that provide avenues for further research and therefore demonstrates the usefulness of the proposed framework.

3.6.1 Additional literature search

The present study systematically evaluates prior empirical literature on algorithm aversion for these aforementioned salient beliefs that characterize attitudes and PBC. Respective literature was identified with the help of a structured literature search on the online database Scopus from the publisher Elsevier. A search query for the term “algorithm aversion” in either title, abstract or keywords revealed 123 peer-reviewed publications until the cut-off month of September 2024. No further restrictions were applied, for example there was no focus on specific journals or research areas. Manual prescreening of the abstracts and introductions of the identified papers for an organizational context as well as for empirical research methods allowed cutting the number of identified and relevant papers to 68. It followed an analysis of these 68 papers and a characterization of the studies according to the proposed causes of algorithm aversion. The following sections summarize and describe the results of the literature search.[18]

3.6.2 Empirical evidence for algorithm aversion from attitudes

A prominent finding of the algorithm aversion literature is the different evaluation of errors from algorithmic and human judgment, where a substantial error can trigger aversion towards algorithmic advisors but not towards human advisors or oneself (e.g., Dietvorst et al. 2018, Prahl & van Swol 2017, Dietvorst et al. 2015). Previous research suggests that this effect occurs because people expect more severe negative consequences from errors caused by an algorithm than by a human (e.g., Filiz et al. 2023, Schneider & Freisinger 2022) or think more about potential negative consequences of using algorithmic advice (Chang & Wang 2023), which leads to an unfavorable ratio of attitudes regarding algorithmic judgment. One explanation for this different error perception is the black-box nature of advanced decision aids, which makes it more difficult to blame algorithmic advisors for an error than to blame human advisors. It is therefore more likely for this uncomfortable responsibility for the error to stay with the decision-maker (Aschauer et al. 2024). Another explanation for the difference in error perception is a diminishing

[18] Note that a full list of the identified papers in included in the appendix.

sensitivity to error gravity, which leads to a preference for the seemingly more volatile advice of human advisors (Dietvorst & Bharti 2020).

Furthermore, decision-makers seem to have troubles in assessing the performance benefits from using a provided decision aid (Tse et al. 2024) and, at first, perceive it as a job threat (Turel & Kalhan 2023). In line with this reasoning, this negative evaluation of a decision aid fades when a decision-maker gains positive performance experience with this aid, for instance through a constant outperforming of a decision-makers judgment in regards to forecast accuracy (e.g., Freisinger et al. 2024, Chacon et al. 2022, Saragih & Morrison 2021). However, note that this ambiguity about performance benefits is context-specific and does, for example, not occur for fund management decisions (Germann & Merkle 2022). Moreover, algorithm aversion is linked to negative emotions and emotional responses in regards to an algorithmic decision aid (Downen et al. 2024, Renier et al. 2021). While these emotions could be interpreted as another form of evaluation of the behavioral option, they likely lay outside of the TPB.

Relying on human judgment might provide some unique benefits, which do not materialize when one chooses to rely on algorithmic judgment. Prior research connects algorithm aversion to self-humanization problem and argues that human advisors but not algorithmic advisors recognize a decision-makers uniqueness, which these decision-makers appreciate (Liu et al. 2023, Heßler et al. 2022). In line with this rationale, there is empirical evidence that algorithm aversion vanishes when a decision aid is presented as a human being (Ganbold et al. 2022, Ochmann et al. 2021), in particular when presented as a woman (Schulte Steinberg & Hohenberger 2023, Borau et al. 2021). For decisions between using the provided advice of a decision aid and relying on one’s own judgment, the latter behavior might trigger a greater feeling of importance, especially for restrictive decision aids (Adam et al. 2024, Neumann et al. 2022a). Consequently, there is a lot of prior research that is able to mitigate algorithm aversion by increasing human decision-makers in the solution of a task (e.g., Fink et al. 2024, Kawaguchi 2021, Snow 2021, Dietvorst et al. 2018).

While disregarded in the proposed framework, there is nevertheless some empirical evidence for algorithm aversion effects arising from social norms. First, there is evidence that an organizational culture of relying on human judgment and a culture of innovation resistance significantly increases the likelihood of algorithm aversion (Mahmud et al. 2023, Neumann et al. 2023a). Moreover, there is some evidence from medicine that physicians can gain reputation from not consulting an available decision aid (Pezzo et al. 2021) and that corporate brands are seen as more authentic when they do not use Generative AI for content creation (Brüns & Meißner 2024, Haupt et al. 2024). In addition, a sentiment analysis of news articles and social media posts on AI projects indicates cautious discussions, indicating a tendency towards algorithm aversion in the society (Oomen et al. 2024).

3.6.3 Empirical evidence for algorithm aversion from PBC

Several studies on algorithm aversion connect the occurrence of the effect to difference in PBC and in particular to concerns about the capabilities and credibility of a decision aid (Neumann et al. 2023a, Chen et al. 2022a, Commerford et al. 2022). First, decision-makers seem to require decision aids to be able to learn from errors and are averse towards aids that seemingly do not possess this capability (Reich et al. 2023, Berger et al. 2021, Dietvorst et al. 2015). Second, there are doubts whether algorithmic decision aids are able to apply fairness principles and ethics for moral decision problems that can lead to algorithm aversion (e.g., Fine et al. 2023, Mok et al. 2023, Jauernig et al. 2022). In addition, there are doubts on whether algorithmic decision aids are able to solve creative tasks, for instance writing texts (Proksch et al. 2024) and there is the expectation that these tools put less effort in creative tasks (Magni et al. 2024). Third, people are averse towards algorithmic advisors with a slow response rate, because they expect fast responses from an algorithmic decision but not from a human decision aid (Efendić et al. 2020). Fourth, the labeling of the decision aid influences the occurrence of algorithm aversion, where decision-makers associate less expertise with "statistical models" and "algorithms" than with "Artificial Intelligence and are more averse towards those (Candrian & Scherer 2024, Hou & Jung 2021, Keding & Meissner 2021).

Moreover, algorithm aversion also significantly depends on the PBC for the human behavioral option. There is empirical evidence for a minimum level of own task expertise or self-confidence that is required for algorithm aversion to occur (Horowitz & Kahn 2024, Logg et al. 2019). This effect is particularly dominant, when the human advisor is perceived as an expert (Lacroux & Martin-Lacroux 2022). In line with this reasoning, prior research finds no algorithm aversion when an alternative human advisor is framed as "another participant" or as "the average guess of 5,000 people", but a significant effect when the respective advisor is framed as an "experienced expert board" (Bogert et al.

2021, Hou & Jung 2021). In addition, algorithm aversion is less likely when an alternative human advisor is perceived as biased and detrimental to a decision-maker (Claudy et al. 2022, Bigman et al. 2021) or likely to be corrupt (Castelo 2024).

3.6.4 Avenues for further research

The proposed framework provides several avenues for further research on algorithm aversion of which the following section provides three examples. First, it provides an explanation for a surprising finding of Dietvorst et al. (2015), where the participants show algorithm aversion despite expecting an algorithm to perform better than themselves in an experimental task. Note that this finding could result from a human-favoring ratio of PBC that nullifies the effect of the algorithm-favoring ratio of attitude-related performance expectancies. One of the applied forecasting tasks in the paper provides a strong argument for the described hypothesis. In particular, around half of the reported studies involved students predicting the future performance of other students (Dietvorst et al. 2015). Evidently, this is a topic that the students should be familiar with and therefore likely possess the necessary skills to solve this task, which in turn leads to a high PBC. Conversely, the findings of Jung & Seiter (2021) show artificially decreasing this confidence in one's perceived capabilities via time-pressure also decreases the likelihood of algorithm aversion effects. Future research could validate the derived hypotheses with the help of a replication study of Dietvorst et al. (2015), where PBC is measured directly. Decades of research on the TPB provide several effective scales to measure PBC in an experimental study (e.g., Fishbein & Ajzen 2010).

Second, the proposed framework indicates that the likelihood for algorithm aversion not only depends on the level of expertise a decision-maker in a domain (e.g., Horowitz & Kahn 2024), but also on his or her general IT-skills. While domain experts should perceive more PBC over human judgment in their domain, IT-averse decision-makers should perceive a low PBC over algorithmic judgment and IT-savvy decision-makers should perceive the opposite. Consequently, the behavioral intention of IT-averse specialist to rely on human judgment is driven by their attitude towards human judgment and their PBC over algorithmic judgment and the respective intention of IT-savvy specialists is driven by their attitudes towards human and algorithmic judgment. Following the proposed rationale, IT-averse specialists should bear an even greater risk of the algorithm aversion than IT-savvy specialists. Future research can validate this hypothesis for example with the help of a comprehensive survey with specialists from corporate practice. There are established measurement scales for IT-skills, for instance the Internet Skills Scale of van Deursen et al. (2016).

Third, recall that there are empirical findings that connect algorithm aversion to social norms. Future research could therefore revise the proposed framework and include social norms. Respective research not only improves the understanding of algorithm aversion, but also contributes to ongoing research on the TPB. Note that recent literature on the effect of social norms on behavioral intentions proposes a converse mediation effect of PBC compared to its effect on attitudes, where a low PBC increases the effects of social norms on behavioral intentions and vice versa (e.g., La Barbera & Ajzen 2020). On the one hand, the respective study of La Barbera & Ajzen (2020) examines individual voting behavior and its transferability to the use of algorithms is ambiguous. On the other hand, this finding is in line with older research on the use of IT-systems, which only observes a significant effect of perceived norms during adoption phases of an IT system (Karahanna et al. 1999)., where the PBC of using the system is likely low. Figure 3-3 illustrates and aggregates the findings of the additional literature search.

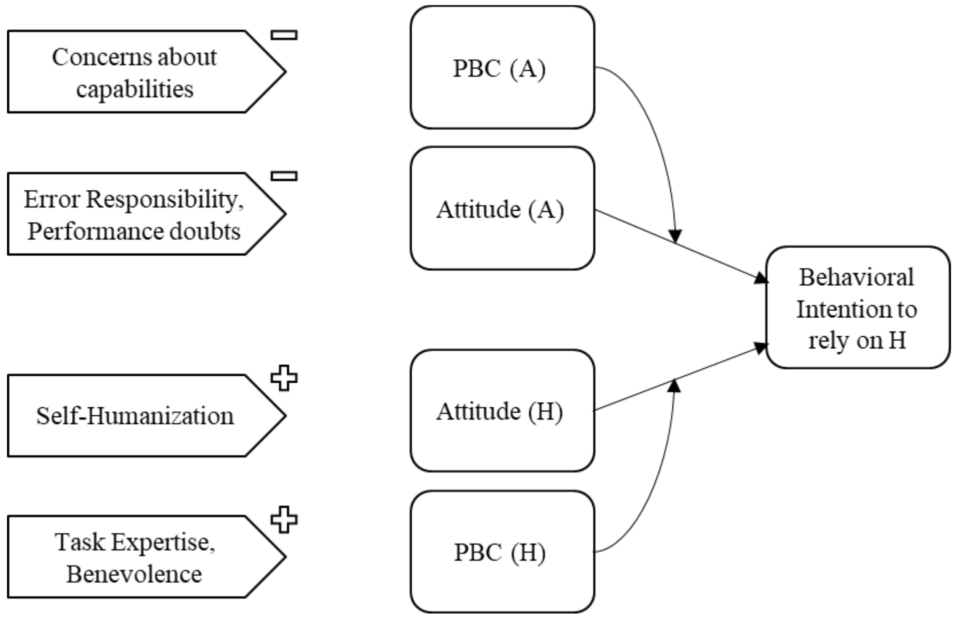

Figure 3-3: The findings of the additional literature search.

3.7 Conclusion

Algorithm aversion is a surprising empirical effect that has inspired over 100 peer-reviewed publications in approximately ten years. However, the understanding of the effect is still limited. Drawing on the TPB of Ajzen (1991), this study therefore proposes a comprehensive theoretical framework that explains the occurrence of algorithm aversion effects in these empirical studies. This framework characterizes two causes of algorithm aversion: algorithm aversion from attitudes and algorithm aversion from PBC. Furthermore, the proposed framework serves a basis to interpret prior research on algorithm aversion and provides avenues for further research. I hope the proposed framework inspires future research on algorithm aversion and helps to explain the observed results.

3.8 Appendix

Table 3-1: Full reference list of the literature search.

Authors	Title	Year	Source title
Adam M.; Diebel C.; Goutier M.; Benlian A.	Navigating autonomy and control in human-AI delegation: User responses to technology- versus user-invoked task allocation	2024	Decision Support Systems
Aschauer F.; Sohn M.; Hirsch B.	Managerial advice-taking—Sharing responsibility with (non)human advisors' trumps decision accuracy	2024	European Management Review
Brüns J.D.; Meißner M.	Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity	2024	Journal of Retailing and Consumer Services
Candrian C.; Scherer A.	How Terminology Affects Users' Responses to System Failures	2024	Human Factors
Castelo N.	Perceived corruption reduces algorithm aversion	2024	Journal of Consumer Psychology
Downen T.; Kim S.; Lee L.	Algorithm aversion, emotions, and investor reaction: Does disclosing the use of AI influence investment decisions?	2024	International Journal of Accounting Information Systems
Fink L.; Newman L.; Haran U.	Let me decide: Increasing user autonomy increases recommendation acceptance	2024	Computers in Human Behavior
Freisinger E.; Unfried M.; Schneider S.	The AI-augmented crowd: How human crowdvoters adopt AI (or not)	2024	Journal of Product Innovation Management
Haupt M.; Freidank J.; Haas A.	Consumer responses to human-AI collaboration at organizational frontlines: strategies to escape algorithm aversion in content creation	2024	Review of Managerial Science
Horowitz M.C.; Kahn L.	Bending the Automation Bias Curve: A Study of Human and AI-Based Decision Making in National Security Contexts	2024	International Studies Quarterly

Magni F.; Park J.; Chao M.M.	Humans as Creativity Gatekeepers: Are We Biased Against AI Creativity?	2024	Journal of Business and Psychology
Oomen T.; Gonçalves J.; Mols A.	Rage Against the Artificial Intelligence? Understanding Contextuality of Algorithm Aversion and Appreciation	2024	International Journal of Communication
Proksch S.; Schühle J.; Streeb E.; Weymann F.; Luther T.; Kimmerle J.	The impact of text topic and assumed human vs. AI authorship on competence and quality assessment	2024	Frontiers in Artificial Intelligence
Rabinovitch H.; Budescu D.V.; Meyer Y.B.	Algorithms in selection decisions: Effective, but unappreciated	2024	Journal of Behavioral Decision Making
Tse T.T.K.; Hanaki N.; Mao B.	Beware the performance of an algorithm before relying on it: Evidence from a stock price forecasting experiment	2024	Journal of Economic Psychology
Chang Y.; Wang R.	Conservatives endorse Fintech? Individual regulatory focus attenuates the algorithm aversion effects in automated wealth management	2023	Computers in Human Behavior
Cheng L.; Chouldechova A.	Overcoming Algorithm Aversion: A Comparison between Process and Outcome Control	2023	Conference on Human Factors in Computing Systems - Proceedings
Filiz I.; Judek J.R.; Lorenz M.; Spiwoks M.	The extent of algorithm aversion in decision-making situations with varying gravity	2023	PLoS ONE
Fine A.; Le S.; Miller M.K.	Content Analysis of Judges’ Sentiments Toward Artificial Intelligence Risk Assessment Tools	2023	Criminology, Criminal Justice, Law and Society
Germann M.; Merkle C.	Algorithm aversion in delegated investing	2023	Journal of Business Economics

Glienke M.; Hartwich N.; Antons D.	Working with AI: How Attitudes Shape Human-AI Collaboration	2023	International Conference on Information Systems, ICIS 2023
Himmelstein M.; Budescu D.V.	Preference for human or algorithmic forecasting advice does not predict if and how it is used	2023	Journal of Behavioral Decision Making
Kim A.; Yang M.; Zhang J.	When Algorithms Err: Differential Impact of Early vs. Late Errors on Users' Reliance on Algorithms	2023a	ACM Transactions on Computer-Human Interaction
Liu N.T.Y.; Kirshner S.N.; Lim E.T.K.	Is algorithm aversion WEIRD? A cross-country comparison of individual-differences and algorithm aversion	2023	Journal of Retailing and Consumer Services
Longoni C.; Cian L.; Kyung E.J.	Algorithmic Transference: People Overgeneralize Failures of AI in the Government	2023	Journal of Marketing Research
Mahmud H.; Islam A.K.M.N.; Mitra R.K.	What drives managers towards algorithm aversion and how to overcome it? Mitigating the impact of innovation resistance through technology readiness	2023	Technological Forecasting and Social Change
Mok L.; Nanda S.; Anderson A.	People Perceive Algorithmic Assessments as Less Fair and Trustworthy Than Identical Human Assessments	2023	Proceedings of the ACM on Human-Computer Interaction
Neumann M.; Niessen A.S.M.; Hurks P.P.M.; Meijer R.R.	Holistic and mechanical combination in psychological assessment: Why algorithms are underutilized and what is needed to increase their use	2023	International Journal of Selection and Assessment
Neumann M.; Niessen A.S.M.; Meijer R.R.	Predicting decision-makers' algorithm use	2023	Computers in Human Behavior
Reich T.; Kaju A.; Maglio S.J.	How to overcome algorithm aversion: Learning from mistakes	2023	Journal of Consumer Psychology
Schulte Steinberg A.L.; Hohenberger C.	Can AI close the gender gap in the job market? Individuals' preferences for AI evaluations	2023	Computers in Human Behavior Reports

Turel O.; Kalhan S.	Prejudiced against the Machine? Implicit Associations and the Transience of Algorithm Aversion	2023	MIS Quarterly: Management Information Systems
Bonezzi A.; Ostinelli M.; Melzner J.	The Human Black-Box: The Illusion of Understanding Human Better Than Algorithmic Decision-Making	2022	Journal of Experimental Psychology: General
Chacon A.; Kausel E.E.; Reyes T.	A longitudinal approach for understanding algorithm use	2022	Journal of Behavioral Decision Making
Claudy M.C.; Aquino K.; Graso M.	Artificial Intelligence Can't Be Charmed: The Effects of Impartiality on Laypeople's Algorithmic Preferences	2022	Frontiers in Psychology
Commerford B.P.; Dennis S.A.; Joe J.R.; Ulla J.W.	Man Versus Machine: Complex Estimates and Auditor Reliance on Artificial Intelligence	2022	Journal of Accounting Research
Daschner S.; Obermaier R.	Algorithm aversion? On the influence of advice accuracy on trust in algorithmic advice	2022	Journal of Decision Systems
Fügener A.; Grahl J.; Gupta A.; Ketter W.	Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation	2022	Information Systems Research
Ganbold O.; Rose A.M.; Rose J.M.; Rotaru K.	Increasing Reliance on Financial Advice with Avatars: The Effects of Competence and Complexity on Algorithm Aversion	2022	Journal of Information Systems
Heßler P.O.; Pfeiffer J.; Hafenbrädl S.	When Self-Humanization Leads to Algorithm Aversion: What Users Want from Decision Support Systems on Prosocial Microlending Platforms	2022	Business and Information Systems Engineering
Jauernig J.; Uhl M.; Walkowitz G.	People Prefer Moral Discretion to Algorithms: Algorithm Aversion Beyond Intransparency	2022	Philosophy and Technology
Lacroux A.; Martin-Lacroux C.	Should I Trust the Artificial Intelligence to Recruit? Recruiters' Perceptions and Behavior When Faced With Algorithm-Based	2022	Frontiers in Psychology

	Recommendation Systems During Resume Screening		
Maasland C.; Weißmüller K.S.	Blame the Machine? Insights From an Experiment on Algorithm Aversion and Blame Avoidance in Computer-Aided Human Resource Management	2022	Frontiers in Psychology
Neumann M.; Niessen A.S.M.; Tendeiro J.N.; Meijer R.R.	The autonomy-validity dilemma in mechanical prediction procedures: The quest for a compromise	2022	Journal of Behavioral Decision Making
Pezzo M.V.; Nash B.E.D.; Vieux P.; Foster-Grammer H.W.	Effect of Having, but Not Consulting, a Computerized Diagnostic Aid	2022	Medical Decision Making
Saragih M.; Morrison B.W.	The Effect of past Algorithmic Performance and Decision Significance on Algorithmic Advice Acceptance	2022	International Journal of Human-Computer Interaction
Schneider S.; Freisinger E.	Overcoming algorithm aversion: the power of task-procedure fit	2022	Academy of Management Annual Meeting Proceedings
Berger B.; Adam M.; Rühr A.; Benlian A.	Watch Me Improve—Algorithm Aversion and Demonstrating the Ability to Learn	2021	Business and Information Systems Engineering
Bigman Y.E.; Yam K.C.; Marciano D.; Reynolds S.J.; Gray K.	Threat of racial and economic inequality increases preference for algorithm decision-making	2021	Computers in Human Behavior
Bonezzi A.; Ostinelli M.	Can Algorithms Legitimize Discrimination?	2021	Journal of Experimental Psychology: Applied
Borau S.; Otterbring T.; Laporte S.; Fosso Wamba S.	The most human bot: Female gendering increases humanness perceptions of bots and acceptance of AI	2021	Psychology and Marketing

Filiz I.; Judek J.R.; Lorenz M.; Spiwoks M.	Reducing algorithm aversion through experience	2021	Journal of Behavioral and Experimental Finance
Hou Y.T.-Y.; Jung M.F.	Who is the Expert? Reconciling Algorithm Aversion and Algorithm Appreciation in AI-Supported Decision Making	2021	Proceedings of the ACM on Human-Computer Interaction
Jung M.; Seiter M.	Towards a better understanding on mitigating algorithm aversion in forecasting: an experimental study	2021	Journal of Management Control
Kawaguchi K.	When will workers follow an algorithm? A field experiment with a retail business	2021	Management Science
Keding C.; Meissner P.	Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions	2021	Technological Forecasting and Social Change
Ochmann J.; Zilker S.; Michels L.; Tiefenbeck V.; Laumer S.	The influence of algorithm aversion and anthropomorphic agent design on the acceptance of AI-based job recommendations	2021	International Conference on Information Systems, ICIS 2020 - Making Digital Inclusive: Blending the Local and the Global
Renier L.A.; Schmid Mast M.; Bekbergenova A.	To err is human, not algorithmic – Robust reactions to erring algorithms	2021	Computers in Human Behavior
Snow T.	From satisficing to artificing: The evolution of administrative decision-making in the age of the algorithm	2021	Data and Policy
Zhang K.	How information processing style shapes people's algorithm adoption	2021	Social Behavior and Personality
Dietvorst B.J.; Bharti S.	People Reject Algorithms in Uncertain Decision Domains Because They Have Diminishing Sensitivity to Forecasting Error	2020	Psychological Science

Efendić E.; Van de Calseyde P.P.F.M.; Evans A.M.	Slow response times undermine trust in algorithmic (but not human) predictions	2020	Organizational Behavior and Human Decision Processes
Niszczota P.; Kaszás D.	Robo-investment aversion	2020	PLoS ONE
Castelo N.; Bos M.W.; Lehmann D.R.	Task-Dependent Algorithm Aversion	2019	Journal of Marketing Research
Dietvorst B.J.; Simmons J.P.; Massey C.	Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them	2018	Management Science
Prahl A.; Van Swol L.	Understanding algorithm aversion: When is advice from automation discounted?	2017	Journal of Forecasting
Dietvorst B.J.; Simmons J.P.; Massey C.	Algorithm aversion: People erroneously avoid algorithms after seeing them err	2015	Journal of Experimental Psychology: General

4. Study 3: Balancing Confidence and Skepticism in Predictive Forecasting - The Role of Perceived Credibility

4.1 Description of the research project

This research project is a collaboration with Alina Timmermann, who is currently a PhD candidate at the "Chair of Modeling & Simulation" of the department of computer science at TU Dortmund University and Prof. Dr. Ann K. Tank from the University of Groningen (at July 2025). The initial idea for this project emerged within the context of the GRK 2193 "Adaptive Intelligence of Factories in a Dynamic and Complex Environment" of the Deutsche Forschungsgemeinschaft (DFG, engl: German Research Foundation), where Alina Timmerman and the author of this dissertation were both active members. The paper will be submitted to the journal "Management Science" (VHB-Jourqual 4: A+ across several sections, Impact Factor (2024): 4.9) in October 2025. The author of this dissertation made several large contributions to the project. First, he has classified the project in the literature and has derived the research hypotheses from prior literature. Second, he has assisted in the operationalization of the derived hypotheses in a behavioral experiment, particularly in the design of the experimental questionnaire. Third, he has statistically analyzed the collected data from participants. Fourth, he has written large parts of earlier manuscripts and most of the following working paper (Sections 4.5.2, 4.5.3 and 4.5.4 were written by another co-author).

The project included two rounds of data collection, where the first round was realized in the behavioral lab of the University of Groningen and the second round was realized via the online platform Prolific.[19] An early abstract of the paper was accepted for the mentoring program of the MAS Midyear Meeting 2024 and discussed with Prof Dr. Andrew Newman (University of South Carolina). A first full manuscript of the project was accepted for presentation at the International Symposium of Accounting Information Systems (ISAIS) 2024 and the European Network for Experimental Accounting Research (ENEAR) Conference 2024. The paper was discussed by Prof. Dr. Markus Arnold

[19] Note that the first round of data collection utilized a complicated 2x5 mixed design. This was criticized at several conferences, where the paper was presented. To reduce complexity, we switched to two substudies with a 2x2 between-subjects design for the second round of data collection. Furthermore, we excluded a third examined forecast metric based on confidence intervals in the second round of data collection. We are currently working on a follow-up paper to the present study that exclusively focusses on the perception of respective confidence intervals with the working title "Prediction Uncertainty Intervals in Industrial Demand Planning". Data collection for this additional paper has already concluded and a first draft of the paper is planned for autumn 2025.

(University of Bern) at ENEAR 2024. After the second round of data collection, a revised abstract highlighting the core contributions of the project was accepted at the New Scholars Consortium of the Joint Midyear Meeting of the AIT and SET Sections of the AAA 2025 including valuable feedback from Prof. Dr, Andrea Kelton (Middle Tennessee State University) and Prof. Dr. Peter Kipp (University of North Texas).

4.2 Abstract

As organizations increasingly adopt predictive analytics for forecasting, managers face the dual challenge of relying on credible forecasts while remaining skeptical of inaccurate ones. This study investigates how different forms of additional information support managers in assessing the credibility of forecasts provided by predictive analytics tools. Drawing on the literature on algorithm aversion, source credibility, and technology trust, we explore the unintended consequences of commonly suggested strategies that aim towards emphasizing a tool's forecasting capabilities in order to overcome phenomenon like algorithm aversion. Using two randomized between-subjects experiments (study 1a, n = 197; study 1b, n = 203), we find that presenting a metric which reflects the inherent predictability of the data on which the forecast is based on along with the output of a forecasting tool, effectively supports participants to more realistically assess the credibility of the provided forecast. We conclude that such a predictability-based metric increases the confidence in accurate forecasts and decreases overconfidence in inaccurate ones and consequently facilitates managers to better evaluate the forecast quality. In contrast, communicating the capabilities of the forecasting tool alone increases the perceived credibility of participants independent of the actual accuracy of the forecast, alluding to the unconditional overestimation of the accuracy of the forecast. Our findings caution against one-sided credibility cues and propose a novel, task-related metric to balance confidence and skepticism in predictive forecasting. This contributes to both theory and practice by offering a more nuanced approach to enhancing human–analytics collaboration in organizational decision-making.[20]

[20] **Data Availability:** Raw data are available from the authors upon request. The data supporting the findings of this study and the final experimental questionnaire are openly available at open science framework (OSF):
https://osf.io/t7a2s/?view_only=99accd7904fa4084837413ab2b5ea601

4.3 Introduction

Recent developments in the field of Artificial Intelligence (AI) allow for very precise and fast data-driven estimations of the future. Advanced predictive analytics tools allow to explore large amounts of unstructured datasets as well as to discover unknown relationships between them and hereby often exceed the capabilities of human forecasters (Hamoudia et al. 2023). Examples include the prediction of future earnings changes (Chen et al. 2022b), the prediction of future market demands and sales (Ensafi et al. 2022) or the prediction of future exchange rates (Chen et al. 2021). However, it is highly unlikely that the advanced tools and methods always predict future developments perfectly. For example, the unpredictable nature of some future events is illustrated by the impact of the COVID19 pandemic (Chharia et al. 2024, Ioannidis et al. 2022). Moreover, the performance of respective tools also critically depends on the time horizon and the quality of the provided data and their technical suitability to estimate certain types of patterns from the provided data (Petropoulos et al. 2022, Fildes 1989). Human judgment – capable of identifying and disregarding respective errors – will therefore remain relevant (De-Arteaga et al. 2020, van Donselaar et al. 2010).[21]

In particular, this creates a challenge for organizations leveraging respective technologies: managers simultaneously need to be able to (1) identify credible predictions and (2) disregard erroneous or inaccurate ones. Although this duality is well recognized in the literature on computer credibility (e.g., Fogg & Tseng 1999) and technology trust (e.g., Lee & See 2004), recent management research on the use of predictive analytics tools has focused mainly on the irrational rejection of credible forecasts, a phenomenon known as *algorithm aversion* (e.g., Greiner et al. 2025, Kawaguchi 2021, Dietvorst et al. 2018). However, we know little about how well decision-makers recognize inaccurate forecasts and what form of decision-support they need to be able to correctly assess the credibility of a forecast (De-Arteaga et al. 2020). In this study we show that previously suggested mitigation strategies to counter algorithm aversion can unintentionally reduce the awareness for inaccurate forecasts. We propose and test an alternative and more effective approach considering the balance between confidence and skepticism in cognitively appraising forecasting credibility.

[21] Van Donselaar et al. (2010) study supermarket managers' adjustments to an automated order advice system and found that due to system inadequacy and incentive misalignment managers adjust the order advices by postponing orders from peak to nonpeak days to achieve a more balanced workload in the store and by factoring in experiences about, e.g., demand uncertainty or seasonality.

One of the central findings of prior research on algorithm aversion is that decision-makers equate forecast errors to missing capabilities of a predictive analytics tool (e.g., Berger et al. 2021, Dietvorst et al. 2015). Prior literature therefore recommends capability-based mitigation strategies to counter algorithm aversion, for example communicating a tools learning capabilities (e.g., Commerford et al. 2024, Reich et al. 2023, Berger et al. 2021, Castelo et al. 2019). While we acknowledge the efficiency of these approaches, we argue that they simultaneously reduce decision-makers awareness of inaccurate forecasts. In line with the *Perfect Automation-Scheme* (Prahl & van Swol 2017, Madhavan & Wiegmann 2007), prior research also identifies irrationally high baseline expectations about the accuracy provided by predictive analytics forecasting tools (Dietvorst et al. 2015) and a significant discounting of a forecast when experiencing a forecasting error (Dietvorst & Bharti 2020). This cognitive scheme indicates that decision-makers underestimate the likelihood of forecast errors which are not only related to the capabilities of the forecasting tool and consequently, highlighting respective capabilities reinforces this biased perception.

Based on these findings, we theorize that next to capability-based indications enabling decision-makers to correctly identify the expectable fundamental difficulty of a forecasting task and in particular the predictability of the input data has the potential to reduce this cognitive bias. A respective predictability-based metric, based on the distribution and quantity of patterns in historic data (Salvino et al. 1995), provides a predictor of forecast accuracy and thus gives an indication for the credibility of the forecast beyond the capabilities of a predictive analytics tool. We hypothesize that a predictability-based metric supports managers in putting the capabilities of a forecasting tool into perspective, especially, in a low predictability scenario.

To test our rationale, we conducted two 2 x 2 randomized between-subjects online experiments for which we recruited 400 participants (197 participants in study 1a and 203 participants in study 1b) via Prolific. In both studies, participants were provided with several time series of past sales and sales forecasts for the next period generated by a predictive analytics tool. They then had to indicate their perception of the credibility of the provided sales forecasts, which serves as the main dependent variable in our studies. The conducted studies differed in the actual credibility of the provided forecasts, where study 1a included only credible forecasts based on highly predictable time series and study 1b included only forecasts based on barely predictable time series, thus including potentially inaccurate forecasts. To test our derived hypotheses, we manipulated the display of additional metrics in each study, in particular the display of a capability-based metric (shown vs. not shown) and the display of a predictability-based metric (shown vs. not shown). While the capability-based metric always confirmed the presence of high forecasting capabilities, the predictability metric indicated a high predictability in study 1a and a low predictability in study 1b. To generate the experimental stimuli, we got access to real sales data from a European manufacturer as well as a real analytics tool.

Consistent with our hypotheses, participants rated provided forecasts as significantly more credible when they are shown a metric signaling a high level of the underlying predictive analytics forecasting capabilities. Confirming our theory-based concerns regarding reduced awareness for forecast errors, this effect occurs for accurate forecasts but also for inaccurate ones. Furthermore, we find evidence for the effectiveness of the predictability based-metric to calibrate the perceived forecast credibility. While displaying a metric which signals a high level of predictability of the input data leads to a significantly higher perceived credibility of the provided forecast, displaying a metric which signals a low level of predictability leads to a significantly lower perceived credibility. Surprisingly, we find an unexpected significant but negative interaction between a high capability-metric and a high predictability metric on forecast credibility, as well as no significant interaction between a high capability-metric and a low predictability metric.

To further analyze the surprising interaction results, we calculate a moderated mediation model for each study with the perceived forecasting capabilities of the predictive analytics tool as mediator and the predictability-based metric as moderator.[22] In study 1a, the results support our rationale and reveal that the positive effects of showing the capability-based metric on perceived credibility can be explained by the mediating effect of the perceived capabilities but we do not find any significant moderation effect of the predictability-based metric. In study 1b, the results again reveal that the positive effect of the displayed capability-based metric on perceived credibility can be explained with the help of the perceived capabilities variable. Displaying the predictability-based metric negatively moderates the mediation effect but also positively moderates the direct effect of perceived capabilities on perceived credibility. We interpret this counter-intuitive positive

[22] We tested for the successful manipulation by analyzing the participants perceived capabilities of the forecasting tool. This measure we included in the moderated mediation model.

interaction as an effect of providing two instead of one metrics and therefore more information.

Our study makes several contributions to existing streams of research across management disciplines. First, our research contributes to the rapidly growing literature around algorithm aversion in management and psychology and possible mitigation strategies (e.g., Mahmud et al. 2022, Burton et al. 2020). While several scholars recommend to mitigate algorithm aversion through the communication of a predictive analytics tool's capabilities (e.g., Greiner et al. 2025, Reich et al. 2023, Commerford et al. 2022, Berger et al. 2021, Dietvorst et al. 2015), our research further examines these strategies and observes unintended and potentially negative consequences in form of an overreliance on respective tools in case of inaccurate forecasts. Second, our study contributes to decades of research on computer credibility (Fogg & Tseng 1999) and technology trust (e.g., Lee & See 2004) describing the risk of decision-makers situationally overusing IT-systems. In particular, we propose a novel approach to assist managers to develop a more realistic perception of the credibility of forecasts from predictive analytics tools. Hereby, we answer the call for further research on respective approaches which specifically recommends examining a forecasting setting (De-Arteaga et al. 2020).

Our results are especially relevant for the growing number of organizations interested in leveraging more advanced predictive analytics tools for their planning processes (e.g., Zorn et al. 2025, Petrie et al. 2024). On the one hand, it emphasizes the need to not only control for an underuse of predictive analytics by managers but also for potential overreliance on inaccurate forecasts. On the other hand, because the predictability of a time-series can be calculated with predefined libraries (e.g., Flood & Grimm 2021) and does not require deep knowledge in computer science, developing and displaying a suitable metric to calibrate the credibility perceptions of organization-intern decision-makers represents a feasible and auspicious approach.

4.4 Background and hypotheses development

4.4.1 Benefits and challenges of predictive analytics

The term predictive analytics describes advanced statistical methods and approaches, which predict future developments of key figures from past data (e.g., Holsapple et al. 2014). Predictive analytics include numerous forecasting methods and algorithms, ranging from simple regression models to advanced machine-learning approaches. A recent forecasting tournament, for example, included over 60 different applied forecasting methods from the participants (Makridakis et al. 2020b). Predictive analytics can complement or even exceed human judgment in many decision contexts and improve overall performance (e.g., Dellermann et al. 2019, Mikalef et al. 2019). Chen et al. (2022b), for example, train a machine-learning model to predict future earnings, benchmark it against the human judgment of financial analysts and observe a significant increase in forecasting accuracy. Ensafi et al. (2022), show that forecasting approaches build upon neural networks further improve classical algorithms in sales forecasting, which are already superior to human judgment. Chen et al. (2021) provide a machine-learning-based forecasting model, which is able to predict highly volatile settings as the Bitcoin exchange rate.

However, forecasts generated by predictive analytics are still subject to error. Certain events remain unpredictable, for example the credit crisis in 2007 (e.g., Bezemer 2010) or the COVID19 pandemic in 2020 (e.g., Chharia et al. 2024, Ioannidis et al. 2022), and have the potential to lower the accuracy of every forecast including predictive analytics tools or methods (Petropoulos et al. 2022). Moreover, the effectiveness of predictive analytics depends on the completeness of the provided information and can lack human-specific knowledge about the forecasting task (Choudhury et al. 2020, Fildes et al. 2009, Donihue 1993). Sales forecasts, for instance, often include oral agreements in planning meetings, which are difficult to systematically consider in a predictive analytics model (e.g., Goretzki & Messner 2016). Additionally, not every predictive analytics method is suited for every forecasting task (Petropoulos et al. 2022). For example, linear regression models are conceptually not able to account for non-linear relations in a dataset (e.g., Zhang 2003).

4.4.2 Incredulity and gullibility errors

To leverage predictive analytics in planning processes, decision-makers are required to incorporate only credible forecasts from predictive analytics into their decisions and to disregard erroneous ones. Due to the aforementioned dependencies of forecasting errors on unpredictable events and the completeness of the provided information, this requirement translates to a constant evaluation of the credibility of every generated forecast.[23]

[23] In a human-computer-interaction context, the credibility of an IT-system can be defined as the believability of its outputs (Fogg & Tseng 1999).

For example, when unpredictable events as the COVID19 pandemic end, previously erroneous forecasts might significantly improve in accuracy. Moreover, a previously accurate forecast could lack critical information about newly implemented tariffs leading to a significant decrease in later forecast accuracy. Worryingly, prior empirical research observes that decision-makers are not always able to distinguish between credible and uncredible forecasts. Following prior literature on computer credibility, we distinguish between *incredulity* and *gullibility errors*. While incredulity errors translate to decision-makers perceiving an accurate forecast as not credible, gullibility errors translate to decision-makers perceiving an erroneous or inaccurate forecast as credible (Fogg & Tseng 1999).

Prior research on predictive analytics has mostly focused on incredulity errors. An interdisciplinary stream of literature centered around *algorithm aversion* constantly reports a situational reluctance of decision-makers to rely on credible forecasts from predictive analytics (e.g., Mahmud et al. 2022, Kawaguchi 2021, Burton et al. 2020, Dietvorst et al. 2018), even if they are aware of their probable credibility (Dietvorst et al. 2015). These incredulity errors are associated with an avoidable loss in forecast accuracy (e.g., Dietvorst et al. 2018, Dietvorst et al. 2015). For example, sales forecasting is expected to greatly benefit from predictive analytics because these can incorporate the large quantity of unstructured social media data into their forecasts exceeding human capabilities (Iftikhar & Khan 2020, Cui et al. 2018). Yet, prior research finds that decision-makers often disregard respective forecasts, particularly when these incorporate respective social media datasets (Fehrenbacher et al. 2023) and predict a downward trend for future sales (Chen et al. 2022a).

While prior literature only sporadically mentions the possibility that decision-makers rely on erroneous forecasts (e.g., Feiler & Tong 2022), it nevertheless indicates a high susceptibility of decision-makers to gullibility errors in credibility evaluation. For example, prior studies observe a situational but significant overweighing of a provided forecast when decision-makers have little to no experience with the underlying analytics tool (e.g., Berger et al. 2021, Prahl & van Swol 2017).[24] Furthermore, decision-makers also place irrational high trust in predictive analytics when respective tools are described as very advanced (Hou & Jung 2021) or when the forecasting task is perceived as objective

[24] Note that rational forecasting would require averaging over a provided forecast from an unknown source and one's own expectations (Dawes 1979).

(Castelo et al. 2019). While gullibility errors are again associated with an avoidable loss in forecast accuracy, they can also decrease the forecasting skills of human-decision-makers in the long term (Sutton et al. 2023, Sparrow et al. 2011, Endsley 1995).

In the following sections, we identify additional evidence for the relevance of gullibility errors in the given forecasting setting. In particular, we argue that the numerous proposed mitigation strategies for incredulity errors can inadvertently trigger gullibility errors. Thus, we propose an alternative mitigation strategy which addresses incredulity errors and gullibility errors simultaneously.

4.4.3 Swinging from incredulity errors to gullibility errors

One of the central findings of the literature stream on incredulity errors is that decision-makers appear to sometimes underestimate the actual capabilities of a predictive analytics tool (Chen et al. 2022a, Commerford et al. 2022). Examples for these erroneous assumptions regarding predictive analytics' capabilities include the perceived lack of learning capabilities (Dietvorst et al. 2015) and the perceived lack of forecasting capabilities in subjective settings (Castelo et al. 2019) among others. To mitigate the observed resistance towards predictive analytics, prior research therefore often recommends to inform decision-makers about the actual capabilities of a predictive analytics tool. These approaches include directly highlighting the actual capabilities of a predictive analytics tool, for example, by communicating an analytics tools learning capability (Reich et al. 2023, Berger et al. 2021) among others. However, similar results can be achieved by including humans into the forecasting process, which possess the supposedly missing capabilities. Examples for this human involvement can include providing data input rights or output adjustment rights to decision-makers (Commerford et al. 2024, Dietvorst et al. 2018).

These mitigation strategies are in line with existing knowledge on individuals' assessment of credibility and their effectiveness is expectable. For example, the computer-credibility-framework predicts that a computer's presumed capabilities or expertise are one of two relevant factors for its later credibility assessment by decision-makers, with the other one being perceived trustworthiness (Fogg & Tseng 1999). Furthermore, the scale-adjustment model identifies the perceived expertise as the major determinant of a source's credibility and persuasion (Birnbaum & Stegner 1979). Lastly these mitigation strategies clearly target the aforementioned forecasting errors related to a predictive analytics tool or methods suitability for a given forecasting task (e.g., Petropoulos et al. 2022).

However, another central finding of the algorithm aversion literature is that these incredulity errors often occur when predictive analytics tools are perceived or experienced as error-prone. For example, Dietvorst et al. (2015) and Berger et al. (2021), only observe a significant reluctance of decision-makers to rely on a provided forecast when they experience the fallibility of the underlying analytics tool beforehand. In line with this reasoning, Logg et al. (2019) never provide decision-makers with performance information of a provided predictive analytics tool and never observe incredulity errors. Dietvorst & Bharti (2020) explain this contingency with a preference of decision-makers for highly accurate forecasts and show that decision-makers have a diminishing sensitivity to forecasting errors, where small forecast errors are penalized more heavily than large ones.

Combining these described findings, we assume that decision-makers on average do not recognize the different sources of forecasting errors and mistakenly confuse forecasting capabilities with highly accurate forecasts and vice versa confuse forecasting errors with a lack of forecasting capabilities. This argument is supported by previous findings of Dietvorst et al. (2015), who observe that their participants expected a predictive analytics tool to hit a perfect prediction every second forecast before learning about the tool's actual performance. As the forecasting tasks of this study included the prediction of future flight passengers in the US, this belief was totally unreasonable for the given context (e.g., Zachariah et al. 2023). Similar overestimations of an IT-system's capabilities are well-known in prior literature and are called the perfect automation scheme (Prahl & van Swol 2017, Madhavan & Wiegmann 2007).

Drawing on this rationale, we hypothesize that communicating a predictive tool's capabilities might convince decision-makers that the provided forecasts are likely accurate and credible. This approach is clearly helpful in situations where this assumption is justified and decision-making can benefit from incorporating the provided forecast. However, the respective approach clearly does not address the underlying perfect automation scheme. It therefore does not provide any guidance to decision-makers in contexts where the provided forecasts are likely inaccurate from other causes than missing forecasting capabilities. Furthermore, a respective capabilities-based strategy will rather reinforce decision-makers erroneous belief that forecast accuracy and forecasting capabilities are necessarily related. In turn this should make decision-makers even more blind to the risk of gullibility errors.[25] We consequently hypothesize that addressing the capabilities of a predictive analytics tools provides an effective strategy to increase its perceived credibility irrespective of whether this increase is justified or not. The first hypothesis states the following:

H1 Highlighting solely the capabilities of a predictive analytics tool increases the perceived credibility of a generated forecast.

4.4.4 Jointly mitigating incredulity and gullibility errors

It follows from the previous section that, instead of informing decision-makers about the capabilities of a predictive analytics tool, it could be fruitful to instead inform decision-makers about the expectable likelihood of forecast errors. Assuming that a forecaster has the necessary forecasting capabilities, the achievable forecasting accuracy is determined by the occurrence and density of observable patterns in past data (e.g., Petropoulos et al. 2022, Makridakis et al. 2020a, Makridakis & Taleb 2009). The balance of unpredictable events and predictable patterns characterizes the actual predictability of a time series of historical values. As *time series predictability* is associated with the presence of predictable patterns, it provides an upper bound to the maximum achievable forecasting performance for a time series irrespective of the applied forecasting method (e.g., Salvino et al. 1995). While a high time series predictability indicates that predictive analytics would in theory be able to provide an informative forecast, a low time series predictability suggests a high likelihood of forecasting errors and a rethinking of whether this time series is worth the effort. Technical developments in computational power and developments around the concept of entropy have led to widespread use of time series predictability in forecasting research (e.g., Soto-Ferrari 2025, Wang & Tan 2021, Namdari & Li 2019) and its diffusion into corporate practice (e.g., Manokhin 2023, Paialunga 2023).

As the time series predictability is clearly an indicator for the probability of forecast errors, we hypothesize that communicating a high predictability of past data increases the perceived credibility of a generated forecast. Conversely, communicating a low predictability of past data decreases the perceived credibility of a generated forecast. Our second hypothesis is divided in a set of two hypotheses – H2a and H2b – and are stated as follows:

[25] Note that this rationale is also in line with the literature on technology dominance (e.g., Sutton et al. 2023), which predicts that automated decision aids can decrease the accuracy of a task, when decision-makers are not capable of interpreting the output correctly.

H2a. Communicating a high predictability of past data increases the perceived credibility of a generated forecast.

H2b. Communicating a low predictability of past data decreases the perceived credibility of a generated forecast.

From a cognitive point of view, it is unlikely that communicating the underlying time series predictability to decision-makers will resolve their desire for perfect forecasts. This is due to sufficient accuracy being a key requirement for forecasts in organizational planning and controlling activities (e.g., Jordan & Messner 2020, Goretzki & Messner 2016, Sivabalan et al. 2009). Decision-makers likely incorporate the provided information regarding the underlying time series predictability into their assessment of the forecasting capabilities of a predictive analytics tool. In particular, if the conveyed time series predictability is in contradiction to supposedly sufficient forecasting capabilities that lead to near-perfect forecast accuracy, decision-makers might doubt whether these assumed forecasting capabilities are actually valid and adjust their respective evaluation downwards.

We expect that this effect will become visible when communicating time-series predictability and forecasting capabilities simultaneously. Conceptually, we assume a conditional moderation effect of the communication of time-series predictability on the relationship between communicating forecasting capabilities and the perceived credibility. In case of a high time series predictability, additionally communicating this feature to decision-makers should not change their credibility evaluation, because it further emphasizes the existing belief that the communicated forecasting capabilities lead to accurate forecasts. In case of a low time series predictability, communicating this low predictability puts any forecasting capabilities into perspective and decision-makers will likely show the aforementioned doubts about their actual effectiveness. The joint communication should therefore significantly decrease the effect of only communicating a predictive analytics tool's capabilities. The third hypothesis states the following:

H3 Communicating the predictability of past data jointly with the capabilities of a predictive analytics tool diminishes the positive effect of communicating these capabilities in case of a low but not in case of a high predictability.

4.5 Method

4.5.1 Design

We conducted an online experiment to empirically test the effects of the discussed additional forecasting metrics on the perceived credibility of an algorithmically generated forecast. The experiment was divided into two sub-parts (study 1a and study 1b), both following a 2 x 2 between-subjects design, resulting in four conditions in each sub-experiment. The sub-experiments differed only in the predictability of the time series underlying the displayed forecast, where study 1a only included predictable time series and study 1b only included unpredictable time series.[26] In each sub-experiment, we manipulated the communication of the underlying tools forecasting capabilities (not shown vs. shown) and the communication of the predictability of the included time series (not shown vs. shown). Both sub-experiments were created using *Qualtrics* and in both sub-studies participants were randomly assigned to one of the between-subjects conditions.

4.5.2 Experimental setting and task

In the experimental setting participants were first introduced to a fictional company, which is currently implementing a new sales analytics tool for sales forecasting. We ask participants to assume the role of a sales planner working for that company, and their task is to evaluate one specific sales analytics tool that the company considers implementing. Participants receive a detailed introduction to the tool and its outputs. Although the company and the situation described is fictional, the sales data set used to create the stimuli in the experimental conditions is from a real industrial company and the provided forecast is derived from a data analytics tool that we developed and specifically trained for the given dataset. In general, the data analytics tool was designed to generate forecasts for multiple time series of product sales simultaneously. In total, we used historical sales data of about 13,000 products for the data analytics tool, which was split in training data and validation data. Forecasts were performed on the company's training dataset using

[26] Note that every forecasting task is based on an underlying time series with a specific predictability. Therefore, the communication of a high (low) time series predictability requires a forecasting task with a predicable (not predictable) time series. Consequently, we separated two sub-experiments based on the underlying input data for the studies stimuli (high *time series predictability* in study 1a and low *time series predictability* in study 1b). See subsection “Predictability classification” for a detailed description.

SARIMA (Seasonal Autoregressive Integrated Moving Average) method and automated hyperparameter optimization using the python library "pmdarima".

In addition to the actual forecast, participants also saw the underlying time series of sales in graph format. Furthermore, depending on the condition they either saw no additional information, the capabilities-based forecast metric, the predictability-based forecast metric, or both additional metrics. Figure 4-1 provides an example of the information provided to participants in the experimental condition in which both additional metrics were shown.[27] Before the experimental task, all participants saw a detailed example of the task with descriptions of the later interface. Afterwards, we tested whether participants understood the provided introduction using several comprehension checks.[28]

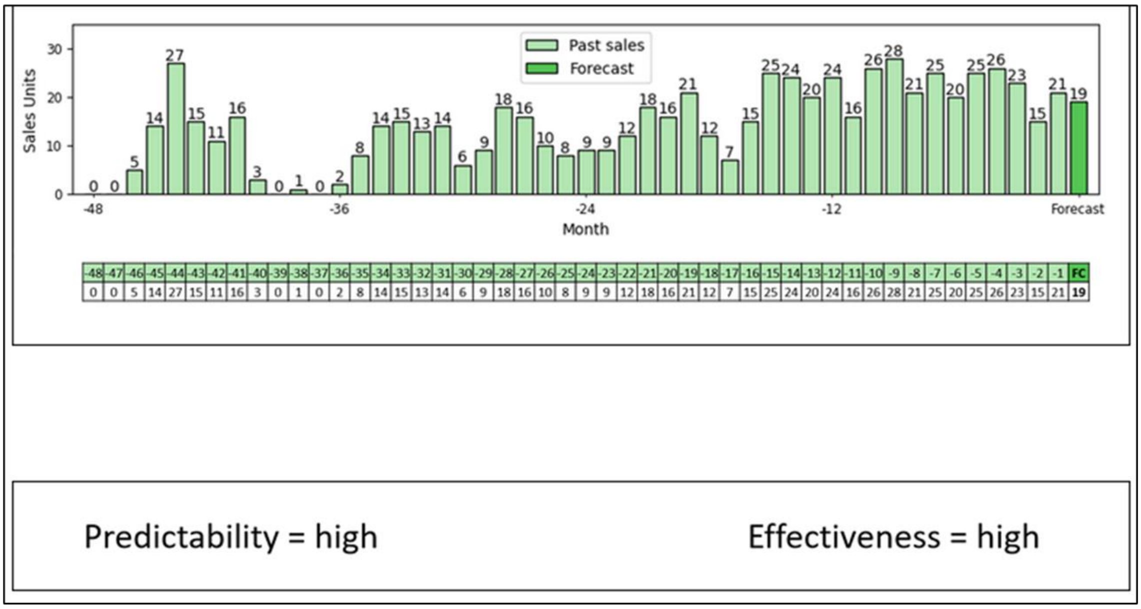

Figure 4-1: The interface during the experiments.[29]

[27] The universities review board approved the experimental procedure and all participants gave their informed consent at the start of the study.
[28] For details see chapter "Manipulation, comprehension and attention checks" below.
[29] Note: The figure shows a screenshot of the designed interface of study 1a in Qualtrics, in the condition in which additional metrics were shown: the predictability-based metric and the capability-based metric. In study 1b the interface was near identical and only predictability was shown as "low". Predictability-based metric = the metric indicates the persistence of predictable patterns in a time series and the likelihood of unpredictable events and is based on the described classification approach to determine the predictability of the included time series. Capabilities-based metric = operationalized by the technical effectiveness; the metric indicates whether the tool is able to identify predictable patterns in past data and to incorporate these patterns into its future predictions and consequently whether the tool has fundamental forecasting capabilities. Dependent variable = Participants rated the perceived credibility four different algorithmically-generated forecasts in random order. Participants to indicate on a seven-point semantic differential: "To what extent do you consider the provided sales forecast to be credible?" in which 1 = not credible and 7 = very credible.

4.5.3 Predictability classification

Our two sub-studies require to divide the past sales data into time series with high and low predictability. To assess the structure of time series, we employ statistical tests that detect patterns or autocorrelation, since the absence of such features is indicative of white noise behavior and is generally associated with increased forecasting difficulty (Shumway & Stoffer 2017). The Weighted Permutation Entropy describes randomity and complexity of time series and is directly linked to their prediction uncertainty (Garland et al. 2014). Ljung-Box detects autocorrelation in a time series, whereby time series with low autocorrelation behave more like white noise and are harder to forecast (Hassani et al. 2025). All measures were calculated for each time series of the company's training dataset. The measures were normalized and aggregated to a Predictability score with 0 representing the lowest and 1 the highest predictability. For a low predictability, time series in the 5th to 10th percentile of the Predictability score were selected. Analogue, for a high predictability, time series in the 90th to 95th percentile of the predictability score were picked. We randomly chose four time series each for our conducted sub-studies.

4.5.4 Independent variables

4.5.4.1. Predictability-based metric

We operationalized the communication of the predictability of past data with the help of one additional forecast metric called *predictability*, which is either displayed or not displayed during forecast evaluation. The metric indicates the persistence of predictable patterns in a time series and the likelihood of unpredictable events and is based on the described classification approach to determine the predictability of the included time series. Consequently, we communicated a time series predictability as being "high" for every forecasting task in the relevant experimental groups of study 1a and as "low" for every forecasting task in the relevant experimental groups of study 1b. In these experimental groups, participants were informed about the meaning of the metric beforehand. In particular, we explained time series predictability as "an inherent feature of a time series. It reflects the persistence of the patterns in the time series and the likelihood of unpredictable events. Therefore, it reflects the extent to which the time series is suitable for a forecasting task". We validated whether participants understood the meaning of the metric with the help of comprehension checks.

4.5.4.2. *Capability-based metric*

We operationalized the communication of the predictive analytics tool's capability using the metric called *technical effectiveness*, which is either displayed or not displayed during forecast evaluation. The metric indicates whether the tool is able to identify predictable patterns in past data and to incorporate these patterns into its future predictions and consequently whether the tool has fundamental forecasting capabilities (Petropoulos et al. 2022). During the development of the forecasting tool, we determined the technical effectiveness of the tool using a statistical analysis of past forecast errors, where the trajectory of past forecast errors should conversely show no signs of a clear trend or seasonality (Hyndman & Athanasopoulos 2018).[30] As our tool showed high technical effectiveness for every examined time series, we communicated a technical effectiveness of "high" for every forecasting task in the relevant experimental groups. In these experimental groups, participants were again informed about the meaning of the metric beforehand. In particular, we explained technical effectiveness as "a feature of the tool that shows how well it can identify and incorporate observable patterns from the time series into its predictions. While errors in forecasts can still happen, these are more likely caused by random, unpredictable events". We again checked whether participants understood the meaning of the metric with the help of several comprehension checks.

4.5.5 Dependent variable

Participants evaluated four different algorithmically-generated forecasts in random order. For every task, we asked participants to indicate their perception of the credibility of the provided forecast. This indication of participants represents our main dependent variable *perceived credibility*. Drawing on prior literature on the credibility of predictive analytics, we measure perceived credibility with the help of the seven-point semantic differential: "To what extent do you consider the provided sales forecast to be credible?" in which 1 = not credible and 7 = very credible (based on Chen et al. 2022a, p. 108). To account for

[30] In detail, this can be validated by performing statistical tests on the residual's trajectory of past forecast errors. The residuals trajectory of a forecast should show no significant autocorrelation (Ljung-Box test), express randomity (Runs Test or Hurst Exponent) and are supposed to be stationary (ADF or KPSS test) which means their statistical properties remain constant over time (Hyndman & Athanasopoulos 2018). While this approach might appear similar to the classification of predictability, recall that predictability is determined by the time series itself, while effectiveness is determined by the trajectory of residuals.

effects of different historic time series trajectories, we calculated and analyzed the average of perceived credibility across all four evaluations.[31]

Other collected variables

To be able to further validate our theoretical rationale, we additionally collected participants perceptions about *capabilities* of the underlying predictive analytics tool to estimate the displayed time series as well as their perception of the *predictability* of the displayed time series. We included two separate Likert-scales and asked participants to indicate their agreement with the statements "The forecast accounts for the predictable patterns in the data" and "The time series itself has a lot of predictable patterns". Moreover, participants filled out an extensive post-experimental questionnaire (PEQ)[32] and had the opportunity to give feedback about the study.[33]

4.5.6 Participants

We recruited 197 participants for study 1a and 203 participants for study 1b via the online labor platform Prolific.[34] On average, the participants in study 1a (study 1b) are 39 (40) years old and 39.6 (37.9) percent of them identify as female. 64 (58) percent were US citizens and 24 (29) percent were UK citizens. 49 (50) percent work for their current employer for more than 5 years and 86.3 (77.8) percent use technology at work more than once a day.[35] Moreover, the participants rate their statistical knowledge on average as 4.56 (4.39) on a scale from "very poor" (1) to "very good" (7). We ran separate ANOVAs to test for a successful randomized assignment of participants to experimental conditions. The results confirm that the randomization was successful, and conditions do not

[31] For example, Fehrenbacher et al. (2023) report a strong effect of the observable time series trend on the later use of an algorithmically-generated forecast, when the predictive analytics tool predicts the breaking of a stable trend. Note that we deliberately chose against including only time series with a clear trend, because we expect our rationale to be independent of the trend in past data.

[32] We collected the highest level of education of the participants, their self-evaluation of statistical knowledge, and their general disposition to technology trust. Prolific provided us with a high amount of demographic data, for instance gender and age, among others.

[33] To validate the relationship between perceived credibility and later use, we also asked participants about their overall evaluation of the provided forecasting tool. Unfortunately, there was an error in the online study, where both ends of the utilized sematic differential stated "The [case company] should not implement the forecasting tool.". While at least one third of our participants noticed this error and contacted us at the end of study explaining their choices, we have no indication from the other participants how they answered this question. Therefore, we excluded this question from further analysis.

[34] To rule out any biases from differences in participants IT-equipment, we restricted participation in our study to computers and laptops.

[35] We required participants to (at least) have a community college degree and (at least) have a part-time job, where they must work with IT-systems among others. A list of requirements is in the Appendix.

systematically differ in participants age (study 1a: $p > 0.54$, study 1b: $p > 0.57$), technology use (study 1a: $p > 0.53$, study 1b: $p > 0.87$) and statistical knowledge (study 1a: $p > 0.26$, study 1b: $p > 0.43$). On average, all 400 participants received £1.50 for participating. The amount represents, based on the median time of 9 minutes and 39 seconds participants needed to finish the study, £9.33 wage per hour.[36]

4.5.7 Manipulation, comprehension and attention checks

To test if the manipulation of the experimental factors was successful and if participants understood the experimental task, we included manipulation, comprehension and attention checks.

Manipulation checks. We tested for the successful manipulation by analyzing the *perceived capabilities* of the forecasting tool and the *perceived predictability* of the time series. The results of the manipulation check show that in study 1a (b) solely communicating the capability-based metric significantly increases the perceived capabilities of the tool (study 1a: $p < 0.01$, study 1b: $p < 0.01$) and that solely communicating the predictability metric significantly increases (significantly decreases) the perceived predictability (study 1a: $p < 0.05$, study 1b: $p < 0.01$), when comparing it to the condition in which participants saw no additional metric. Based on these results we conclude that our manipulation of the *capabilities-based metric* and the *predictability-based metric* was successful in both sub-studies.

Comprehension checks. After participants read the detailed task instructions, we asked them to answer some control questions to test whether they understood the descriptions about 1) the past sales numbers and the forecast, 2) the predictability-based metric and 3) the capability-based metric. All participants saw an example interface of the tool, showing a row of past sales data and a forecast the tool provided for the next period and we asked them to select the correct answer to the following questions: a) "*In the example above, what is the number of sales from 36 months ago?*" and b) "*In the example above, what is the forecast for next month?*". For both questions also the answer option "*I don't know*." was provided. In both studies the majority of participants were able to answer both questions correctly (study 1a: a) correct answers = 191 of 197 participants, b) correct answers

[36] Prolific recommends to pay participants at least £9.00 per hour, while the minimum pay allowed is £6.00 per hour.

= 191 of 197 participants; study 1b: a) correct answers = 197 of 203 participants, b) correct answers = 191 of 203 participants).

Further, depending on the experimental group, we asked participants to identify what the metric for a provided forecast example expresses. For the predictability-based metric and the capability-based metric we asked one specific question referring to the example shown "*In the provided example, is the sales number for next month likely to be predictable or not likely to be predictable?*" / "*In the provided example, is the provided forecast for next month likely to be effective or not likely to be effective?"* and one general question about the meaning of this metric "*What does the provided predictability feature indicate?*" / *"What does the provided effectiveness feature indicate?"*. Again, for both questions (and both metrics) the answer option "*I don't know*." was provided. In both studies the majority of participants who saw the additional metrics were able to answer both questions correctly (study 1a, predictability-based metric: specific) correct answers = 90 of 91 participants, general) correct answers = 88 of 91 participants; capability-based metric: specific) correct answers = 88 of 90 participants, general) correct answers = 84 of 90 participants; study 1b: predictability-based metric: specific) correct answers = 101 of 109 participants, general) correct answers = 103 of 109 participants; capability-based metric: specific) correct answers = 102 of 107 participants, general) correct answers = 100 of 107 participants).

Attention checks. We included seven general attention checks, which were spread across each of both studies. In five of the attention checks participants had to choose the correct color or number asked for in the question or given in the text above. Further, we included one logic question, asking participants to indicate whether they agree or disagree with the following statement: "I dig to the Earth's core every night to charge my phone." To detect AI agents and other bots, we included a question asking whether a bot is answering the questionnaire. In total, only 20 participants (8 in study 1a and 12 in study 1b) failed at least one of the 7 attention checks; most of them (15 participants) failed the logic test.

Before analyzing the collected data, we strictly excluded the data of 26 (36) participants or 13 percent (18 percent) of the participants from study 1a (b), who failed either any of the comprehension checks and/or any of the attention checks, resulting in a final sample of 171 (167) participants in study 1a (b).[37]

4.6 Results

4.6.1 Study 1a

In Table 4-1 are descriptive statistics for the dependent variable across treatment groups for study 1a. We test our hypotheses using a 2 x 2 ANOVA with the *capabilities-based metric* (shown vs. not shown) and *predictability-based metric* (shown vs. not shown) as manipulated factors and the *perceived credibility* as the dependent variable. In hypothesis 1 (H1) we stated that providing a metric that is indicating the capabilities of a predictive analytics tool increases the perceived credibility of a forecast. Supporting H1, our findings show that participants on average perceive the credibility of the provided forecasts as higher when the capabilities-based metric is shown (mean = 5.43, SD = 0.84) compared to when the metric is not shown (mean = 5.10, SD = 0.82) ($F(1,167) = 9.99$; $p < 0.01$). Hypothesis 2a (H2a) predicts that providing a metric indicating a high time series predictability increases the perceived credibility of a forecast. Supporting H2a, we find that, on average, participants perceive the credibility of the provided forecasts as higher when the predictability-based metric is shown (mean = 5.42, SD = 0.80) compared to when the metric is not shown (mean = 5.03, SD = 0.86) ($F(1,167) = 10.11$; $p < 0.01$). However contradicting hypothesis 3, the interaction effect between the capabilities-based metric and the predictability-based metric is marginally significant ($F(1,167) = 3.30$; $p = 0.07$). While communicating one metric increases perceived credibility by 0.62 on average, displaying a second metric only increases perceived credibility by 0.16. Figure 4-2 illustrates the results.[38]

4.6.2 Study 1b

In Table 4-2 are descriptive statistics for the dependent variable across treatment groups in study 1b. Again, we test our hypotheses using a 2 x 2 ANOVA with the capabilities-based metric (shown vs. not shown) and predictability-based metric (shown vs. not shown) as manipulated factors and the perceived credibility as the dependent variable.

[37] Note that we reran all of our statical analyses with the full sample and do not find any notable differences in the results.

[38] A power analysis indicates medium effect sizes for technical effectiveness shown ($\eta^2 = 0.06$) and time series predictability ($\eta^2 = 0.06$) and a small effect size for the interaction ($\eta^2 = 0.02$).

Also in study 1b participants on average perceive the credibility of the provided forecasts as higher when the capabilities-based metric is shown (mean = 4.26, SD = 1.24) compared to when the metric is not shown (mean = 3.68, SD = 1.25) ($F(1,163) = 11.74$; $p < 0.01$), thus supporting H1. H2b states that signaling a low time series predictability decreases the perceived credibility of a forecast. Supporting H2b, we find that, on average, participants perceive the credibility of the provided forecasts as lower when the predictability-based metric is shown (mean = 3.44, SD = 1.14) compared to when the metric is not shown (mean = 4.52, SD = 1.28) ($F(1,163) = 36.78$; $p < 0.01$). Lastly and against our assumptions stated in hypothesis 3, the interaction effect is not significant ($F(1,163) = 0.00$; $p = 0.97$). Figure 4-3 illustrates the findings in study 1b.[39]

4.6.3 Discussion of the findings

Our findings show that decision-makers perceive a predictive analytics tool forecast as more credible when they are informed about the tool's forecasting capabilities. However, we also observe a tendency of decision-makers to disregard potential forecast errors of a respective tool when being informed about a tool's forecasting capabilities. Furthermore, our study provides empirical evidence that communicating the innate predictability of a forecasting task calibrates decision-makers perception of a tool's credibility more effectively. While a respective approach increases the perceived credibility for likely accurate forecasts, it decreases the perceived credibility for probable inaccurate ones. However, we were not able to confirm that time series predictability can be communicated in addition to capability-based metrics and approaches. In particular, we find that indicating forecasting capabilities increases the perceived credibility of a forecast, despite indicating a low time series predictability.

[39] Power analysis indicates a large effect size for *time series predictability* ($\eta^2 = 0.18$), a medium effect size for *technical effectiveness* ($\eta^2 = 0.07$) and no effect size for their interaction ($\eta^2 = 0.00$).

Table 4-1: Results for study 1a.

Descriptive Statistics: mean (standard error) [n]

	Effectiveness metric not shown	Effectiveness metric shown	Overall
Predictability metric not shown	4.74 (0.77) [47]	5.35 (0.82) [43]	5.03 (0.85) [90]
Predictability metric shown	5.36 (0.75) [49]	5.52 (0.86) [32]	5.42 (0.80) [81]
Overall	5.10 (0.82) [96]	5.43 (0.84) [75]	5.22 (0.85) [171]

Conventional ANOVA

Source of Variance	Sum of Squares	df	F	Sig.
Predictability metric	6.43	1	10.11	<0.01
Effectiveness metric	6.35	1	9.99	<0.01
Predictability x Effectiveness	2.10	1	3.30	0.07
Error	106.19	167		

Table 4-2: Results for study 1a.

Descriptive Statistics: mean (standard error) [n]

	Effectiveness metric not shown	Effectiveness metric shown	Overall
Predictability metric not shown	4.22 (1,23) [42]	4.84 (1.28) [40]	4.52 (1.28) [82]
Predictability metric shown	3.11 (1.24) [40]	3.74 (0.96) [45]	3.44 (1.14) [85]
Overall	3.68 (1.35) [82]	4.26 (1.24) [85]	3.97 (1.32) [167]

Conventional ANOVA

Source of Variance	Sum of Squares	df	F	Sig.
Predictability metric	50.98	1	36.78	<0.01
Effectiveness metric	16.27	1	11.74	<0.01
Predictability x Effec-tiveness	0.00	1	0.00	0.97
Error	225.95	163		

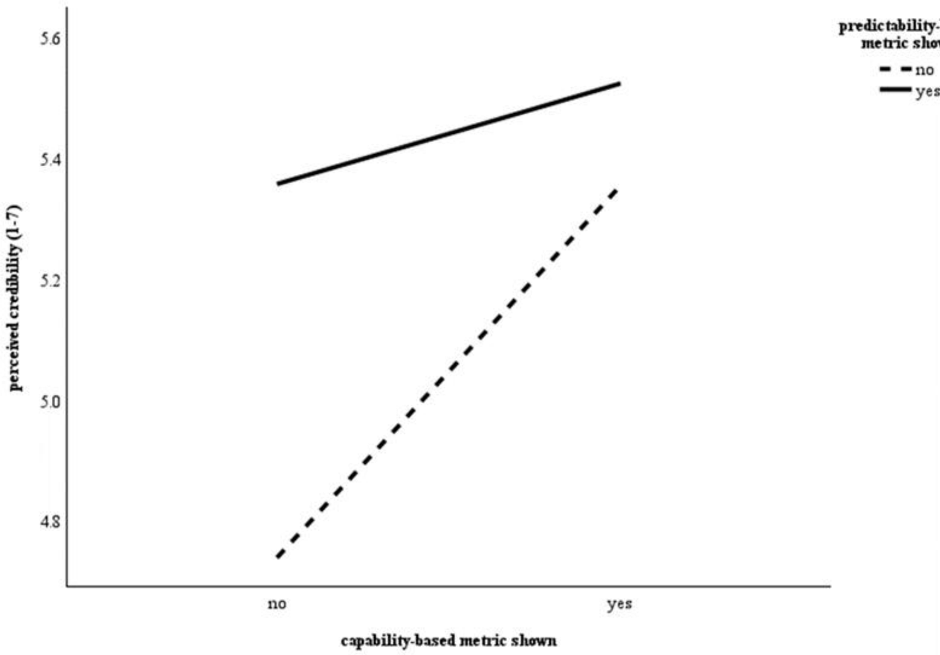

Figure 4-2: Observed effects in study 1a.[40]

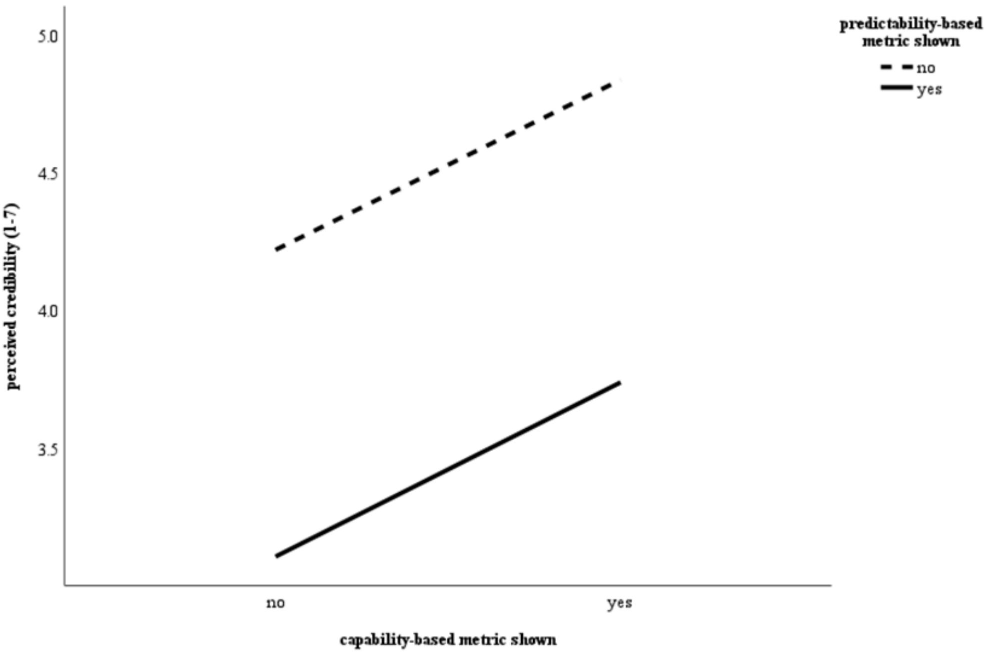

Figure 4-3: Observed effects in study 1b.[41]

[40] Note: This graph visualizes the observed effects of the displayed metrics on the perceived credibility of a provided forecast. See Table 4-1 for a more detailed statistical analysis of the results.
[41] Note: This graph visualizes the observed effects of the displayed metrics on the perceived credibility of a provided forecast. See Table 4-2 for a more detailed statistical analysis of the results.

4.7 Additional analyses

Although our results support our theoretical rationale and most of the derived hypotheses, we further observe an unexpected and marginally significant interaction effect in study 1a, and unexpectedly no significant interaction effect in study 1b. Our theoretical rationale suggests that because time series predictability is a feature independent from a predictive analytics tool, only the capability-based metric technical effectiveness directly influences the perceived capabilities of this tool. The predictability-based metric likely serves as a moderator putting these perceived capabilities into perspective. Furthermore, while previous research suggests that positive moderation effects (as assumed in study 1a) are not distinguishable from an additive model in an ANOVA, weakening moderation effects (as assumed in study 1b) should be (see La Barbera & Ajzen 2020 for a similar problem). To examine this in more detail, for each study, we conducted a moderated mediation analysis, where the *perceived capabilities* is treated as a mediator for the relationship between the communication of the *capability-based metric* (shown vs. not shown) on *perceived credibility* and all relationships are moderated by the communication of the *predictability-based metric* (shown vs. not shown). We used the SPSS Process Macro (model 59) based on Hayes & Little (2018).

Figure 4-4 illustrates the results of the moderated mediation analysis for study 1a. In line with our rationale, we find that the effect of showing the capability-based metric on the perceived credibility was positively and significantly mediated by the perceived capabilities (95% confidence interval of 0.21 to 0.71). We also observe a direct effect of the predictability-based metric on perceived capabilities ($\mathrm{a}_3 = 0.70,\ \mathrm{p} < 0.01$). However, beyond this direct effect, we do not find any significant moderation effect of the predictability-based metric. Figure 4-5 illustrates the moderated mediation model for study 1b. Again, we find that the effect of the capability-based metric on perceived credibility was significantly mediated by perceived capabilities (95% confidence interval of 0.09 to 1.02). We also find the predicted negative moderation effect of the predictability-based metric on the relationship between the capability-based metric and perceived capabilities ($a_2 = -0.94,\ p < 0.05$) and the relationship between perceived capabilities and perceived credibility ($b_2 = -0.42, p < 0.01$). Moreover, we observe an additional positive moderation effect of the predictability-based metric on the relationship between the capability-based metric and perceived credibility ($c_2 = 0.70, p < 0.01$).

The moderated mediation model for study 1a confirms our rationale and suggests that there is indeed no significant interaction effect of displaying both metrics. We therefore explain the almost significant interaction effect in the conducted ANOVA with the help of scale boundaries. Note that displaying only one metric already increases the perceived credibility of a forecast from around 4.5 to almost 5.5 of a maximum of 7. As this is near the upper boundary of the scale, response biases like the tendency to avoid scale ends (Paulhus 1984), might have prevented a regular maximum credibility evaluation. Furthermore, the moderated mediation model for study 1b again confirms our rationale and indicates that there is a significant interaction effect of displaying both metrics, which lowers the perceived credibility instead of increasing it. However, this effect is not visible in the conducted ANOVA because there is also a direct and positive interaction effect on perceived credibility. We assume that this positive interaction could arise from providing two metrics and therefore more information (similar to Castelo et al. 2019). In particular, some participants could appreciate the information that a time series is unpredictable, despite the negative effect on forecast accuracy. However, as this positive interaction effect is completely unexpected, it is an interesting subject for further research.

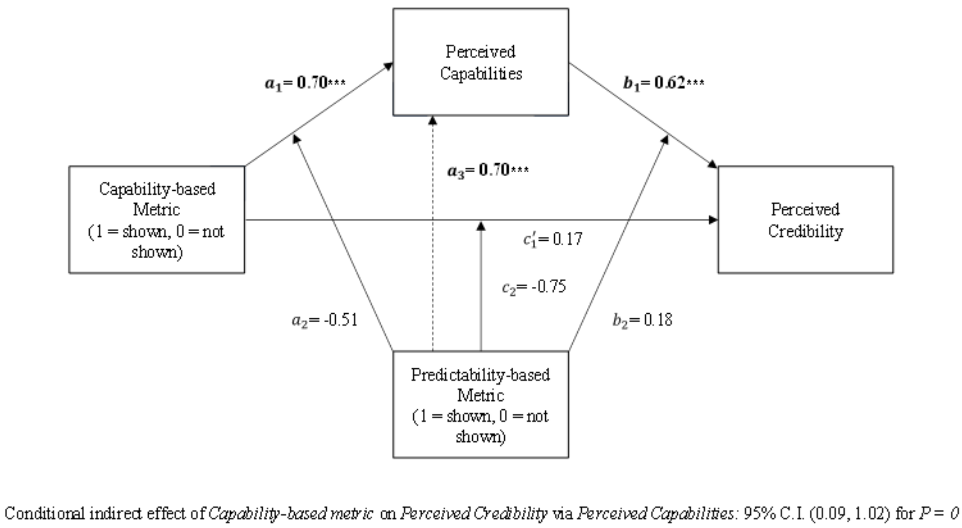

Figure 4-4: Moderated Mediation Model Study 1a.[42]

[42] **, *** Represent significance levels of 0.05, and 0.01, respectively, using two-tailed p-values. Note: The moderated mediation model is based on Model 59 from the Hayes PROCESS Macro (5,000 bootstrap samples) on data of n = 171 participants. P*redictability-based metric* = the metric indicates the persistence of predictable patterns in a time series and the likelihood of unpredictable events and is based on the described classification approach to determine the predictability of the included time series. C*apability-based metric* = operationalized by the technical effectiveness; the metric indicates whether the tool is able to identify predictable patterns in past data and to incorporate these patterns into its future predictions and

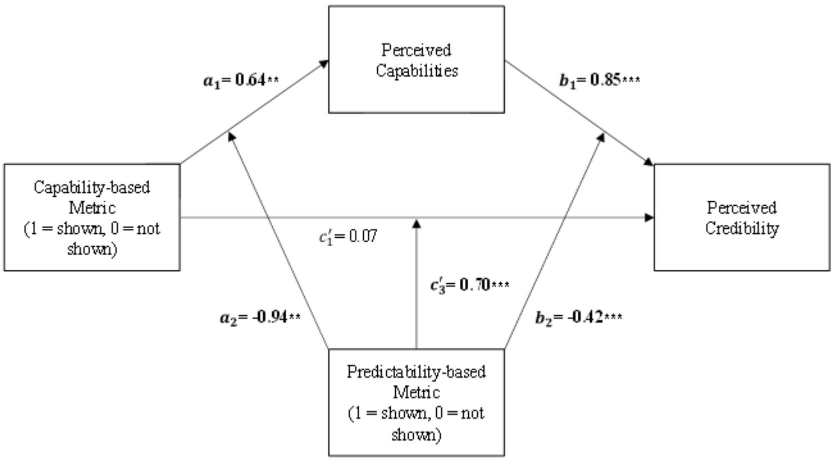

Figure 4-5: Moderated Mediation Model Study 1b.[43]

4.8 Conclusion

Predictive analytics tools have the potential to significantly improve planning and decision-making in organizations. However, this requires decision-makers to use these tools correctly and in particular to distinguish between credible and erroneous forecasts. Unfortunately, decision-makers appear to have difficulties with this task and require additional assessment-support. Prior research has mostly focused on the irrational rejection of credible forecasts and rarely addressed the risk of decision-makers' relying too heavily on erroneous forecasts. This study addresses this gap in prior literature and also answers a direct call for more research on effective decision support in a forecasting setting (De-Arteaga et al. 2020). Using a sales forecasting context, we conduct an online experiment with over 400 participants, which is split into two sub-studies. We utilize a real predictive

consequently whether the tool has fundamental forecasting capabilities. *Dependent variable* = Participants rated the *perceived credibility* four different algorithmically-generated forecasts in random order. Participants to indicate on a seven-point semantic differential: "To what extent do you consider the provided sales forecast to be credible?" in which 1 = not credible and 7 = very credible.

[43] **, *** Represent significance levels of 0.05, and 0.01, respectively, using two-tailed p-values. Note: The moderated mediation model is based on Model 59 from the Hayes PROCESS Macro (5,000 bootstrap samples) on data of n = 167 participants. Predictability-based metric = the metric indicates the persistence of predictable patterns in a time series and the likelihood of unpredictable events and is based on the described classification approach to determine the predictability of the included time series. Capability-based metric = operationalized by the technical effectiveness; the metric indicates whether the tool is able to identify predictable patterns in past data and to incorporate these patterns into its future predictions and consequently whether the tool has fundamental forecasting capabilities. Dependent variable = Participants rated the perceived credibility four different algorithmically-generated forecasts in random order. Participants to indicate on a seven-point semantic differential: "To what extent do you con-sider the provided sales forecast to be credible?" in which 1 = not credible and 7 = very credible.

analytics tool providing sales forecasts based on real historical sales data of a European manufacturing company.

In summary, our experiments show that increasing the perceived credibility of a forecast through the communication of a predictive analytics tool's capabilities can unintentionally make decision-makers less sensitive to potential forecast errors of the tool. Moreover, we propose an alternative approach to calibrate credibility perceptions based on time series predictability. Contrary to established capability-based approaches, we show that communicating the underlying time series predictability of the sales data simultaneously addresses both types of biased credibility perceptions in decision-makers. Lastly, we provide a theoretical explanation for our observed results with the perceived capabilities of a predictive analytics tool as a moderator. We also find an unexpected positive interaction effect, which is an interesting subject for further research. Our findings are relevant for the current management literature on algorithm aversion. In particular, they emphasize the shortcomings of prior mitigation approaches and provide an alternative, which is based on established knowledge from information systems research (e.g., Lee and See 2004, Fogg and Tseng 1999). Further, our study provides a basis for developing an effective implementation strategy for organizations wanting to leverage predictive analytics.

Despite these significant contributions, our study has limitations that provide opportunities for further research. First, in our experiments we use a relatively easy measure of time series predictability. However, while we are aware that even more sophisticated measures exist, i.e. the *coefficient of prescriptiveness* based on Bertsimas & Kallus (2020), which gives an indication of the normative content of data and the effectiveness of a decision by complementing internal data with public online data, the aim of our study is to show that providing additional measures, even simple measures, support users of a forecasting tool to better interpret the credibility of the provided forecast and reduce gullibility errors, decision-makers perceiving an erroneous or inaccurate forecast as credible (Fogg & Tseng 1999). Although we strictly excluded the responses of participants from the analyses who failed any of the comprehension test, we want to acknowledge that comprehension rate was very high, which is advantageous when using easy-to-understand measures in an experimental setting. Further research could therefore validate our rationale using more sophisticated measures like the coefficient of prescriptiveness. However, since to all participants understood the provided measures, another idea could be to use a different operationalization of the predictability-based measure for example based on confidence interval of a forecast (e.g., Hyndman & Athanasopoulos 2018), because these intervals visualize predictability through their width of prediction, thus the visual manipulation might be even easier to understand and increase comprehension of participants.

Second, we find the aforementioned direct and positive interaction effect of communicating both metrics on the perceived credibility of a provided forecast. As we did not expect to observe a respective effect, our experimental design does not allow to analyze this effect in more detail and our explanation for this effect are completely conceptual. Further research could examine weather this effect indeed occurs because of the provision of additional information (e.g., Castelo et al. 2019), even when this information indicates a low forecasting accuracy. Alternatively, a deeper understanding of our observation could also be achieved by conducting in-depth interviews with decision-makers in which they explain their rationale in credibility evaluation.

Furthermore, our experimental setting assumes the introduction of a new predictive analytics tool into the planning processes of the exemplary organization. While this design choice is in line with the current adoption activities in corporate practice, our participants likely only have superficial expectations about the tool's capabilities, which can change with performance experience (Proeger & Meub 2014). Furthermore, prior research in information systems identifies different beliefs about an IT-system pre- and post-adoption (e.g., Karahanna et al. 1999). It would therefore be interesting to examine the effects of the described mitigation strategies function for established tools, where decision-makers already have a strong opinion about their capabilities.

4.9 Appendix

Requirements for eligible participants in both studies:

- Fluent languages: English
- Technology use at work: 4 or 6 times a week, about once a day, more than once a day
- Highest education level completed: Technical/community college, Undergraduate degree (BA/BSc/other), Graduate degree (MA/MSc/MPhil/other), Doctorate degree (PhD/other)
- Number of previous submissions: 3–10000
- Employment Status: Full-Time, Part-Time

- Work Function: Account Management, Administration/ Personal Assistant, CX / Customer Experience / Support, Data Analysis, Healthcare Professional, Education Professional, Engineering (e.g., software), Finance or Accounting, Human Resources, IT / Information Networking / Information Security, Marketing, Operations, Product or Product Management, Project or Program Management, Public Relations / Communications, Sales / Business Development
- Employment-Sector: Business Management & Administration, Finance, Government & Public Administration, Information Technology, Manufacturing, Marketing & Sales, Retail, Science, Technology, Engineering & Mathematics, Transportation, Distribution & Logistics
- First Language: English
- Organizational tenure: 2-4 months, 5-6 months, 7-12 months, 1-2 years, 2-5 years, more than 5 years
- Approval Rate: 100–100
- Exclude participants from other studies: Experiment Sales Analytics- revised - online - test
- Weekly working hours on Prolific: 0 - 5 hours, 6 - 10 hours, 11 - 20 hours, Other

5. General discussion

5.1 Theoretical and practical contributions

The following section discusses how the included studies in this dissertation contribute to the overall research questions. It hereby does not repeat the study-specific contributions that are already stated in the respective sections of the included papers. Instead, it focusses on the most fundamental contributions that emerge when all three studies are viewed together. It distinguishes between contributions regarding the first research question, contributions regarding the second research question and additional implications for corporate practice.

5.1.1 Findings on the first research question

The first overall research question states: “What are relevant factors that encourage or deter managers from incorporating predictive analytics tools into planning activities?”. To answer this research question, the following section again draws on the TPB and distinguishes between managers PBC and attitudes.[44] This dissertation identifies PBC as a central factor for the intention to use predictive analytics tools in planning and forecasting tasks. PBC entails three different dimensions: (1) beliefs about one’s own capability to operate respective tools, (2) beliefs about a tool’s forecasting capabilities and (3) beliefs about the suitability of the available datasets for forecasting. While study 2 explains and transfers PBC to the given contexts, study 1 and study 3 justify these identified dimensions with the help of focused interviews and a laboratory experiment respectively. The first identified dimension is to be expected and, for example, in line with the TPB (e.g., Ajzen 2002), the reliance model of the TTD (e.g., Sutton et al. 2023) and prior research on technology acceptance (e.g., Venkatesh et al. 2003). The third dimension is in line with prior research in accounting that identifies data quality as an important success factor for data analytics (e.g., Appelbaum et al. 2017), AIS and other information systems (e.g., Bai et al. 2012) and management accounting in general (e.g., Knauer et al. 2020). However, study 3 illustrates that not only the actual data quality is important for the success of data analytics, but also the perceived data quality by management.

[44] While study 2 would provide some empirical results for the influence of social norms on the use of predictive analytics, these findings are highly dependent on the special context of the examined federal agency and their generalization is challenging. However, social norms are likely an important determinant of the actual use of respective tools (see section 5.2.1 for a more detailed discussion).

The second dimension seemingly contradicts the technology acceptance literature, which classifies respective beliefs into the attitude-based “performance expectancy” regarding a technology (e.g., Venkatesh et al. 2003). However, recall that PBC and attitude are conceptually connected, where PBC entails the perceived capability to perform of a behavior per se and “attitude” represents the evaluation of the consequences resulting from the performance of a behavior (Fishbein & Ajzen 2010). For the given context of predictive analytics, this translates to the expectation of forecasting capabilities per se (PBC) and the expected accuracy from these forecasting capabilities (Attitude). The included studies in this dissertation therefore suggest that decision-makers only assess the existence of forecasting capabilities when considering the use of a predictive analytics tool and not their consequences. Drawing on the rationale of study 3, this behavior likely occurs because of a perfect automation scheme (Prahl & van Swol 2017, Madhavan & Wiegmann 2007), where decision-makers irrationally and subconsciously expect very accurate forecasts from a predictive analytics tool with perceived forecasting capabilities.

Furthermore, this dissertation illustrates the diversity of beliefs that managers and other decision-makers incorporate into their evaluation of a predictive analytics tool and consequently their attitude towards a respective tool. Attitudes incorporate (1) beliefs regarding forecast accuracy, (2) beliefs about the compatibility with existing processes and workflows in an organization, (3) beliefs about error responsibility in case of a forecast error and (4) beliefs about the innovativeness and image of the organization among others. In total, every dimension of attitude relates to an own literature stream in information systems and psychology that have not yet been combined in prior research. First, forecast accuracy can be seen as a contextualization of the perceived usefulness construct from TAM and UTAUT (e.g., Venkatesh et al. 2003, Davis 1989). Second, the compatibility beliefs relate to the topic task-technology-fit and for example the research on the technology-to-performance chain (e.g., Fuller & Dennis 2009, Goodhue & Thompson 1995). Third, beliefs regarding error responsibility fit into the research on blame avoidance from the field of public management (Aschauer et al. 2024, Artinger et al. 2019, Weaver 1986). Fourth, beliefs regarding the innovativeness of an organization relate to research on corporate image and job satisfaction (e.g., Riordan et al. 1997) or employee commitment (e.g., Collier & Esteban 2007).

Lastly, this dissertation characterizes the complex interaction of PBC and attitudes when managers and other decision-makers have to decide between human and algorithmic judgment. Study 1 and study 3 show that the evaluation of PBC occurs before the formation of the attitudes towards human or algorithmic judgment. Study 2 further illustrates that the behavioral choice depends on the on the ratios of PBC and attitudes towards the judgment options instead of the absolute differences in cognitive factors (see study 2). This second characteristic is relevant in contexts, where PBC over human and algorithmic judgment is low or attitudes towards human or algorithmic judgment are negative, because the actual behavioral choice can be determined by very small absolute differences in cognitive factors (see Figure 3-2). Furthermore, this characteristic could explain some of the very strong empirical effects of the algorithm aversion literature (e.g., Berger et al. 2021, Dietvorst et al. 2015). Due to the dominance of experimental studies in the literature stream, most participants were confronted with artificial forecasting tasks and ambiguous forecasting tools, where PBC towards human and algorithmic judgment is likely low.[45] Following the developed framework in study 2, even a small decrease in attitude after experiencing the error-proneness of an algorithm can therefore be enough to reverse a choice from algorithmic to human judgment (e.g., as observed by Dietvorst et al. 2015).[46]

5.1.2 Findings on the second research question

The second overall research question states: "How can management accounting influence managers' use of predictive analytics in these tasks?" To answer this research questions, the following section draws on the management accounting artifacts described in section 1.3.2: applying XAI methods, solely communicating a generated forecast or communicating capability-based and predictability-based metrics. First, XAI methods provide a solid foundation to assist managers in planning and the necessary decision-making activities. In line with the propositions of prior research, global XAI methods, e.g., the examined ALE plots, indicate the underlying reasoning of a generated forecast and therefore convey credibility to decision-makers (e.g., Arrieta et al. 2020). In addition to this interpretability, study 1 shows these global XAI methods can also inform decision-makers about previously unknown relationships in the analyzed data sets. In the examined federal agency, for example, many interview partners were interested in the predicted influence

[45] For instance, recall that Dietvorst & Bharti (2020) require their participants to estimate the crop harvest and mining output on an alien planet.
[46] While Dietvorst et al. 2015 do not directly measure the attitudes of their participants, they incentivize them to provide accurate forecasts and do not provide other contexts. Their collected variables related to performance expectancy should therefore be suitable indicators for attitude (e.g., Ajzen & Fishbein 1969).

of age on gender on employee turnover.[47] However, our interview partners had significant doubts whether the use of local XAI methods, e.g., the examined SHAP values, can legally be used to analyze HR data. Considering that respective methods fall under the General Data Protection Regulation (GDPR) and AI Act of the European Union (EU) (Kusche 2024, Albrecht 2016), their use in personnel decisions seems highly unlikely in European countries.[48]

Second, it follows that solely communicating a forecast generated by predictive analytics to management can trigger algorithm aversion and the approach can therefore be seen as ineffective (see study 2).[49] In particular, with over 50 identified antecedents of algorithm aversion (e.g., Mahmud et al. 2022), it is not feasible to consider all of them in the communication with management. Furthermore, a respective communication approach bears the risk of decreasing the long-term success of predictive analytics in an organization. Among others, prior research identifies motivators and leaders as important forces that determine the actual potential for change in a management accounting department (e.g., Kasurinen 2002). If a respective communication strategy triggers algorithm aversion in managers, they will conceptually prefer existing sources of advice to this new sort of advice (see study 2). In turn, this likely decreases their support of a new predictive analytics tool (leaders) and management accountings motivation to use that tool. In case the sole communication of a forecast cannot be avoided, management accountants must be aware that the likelihood of algorithm aversion depends on the aforementioned ratios of attitudes and PBC and that an effective communication strategy must address these ratios (see study 2).

Third, communicating capability-based and predictability-based metrics leads to intentional and unintentional changes in managers' use of predictive analytics tools. In case of capability-based metrics, study 3 shows that a respective metric improves managers' perceived credibility of a respective tool no matter if justified or unjustified (see study 3). The observed effect of communicating a respective metric further validates prior research on effective mitigation strategies for algorithm aversion (e.g., Reich et al. 2023, Berger et al. 2021, Castelo et al. 2019, Dietvorst et al. 2015). However, study 3 also shows that

[47] Note that for the given HR context, this interest in age and gender bears the risk of unthinking and potentially biased decision making (e.g., Newman et al. 2020).
[48] Note that there are nevertheless reasonable effective use cases of local XAI methods, for example to examine the likelihood of material restatements (Zhang et al. 2022).
[49] Unless, of course, management accounting wants to reduce the use of these tools by management.

communicating a respective metric also increases the perceived credibility of a forecast in unjustified contexts. This finding questions the assumption of prior research that increasing the use of predictive analytics tools is always desirable (e.g., Greiner et al. 2025, Dietvorst et al. 2015). In case of predictability-based metrics, study 3 shows that a respective metric calibrates managers' perceived credibility of a respective tool in accordance to its actual credibility and is therefore superior to a capability-based metric (see study 3). Note that respective predictability-based metrics draw on knowledge from physics and mathematics (e.g., Paialunga 2023, Flood & Grimm 2021, Salvino et al. 1995) and we are among the first to transfer this knowledge into an accounting and business context.

In total, Figure 5-1 illustrates and summarizes the contributions of the three studies included in this dissertation to the overarching research questions.

	RQ1: What are relevant factors that encourage or deter managers from incorporating predictive analytics tools into planning activities?	RQ2: How can management accounting influence managers' use of predictive analytics in these tasks?
Study 1: Exploring the Individual Adoption of Human Resource Analytics - Behavioral Beliefs and the Role of Machine Learning Characteristics	• PBC-related beliefs: capabilities to use a tool, tool's forecasting capabilities, quality of the available datasets. • Attitude-related beliefs: tool accuracy, tool compatibility, error responsibility, image. • (Norm-related beliefs: culture, legal framework)	• XAI methods do not only convey transparency but also inform decision-makers about unknown and potentially important relationships in the data. • However, XAI methods do not induce considerations regarding the fairness of HR forecasts.
Study 2: A Comprehensive Framework for Algorithm Aversion in Business Contexts	• The choice between algorithmic and human judgment is determined by the relative ratio of attitudes towards algorithmic and human judgement compared to the relative ratio of attitudes towards algorithmic and human judgement.	• The sole communication of a generated forecast can trigger algorithm aversion. • Management accounting must address ratios of attitudes towards and PBC over algorithmic and human judgement.
Study 3: Balancing Confidence and Skepticism in Predictive Forecasting - The Role of Perceived Credibility	• The perceived capabilities of a tool are a strong predictor of perceived credibility (and later use). • Decision makers erroneously associate forecasting capabilities with accurate forecasts.	• Capability-based metrics always improve the perceived credibility of a provided, even when this thought is not justified. • Predictability-based-metrics calibrate the perceived credibility of a provided forecast in accordance with the likely forecast accuracy.

Figure 5-1: Research contributions of this dissertation split into studies and research questions.[50]

5.1.3 Implications for corporate practice and education

The overall theoretical contributions discussed in this section have additional implications for corporate practice and education, three of which are discussed in the following section. First, the effective use of predictive analytics tools requires comprehensive decision support for managers, which exceeds the simple provision of generated forecasts.

[50] Note that because only Study 1 considered the effects of social norms, these results were not discussed in the previous paragraphs. Instead, a discussion of these factors is included in the next section.

Drawing on the findings of all three studies, this decision support should address biased perceptions of the forecasting capabilities of a predictive analytics tool, which range from underestimation to overestimation combined with subsequent disappointment. Study 1 and study 3 characterize two feasible approaches for the respective and required decision-support. On the one hand, XAI methods allow to illustrate and characterize the underlying reasons for a ML-based and often opaque forecast. At least in the examined context of personnel planning, global methods appear to be superior to local methods. On the other hand, predictability-based metrics inform management about the fundamental predictability of a time series, or, in other words, the difficulty of forecasting tasks, and calibrate their perceived credibility of a provided forecast. While study 3 examines the provision of this information in the context of algorithmically generated forecasts, time series predictability is defined independent from the applied forecasting method (e.g., Salvino et al. 1995). This information could therefore be interesting for the evaluation of forecasts from a human expert too.

Second, to effectively assist managers in interpreting the forecasts generated by predictive analytics tools, management accounting should acquire sound knowledge on these tools. While this dissertation characterizes two different approaches for respective decision support, these approaches nevertheless require a certain degree of technical understanding. For instance, the examined ALE-plots in study 1 can easily be misunderstood as factors from a multiple linear regression model (see Figure 2-1) and effective decision support must correct the managers falling for this misinterpretation. Moreover, disregarding predictive analytics and refusing to acquire the necessary competencies can decrease the importance of management accounting in the long term. Considering the aforementioned calculation speed and achievable accuracy with respective tools, it appears unlikely that corporations forego the competitive advantage (e.g., Chen et al. 2022b, Vasarhelyi et al. 2015). Instead, areas of responsibility might shift from the management accounting department to other departments, for example a new data science department (e.g., Pabinger et al. 2021, Freistühler et al. 2019, Steiner & Welker 2016). As the acquisition of new knowledge likely involves the hiring of new employees (Cirillo et al. 2024), this shift in the necessary competences of management accounting provides an opportunity for university graduates to specializing in respective topics.

Third, university education must acknowledge the digital transformation of management accounting practice and adjust their teaching programs.[51] Hereby, an effective approach could be to introduce more case studies and assignments into the curriculum. Compared to the European accounting courses, for example the business schools in the United States follow a more hands-on teaching approach and use case studies to confront their students with predictive analytics and other IT-tools (e.g., Nickell et al. 2023, Parlier & Lee 2023, Polimeni & Burke 2021, Qasim et al. 2020). In total there are three academic journals, the "Journal of Emerging Technologies in Accounting", the "Journal of Accounting Education" and "Issues in Accounting Education", which publish respective case studies and also provide sample solutions to lecturers. However, respective changes in the management accounting curriculum require lecturers, which have experience with these IT-systems. Therefore, faculties must either consider hiring lecturers specializing in AIS and other IT-systems or finding suitable and reliable partners in corporate practice supervising the new assignments.

5.2 Limitations and further research

Although this dissertation makes several contributions to research and corporate practice, every form of research is subject to limitations. The following section therefore discusses the methodological and conceptual limitations of the included studies in this dissertation in answering the overall research questions. Furthermore, it characterizes avenues for further research on these overall research questions and beyond.

5.2.1 External validity and potential moderator variables

While this dissertation provides several approaches for the steering of managers' use of predictive analytics in planning, the utilized research methods are all limited by a low external validity.[52] Study 1 and study 3 draw on an in-depth case or a laboratory setting. Both of the latter research methods have been criticized for their lacking ability to identify all background factors that influence an effect or behavior in practice (Quintão et al. 2020, Lynch 1982). The external validity of study 2 is difficult to deduce, because no primary data was collected and the reasoning consequently depends on the external validity of the cited studies (Lynch 1999). It is therefore ambiguous whether this dissertation identifies

[51] Note that the following paragraph focuses on the German education system at public universities, which the author of this dissertation is very familiar with.
[52] For a definition, see for example Döring (2023).

all moderator variables that influence managers' use of predictive analytics in planning. This limitation of the included studies in this dissertation provides an avenue for further research. In general, the proposed approaches in this dissertation could be tested with research methods that possess a high external validity, for instance large-scale surveys or qualitative expert interviews (e.g., Döring 2023).

In particular, examining the effects of social norms on managers' use of predictive analytics tools seems fruitful. Recall that the TPB predicts a moderation effect of PBC on the relationship between social norms and behavioral intention, where the direct effect of norms is greater with less PBC (La Barbera & Ajzen 2020) and that study 1 observed norm-driven opinions. Social norms could therefore be an important factor that convince or deter managers from using predictive analytics tools in their implementation phase, where PBC regarding the new tools is low. Respective research would be in line with the idea of steering an organization with the help of cultural controls (e.g., Malmi & Brown 2008). Moreover, while these social norms are likely related to the culture of an organization and legal guidelines (e.g., see study 1), they likely also depend on the culture of the underlying country (e.g., Hofstede 2011). For instance, Castelo 2024 shows that algorithm aversion is significantly less common for people, which were raised in countries with high levels of corruption. Moreover, Schepers & Wetzels (2007) unexpectedly find social norms being less important for technology acceptance in non-western and collectivistic cultures than in western and individualistic ones.

5.2.2 Long-term effects of algorithmic decision support

As longitudinal studies exceed the possible timeframe of a dissertation, it does not examine potential long-term effects of the provision of data analytics to managements quality of decision-making. However, the aforementioned Theory of Technology Dominance predicts that users of technology can lose the capabilities required for certain tasks, when these are automated by this technology, which is called "deskilling". (Sutton et al. 2023, Arnold & Sutton 1998). Deskilling can be problematic for the given context, because predictive analytics tools are always error-prone and detecting these errors requires human expertise (De-Arteaga et al. 2020). Moreover, deskilling effects exceed the scope of the developed calibration approach for perceived tool credibility in study 3. Recall that informing managers about the predictability of a time series informs them about the likelihood of unavoidable forecast errors and not about actually erroneous forecasts.

Consequently, detecting these errors will still require human forecasting expertise. While one could argue that deskilling is not relevant for managers, which were always supported by specialist and never acquired forecasting skills, predictive analytics will likely deskill these specialists and decrease their capabilities to provide helpful decision-support.

To better understand respective long-term effects, future research could specifically examine setting in which predictive analytics tools are well-established. A fitting example is the use of audience analytics by newspapers and television broadcasters, where these tools inform editorial decision-making regarding the behavior and preferences of their audience (e.g., Lee & Peng 2024, Cherubini & Nielsen 2016). While respective tools have become common in newsrooms and are often used by journalists (Lamot & Paulussen 2020), there is an ongoing debate about the consequences of respective tools for the profession. One the one hand, audience analytics help journalists in better understanding their consumers wishes and to optimize their news accordingly (e.g., Belair-Gagnon et al. 2020). On the other hand, there are ethical concerns that audience analytics draw away the focus from balanced news coverage and remove editors independence (e.g., Dodds et al. 2023, Tandoc & Thomas 2015). Considering the previous argument of deskilling, the journalistic context allows to examine whether editors and other decision-makers are still able to identify the relevant information for the general public interest.

5.2.3 Additional avenues for further research

This dissertation provides several other avenues for further research. For instance, analytical modelling of algorithm aversion and similar behaviors still appears to be fruitful. Although study 2 changed from an analytical paper to a conceptual paper during the research project, this change was motivated by the conventions in accounting research less by content-related reasons. In particular, prior research on algorithm aversion identifies a growing number of non-linear relationships that exceed human comprehension. For instance, Allen & Choudhury (2022) find that the performance increase from using an algorithmic decision aid follows an inverted U-shape over the range of domain expertise, where decision-makers with a medium level of expertise benefit the most from the offered aid. Dietvorst & Bharti (2020) show that decision-makers have a diminishing sensitivity to forecasting errors from a predictive analytics tool, where they penalize each additional margin of error less than the one before. While analytical research in accounting mostly centers around incentivization and incentive systems (e.g., Dikolli et al. 2013, Heinle et al. 2012), there are already some analytical studies on algorithm aversion in the sub discipline of mathematical psychology (Sinclair-Desgagné 2024, Kumar et al. 2021) calling for further analytical research.

Furthermore, management accounting evidently must embrace predictive analytics tools itself to be able to effectively assist managers in their use of respective tools and their outputs. However, prior research in accounting raises doubts whether management accountants are open to do so. First, management accounting is generally known for stable routines and a resistance to change (e.g., Granlund 2001). For instance, this stability shows in the preference of management accountants for spreadsheet software (Church et al. 2022, Schmidt et al. 2020b). Second, management accounting has been reluctant to use advanced functionalities of ERP systems during their implementation phase 20 years ago (e.g., Hyvönen et al. 2009). Third, more recent literature suggests that management accountants working with data analytics tools still mostly focused on the descriptive functionalities of these tools (e.g., Spraakman et al. 2021). These findings of prior research motivate another research project of the author with the working title "Stuck in the Spreadsheet Era - An Experimental Investigation of Accountants' Resistance to Abandon Microsoft Excel", which is a collaboration with Alina Bieniek from the Chair of Management Accounting and Control at TU Dortmund University. The project is still in an earlier stage and the data collection is scheduled for autumn/winter 2025. However, we already presented the hypotheses and planned experimental design in an early-stage research session at the Joint Midyear Meeting of the AIS/SET sections of the AAA 2025.

6. Conclusion

This dissertation provides a comprehensive perspective on the factors that encourage or deter managers from incorporating predictive analytics tools into planning activities and illustrates strategies for management accounting to influence managers' use of these tools. It is motivated by the emergence of predictive analytics tools in corporate practice and the insufficient accounting-based research on this topic, in particular regarding managers' perception and use of these new and alternative sources of advice (see section 1.1). This dissertation includes three different research studies, which addresses three different sub-topics (see section 1.2) and hereby examines different artifacts for decision-support (see section 1.3). All three included studies rely on different research methods: While study 1 is a qualitative in-depth case study in a German federal agency, study 2 is conceptual and provides a framework to structure und better understand existing research findings and study 3 is a behavioral laboratory experiment that was conducted with participants from the online labor platform Prolific.

Overall, the three studies identify significant biases in managers perception of predictive analytics tools and their outputs. A central misperception resolves around the actual capabilities of respective tools, where managers tend to either underestimate or overestimate these capabilities based on subjective criteria. Moreover, this dissertation identifies numerous attitudes regarding predictive analytics that influence the use of these tools (see section 5.1.1). Furthermore, the included studies in this dissertation identify XAI methods and predictability-based metrics as effective approaches to influence managers' use of predictive analytics tools in a beneficial way. However, the included studies also illustrate the challenges that arise from solely communicating an algorithmically generate forecast to management (see section 5.1.2). Based on these contributions, the dissertation provides additional and more operational recommendations for corporate practice and university education (see section 5.1.3). Lastly, it also discusses limitations of the included studies and provides avenues for further research on the general topic (see section 5.2).

This dissertation closes with the quotation from the beginning and also concludes that "control cannot be studied apart from technology and context." (cf. Dechow & Mouritsen 2005, p. 691). Although research on predictive analytics and similar topics in prime accounting journals is still rare, the ever accelerating technological progress and the enormous success of GenAI provide hope that this will change in the foreseeable future.

References

Abbas, K. (2025). Management Accounting and Artificial Intelligence: A Comprehensive Literature Review and Recommendations for Future Research. *The British Accounting Review,* 101551. https://doi.org/10.1016/j.bar.2025.101551.

Adam, M., Diebel, C., Goutier, M., & Benlian, A. (2024). Navigating autonomy and control in human-AI delegation: User responses to technology- versus user-invoked task allocation. *Decision Support Systems*, *180,* 114193. https://doi.org/10.1016/j.dss.2024.114193.

Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl, & J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11–39). Berlin, Germany: Springer.

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-t.

Ajzen, I. (2002). Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior. *Journal of Applied Social Psychology*, *32*(4), 665–683. https://doi.org/10.1111/j.1559-1816.2002.tb00236.x.

Ajzen, I. (2020). The Theory of Planned Behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, *2*(4), 314–324. https://doi.org/10.1002/hbe2.195.

Ajzen, I., & Fishbein, M. (1969). The Prediction of Behavioral Intentions in a Choice Situation. *Journal of Experimental Social Psychology*, *5*(4), 400–416. https://doi.org/10.1016/0022-1031(69)90033-X.

Albrecht, J. P. (2016). How the GDPR Will Change the World. *European Data Protection Law Review*, *2*(3), 287–289. https://doi.org/10.21552/edpl/2016/3/4.

Al-Htaybat, K., & Alberti-Alhtaybat, L. von (2017). Big Data and Corporate Reporting: Impacts and Paradoxes. *Accounting, Auditing & Accountability Journal*, *30*(4), 850–873. https://doi.org/10.1108/AAAJ-07-2015-2139.

Allen, R., & Choudhury, P. (2022). Algorithm-Augmented Work and Domain Experience: The Countervailing Forces of Ability and Aversion. *Organization Science*, *33*(1), 149–169. https://doi.org/10.1287/ORSC.2021.1554.

Alon-Barkat, S., & Busuioc, M. (2023). Human–AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice. *Journal of Public Administration Research and Theory*, *33*(1), 153–169. https://doi.org/10.1093/jopart/muac007.

Alpar, P., & Schulz, M. (2016). Self-Service Business Intelligence. *Business & Information Systems Engineering*, *58*(2), 151–155. https://doi.org/10.1007/s12599-016-0424-6.

Andreassen, P. B., & Kraus, S. J. (1990). Judgmental Extrapolation and the Salience of Change. *Journal of Forecasting*, *9*(4), 347–372. https://doi.org/10.1002/for.3980090405.

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and Analytics: Why HR is Set to Fail the Big Data Challenge. *Human Resource Management Journal*, *26*(1), 1–11. https://doi.org/10.1111/1748-8583.12090.

Apley, D. W., & Zhu, J. (2020). Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(4), 1059–1086. https://doi.org/10.1111/rssb.12377.

Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of Business Analytics and Enterprise Systems on Managerial Accounting. *International Journal of Accounting Information Systems*, *25*, 29–44. https://doi.org/10.1016/j.accinf.2017.03.003.

Arnold, V., Collier, P. A., Leech, S. A., & Sutton, S. G. (2004). Impact of Intelligent Decision Aids on Expert and Novice Decision-Makers' Judgments. *Accounting & Finance*, *44*(1), 1–26. https://doi.org/10.1111/j.1467-629x.2004.00099.x.

Arnold, V., & Sutton, S. G. (1998). The Theory of Technology Dominance: Understanding the Impact of Intelligent Decision Aids on Decision Makers' Judgments. *Advances in Accounting Behavioral Research*, *1*, 175–194. https://www.researchgate.net/publication/284107668 _The_theory_of_technology_dominance_Understanding_the_impact_of_intelligent_decision_aids_on_decision_makers%27_judgments.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Artinger, F. M., Artinger, S., & Gigerenzer, G. (2019). C. Y. A.: Frequency and Causes of Defensive Decisions in Public Administration. *Business Research*, *12*(1), 9–25. https://doi.org/10.1007/s40685-018-0074-2.

Aschauer, F., Sohn, M., & Hirsch, B. (2024). Managerial Advice-Taking—Sharing Responsibility with (Non)Human Advisors trumps Decision Accuracy. *European Management Review*, *21*(1), 186–203. https://doi.org/10.1111/emre.12575.

Ashton, R. H. (1990). Pressure and Performance in Accounting Decision Settings: Paradoxical Effects of Incentives, Feedback, and Justification. *Journal of Accounting Research*, *28*, 148. https://doi.org/10.2307/2491253.

Austin, A. A., Carpenter, T. D., Christ, M. H., & Nielson, C. S. (2021). The Data Analytics Journey: Interactions Among Auditors, Managers, Regulation, and Technology. *Contemporary Accounting Research*, *38*(3), 1888–1924. https://doi.org/10.1111/1911-3846.12680.

Bai, X., Nunez, M., & Kalagnanam, J. R. (2012). Managing Data Quality Risk in Accounting Information Systems. *Information Systems Research*, *23*(2), 453–473. https://doi.org/10.1287/isre.1110.0371.

Bandura, A. (1977). Self-Efficacy: toward a Unifying Theory of Behavioral Change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037//0033-295x.84.2.191.

Basu, S., Majumdar, B., Mukherjee, K., Munjal, S., & Palaksha, C. (2023). Artificial Intelligence–HRM Interactions and Outcomes: A Systematic Review and Causal Configurational Explanation. *Human Resource Management Review*, *33*(1), 100893. https://doi.org/10.1016/j.hrmr.2022.100893.

Belair-Gagnon, V., Zamith, R., & Holton, A. E. (2020). Role Orientations and Audience Metrics in Newsrooms: An Examination of Journalistic Perceptions and their Drivers. *Digital Journalism*, *8*(3), 347–366. https://doi.org/10.1080/21670811.2019.1709521.

Berger, B., Adam, M., Rühr, A., & Benlian, A. (2021). Watch me improve—Algorithm aversion and demonstrating the ability to learn. *Business and Information Systems Engineering*, *63*(1), 55–68. https://doi.org/10.1007/s12599-020-00678-5.

Bertsimas, D., & Kallus, N. (2020). From Predictive to Prescriptive Analytics. *Management Science*, *66*(3), 1025–1044. https://doi.org/10.1287/mnsc.2018.3253.

Bezemer, D. J. (2010). Understanding Financial Crisis Through Accounting Models. *Accounting, Organizations and Society*, *35*(7), 676–688. https://doi.org/10.1016/j.aos.2010.07.002.

Bhimani, A., & Willcocks, L. (2014). Digitisation, 'Big Data' and the Transformation of Accounting Information. *Accounting and Business Research*, *44*(4), 469–490. https://doi.org/10.1080/00014788.2014.910051.

Bigman, Y. E., Yam, K. C., Marciano, D., Reynolds, S. J., & Gray, K. (2021). Threat of Racial and Economic Inequality increases Preference for Algorithm Decision-Making. *Computers in Human Behavior*, *122,* 106859. https://doi.org/10.1016/j.chb.2021.106859.

Birnbaum, M. H., & Stegner, S. E. (1979). Source Credibility in Social Judgment: Bias, Expertise, and the Judge's Point of View. *Journal of Personality and Social Psychology*, *37*(1), 48–74. https://doi.org/10.1037/0022-3514.37.1.48.

Bogert, E., Schecter, A., & Watson, R. T. (2021). Humans Rely More on Algorithms Than Social Influence as a Task Becomes More Difficult. *Scientific Reports*, *11*(1), 8028. https://doi.org/10.1038/s41598-021-87480-9.

Böhmer, N., & Schinnenburg, H. (2023). Critical Exploration of AI-Driven HRM to Build Up Organizational Capabilities. *Employee Relations: The International Journal*, *45*(5), 1057–1082. https://doi.org/10.1108/ER-04-2022-0202.

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.

Bonezzi, A., & Ostinelli, M. (2021). Can Algorithms Legitimize Discrimination? *Journal of Experimental Psychology: Applied*, *27*(2), 447–459. https://doi.org/10.1037/xap0000294.

Bonezzi, A., Ostinelli, M., & Melzner, J. (2022). The Human Black-Box: The Illusion of Understanding Human better than Algorithmic Decision-Making. *Journal of Experimental Psychology: General*, *151*(9), 2250–2258. https://doi.org/10.1037/xge0001181.

Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most Human Bot: Female Gendering increases Humanness Perceptions of Bots and Acceptance of AI. *Psychology and Marketing*, *38*(7), 1052–1068. https://doi.org/10.1002/mar.21480.

Boritz, J. E., & Stratopoulos, T. C. (2023). AI and the Accounting Profession: Views from Industry and Academia. *Journal of Information Systems*, *37*(3), 1–9. https://doi.org/10.2308/ISYS-2023-054.

Bratten, B., Gaynor, L. M., McDaniel, L., Montague, N. R., & Sierra, G. E. (2013). The Audit of Fair Values and Other Estimates: The Effects of Underlying Environmental, Task, and Auditor-Specific Factors. *Auditing -A Journal of Practice & Theory*, *32*(Supplement 1), 7–44. https://doi.org/10.2308/ajpt-50316.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brüggen, A., Grabner, I., & Sedatole, K. L. (2021). The Folly of Forecasting: The Effects of a Disaggregated Demand Forecasting System on Forecast Error, Forecast Positive Bias, and Inventory Levels. *The Accounting Review*, *96*(2), 127–152. https://doi.org/10.2308/tar-2018-0559.

Brüns, J. D., & Meißner, M. (2024). Do You Create Your Content Yourself? Using Generative Artificial Intelligence for Social Media Content Creation diminishes

Perceived Brand Authenticity. *Journal of Retailing and Consumer Services*, *79*, 103790. https://doi.org/10.1016/j.jretconser.2024.103790.

Brynjolfsson, E., Hitt, L., & Kim, H. (2011). Strength in Numbers: How Does Data-Driven Decision-Making Affect Firm Performance? *ICIS 2011 Proceedings*. https://aisel.aisnet.org/icis2011/proceedings/economicvalueIS/13.

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

Burton, J. W., Stein, M.-K., & Jensen, T. B. (2020). A Systematic Review of Algorithm Aversion in augmented Decision Making. *Journal of Behavioral Decision Making*, *33*(2), 220–239. https://doi.org/10.1002/bdm.2155.

Busuioc, M. (2021). Accountable Artificial Intelligence: Holding Algorithms to Account. *Public administration review*, *81*(5), 825–836. https://doi.org/10.1111/puar.13293.

Candrian, C., & Scherer, A. (2024). How Terminology Affects Users' Responses to System Failures. *Human Factors*, *66*(8), 2082–2103. https://doi.org/10.1177/00187208231202572.

Carmeli, A., & Weisberg, J. (2006). Exploring Turnover Intentions Among Three Professional Groups of Employees. *Human Resource Development International*, *9*(2), 191–206. https://doi.org/10.1080/13678860600616305.

Casas-Arce, P., Cheng, M. M., Grabner, I., & Modell, S. (2022). Managerial Accounting for Decision-Making and Planning. *Journal of Management Accounting Research*, *34*(1), 1–7. https://doi.org/10.2308/jmar-10784.

Casas-Arce, P., Martnez-Jerez, F. A., & Narayanan, V. G. (2017). The Impact of Forward-Looking Metrics on Employee Decision-Making: The Case of Customer Lifetime Value. *The Accounting Review*, *92*(3), 31–56. https://doi.org/10.2308/accr-51554.

Castelo, N. (2024). Perceived Corruption Reduces Algorithm Aversion. *Journal of Consumer Psychology*, *34*(2), 326–333. https://doi.org/10.1002/jcpy.1373.

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788.

Chacon, A., Kausel, E. E., & Reyes, T. (2022). A Longitudinal Approach for Understanding Algorithm Use. *Journal of Behavioral Decision Making*, *35*(4), e2275. https://doi.org/10.1002/bdm.2275.

Chang, Y., & Wang, R. (2023). Conservatives Endorse Fintech? Individual Regulatory Focus Attenuates the Algorithm Aversion Effects in Automated Wealth Management. *Computers in Human Behavior*, *148*, 107872. https://doi.org/10.1016/j.chb.2023.107872.

Chatterjee, S., Rana, N. P., Dwivedi, Y. K., & Baabdullah, A. M. (2021). Understanding AI Adoption in Manufacturing and Production Firms Using an Integrated TAM-TOE Model. *Technological Forecasting and Social Change*, *170*, 120880. https://doi.org/10.1016/j.techfore.2021.120880.

Chen, C. X., Hudgins, R., & Wright, W. F. (2022a). The Effect of Advice Valence on the Perceived Credibility of Data Analytics. *Journal of Management Accounting Research*, *34*(2), 97–116. https://doi.org/10.2308/JMAR-2020-015.

Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine Learning Model for Bitcoin Exchange Rate Prediction Using Economic and Technology Determinants. *International Journal of Forecasting*, *37*(1), 28–43. https://doi.org/10.1016/j.ijforecast.2020.02.008.

Chen, X. I., Yang, H. A., Dou, Y., & Lev, B. (2022b). Predicting Future Earnings Changes Using Machine Learning and Detailed Financial Data. *Journal of Accounting Research*, *60*(2), 467–515. https://doi.org/10.1111/1475-679X.12429.

Cheng, L., & Chouldechova, A. (2023). Overcoming Algorithm Aversion: A Comparison between Process and Outcome Control. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems 2023*, 1–27. https://doi.org/10.1145/3544548.3581253.

Cherubini, F., & Nielsen, R. K. (2016). *Editorial Analytics: How News Media are Developing and Using Audience Data and Metrics*. Oxford: Reuters Institute.

Chharia, A., Jeevan, G., Jha, R. A., Liu, M., Berman, J. M., & Glorioso, C. (2024). Accuracy of US CDC COVID-19 Forecasting Models. *Frontiers in Public Health*, *12,* 1359368. https://doi.org/10.3389/fpubh.2024.1359368.

Choudhury, N. (1986). In Search of Relevance in Management Accounting Research. *Accounting and Business Research*, *17*(65), 21–32. https://doi.org/10.1080/00014788.1986.9729778.

Choudhury, P., Allen, R. T., & Endres, M. G. (2021). Machine Learning for Pattern Discovery in Management Research. *Strategic Management Journal*, *42*(1), 30–57. https://doi.org/10.1002/smj.3215.

Choudhury, P., Starr, E., & Agarwal, R. (2020). Machine Learning and Human Capital Complementarities: Experimental Evidence on Bias Mitigation. *Strategic Management Journal*, *41*(8), 1381–1411. https://doi.org/10.1002/smj.3152.

Chowdhury, S., Joel-Edgar, S., Dey, P. K., Bhattacharya, S., & Kharlamov, A. (2022). Embedding Transparency in Artificial Intelligence Machine Learning Models: Managerial Implications on Predicting and Explaining Employee Turnover. *The International Journal of Human Resource Management*, 1–32. https://doi.org/10.1080/09585192.2022.2066981.

Church, K. S., Riley, J., & Schmidt, P. J. (2022). Has Excel Become a "Golden Hammer": The Paradox of Data Analytics in SME Clusters. *Journal of Emerging Technologies in Accounting*, *19*(2), 211–234. https://doi.org/10.2308/JETA-2020-051.

Cirillo, V., Mina, A., & Ricci, A. (2024). Digital Technologies, Labor Market Flows and Training: Evidence from Italian Employer-Employee Data. *Technological Forecasting and Social Change*, *209*, 123735. https://doi.org/10.1016/j.techfore.2024.123735.

Claudy, M. C., Aquino, K., & Graso, M. (2022). Artificial Intelligence Can't Be Charmed: The Effects of Impartiality on Laypeople's Algorithmic Preferences. *Frontiers in Psychology*, *13,* 898027. https://doi.org/10.3389/fpsyg.2022.898027.

Collier, J., & Esteban, R. (2007). Corporate Social Responsibility and Employee Commitment. *Business Ethics: A European Review*, *16*(1), 19–33. https://doi.org/10.1111/j.1467-8608.2006.00466.x.

Commerford, B. P., Dennis, S. A., Joe, J. R., & Ulla, J. W. (2022). Man Versus Machine: Complex Estimates and Auditor Reliance on Artificial Intelligence. *Journal of Accounting Research*, *60*(1), 171–201. https://doi.org/10.1111/1475-679X.12407.

Commerford, B. P., Eilifsen, A., Hatfield, R. C., Holmstrom, K. M., & Kinserdal, F. (2024). Control Issues: How providing Input affects Auditors' Reliance on Artificial Intelligence. *Contemporary Accounting Research*, *41*(4), 2134–2162, 1911-3846.12974. https://doi.org/10.1111/1911-3846.12974.

Coolen, P., van den Heuvel, S., van de Voorde, K., & Paauwe, J. (2023). Understanding the Adoption and Institutionalization of Workforce Analytics: A Systematic Literature Review and Research Agenda. *Human Resource Management Review*(100985). https://doi.org/10.1016/j.hrmr.2023.100985.

Coyne, J. G., Summers, S. L., Williams, B., & Wood, D. A. (2010). Accounting Program Research Rankings by Topical Area and Methodology. *Issues in Accounting Education*, *25*(4), 631–654. https://doi.org/10.2308/iace.2010.25.4.631.

Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2018). The Operational Value of Social Media Information. *Production and Operations Management*, *27*(10), 1749–1769. https://doi.org/10.1111/poms.12707.

Daschner, S., & Obermaier, R. (2022). Algorithm Aversion? On the Influence of Advice Accuracy on Trust in Algorithmic Advice. *Journal of Decision Systems*, *31*(S1), 77–97. https://doi.org/10.1080/12460125.2022.2070951.

Davenport, T. H., Harris, J., & Shapiro, J. (2010). Competing on Talent Analytics. *Harvard Business Review*, *88*(10), 52-8, 150. https://www.researchgate.net/publication/47369355_Competing_on_talent_analytics.

Davis, F. D. (1985). *A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results*. Massachusetts Institute of Technology.

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.

Dawes, R. M. (1979). The Robust Beauty of Improper Linear Models in Decision Making. *American Psychologist*, *34*(7), 571–582. https://doi.org/10.1037/0003-066X.34.7.571.

De-Arteaga, M., Fogliato, R., & Chouldechova, A. (2020). A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems 2020*, 1–12. https://doi.org/10.1145/3313831.3376638.

Dechow, N., & Mouritsen, J. (2005). Enterprise Resource Planning Systems, Management Control and the Quest for Integration. *Accounting, Organizations and Society*, *30*(7), 691–733. https://doi.org/10.1016/j.aos.2004.11.004.

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid Intelligence. *Business and Information Systems Engineering*, *61*(5), 637–643. https://doi.org/10.1007/s12599-019-00595-2.

Desouza, K. C., Dawson, G. S., & Chenok, D. (2020). Designing, Developing, and Deploying Artificial Intelligence Systems: Lessons from and for the Public Sector. *Business Horizons*, *63*(2), 205–213. https://doi.org/10.1016/j.bushor.2019.11.004.

Di Vaio, A., Hassan, R., & Alavoine, C. (2022). Data Intelligence and Analytics: A Bibliometric Analysis of Human–Artificial Intelligence in Public Sector Decision-Making Effectiveness. *Technological Forecasting and Social Change*, *174*(121201). https://doi.org/10.1016/j.techfore.2021.121201.

Dietvorst, B. J., & Bharti, S. (2020). People Reject Algorithms in Uncertain Decision Domains Because They Have Diminishing Sensitivity to Forecasting Error. *Psychological Science*, *31*(10), 1302–1314. https://doi.org/10.1177/0956797620948841.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Er. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming Algorithm Aversion: People will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science*, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643.

Dikolli, S. S., Hofmann, C., & Pfeiffer, T. (2013). Relative Performance Evaluation and Peer-Performance Summarization Errors. *Review of Accounting Studies*, *18*(1), 34–65. https://doi.org/10.1007/s11142-012-9212-9.

Dodds, T., Vreese, C. de, Helberger, N., Resendez, V., & Seipp, T. (2023). Popularity-driven Metrics: Audience Analytics and Shifting Opinion Power to Digital Platforms. *Journalism Studies*, *24*(3), 403–421. https://doi.org/10.1080/1461670X.2023.2167104.

Donihue, M. R. (1993). Evaluating the Role Judgment Plays in Forecast Accuracy. *Journal of Forecasting*, *12*(2), 81–92. https://doi.org/10.1002/for.3980120203.

Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Downen, T., Kim, S., & Lee, L. (2024). Algorithm Aversion, Emotions, and Investor Reaction: Does disclosing the use of AI influence Investment Decisions? *International Journal of Accounting Information Systems*, *52*, 100664. https://doi.org/10.1016/j.accinf.2023.100664.

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., & Ranjan, R. (2023). Explainable AI (XAI): Core Ideas, Techniques, and Solutions. *ACM Computing Surveys*, *55*(9), 1–33. https://doi.org/10.1145/3561048.

EAA Congress (2024). Submission/Review Process - EAA Congress 2024. https://eaa-online.org/congress-2024/submission-review-process/. Accessed 24.08.2025.

Edwards, M. R., Charlwood, A., Guenole, N., & Marler, J. (2022). HR Analytics: An Emerging Field finding its Place in the World Alongside Simmering Ethical Challenges. *Human Resource Management Journal,* 1748-8583.12435. https://doi.org/10.1111/1748-8583.12435.

Efendi, J., Mulig, E. V., & Smith, L. M. (2006). Information Technology and Systems Research Published in Major Accounting Academic and Professional Journals. *Journal of Emerging Technologies in Accounting*, *3*(1), 117–128. https://doi.org/10.2308/jeta.2006.3.1.117.

Efendić, E., van de Calseyde, P., & Evans, A. M. (2020). Slow Response Times Undermine Trust in Algorithmic (but not Human) Predictions. *Organizational Behavior and Human Decision Processes*, *157*, 103–114. https://doi.org/10.1016/j.obhdp.2020.01.008.

Einola, K., & Khoreva, V. (2023). Best Friend or Broken Tool? Exploring the Co-Existence of Humans and Artificial Intelligence in the Workplace Ecosystem. *Human Resource Management*, *62*(1), 117–135. https://doi.org/10.1002/hrm.22147.

Ellmer, M., & Reichel, A. (2021). Staying Close to Business: the Role of Epistemic Alignment in Rendering HR Analytics Outputs Relevant to Decision-Makers. *The International Journal of Human Resource Management*, *32*(12), 2622–2642. https://doi.org/10.1080/09585192.2021.1886148.

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543.

Ensafi, Y., Amin, S. H., Zhang, G., & Shah, B. (2022). Time-Series Forecasting of Seasonal Items Sales Using Machine Learning – A Comparative Analysis. *International Journal of Information Management Data Insights*, *2*(1), 100058. https://doi.org/10.1016/j.jjimei.2022.100058.

Fehrenbacher, D. D., Ghio, A., & Weisner, M. (2023). Advice Utilization From Predictive Analytics Tools: The Trend is Your Friend. *European Accounting Review*, *32*(3), 637–662. https://doi.org/10.1080/09638180.2022.2138934.

Feiler, D., & Tong, J. (2022). From Noise to Bias: Overconfidence in New Product Forecasting. *Management Science*, *68*(6), 4685–4702. https://doi.org/10.1287/mnsc.2021.4102.

Feizabadi, J. (2022). Machine Learning Demand Forecasting and Supply Chain Performance. *International Journal of Logistics Research and Applications*, *25*(2), 119–142. https://doi.org/10.1080/13675567.2020.1803246.

Fildes, R. (1989). Evaluation of Aggregate and Individual Forecast Method Selection Rules. *Management Science*, *35*(9), 1056–1065. https://doi.org/10.1287/mnsc.35.9.1056.

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective Forecasting and Judgmental Adjustments: an empirical Evaluation and Strategies for Improvement in Supply-Chain Planning. *International Journal of Forecasting*, *25*(1), 3–23. https://doi.org/10.1016/j.ijforecast.2008.11.010.

Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2021). Reducing Algorithm Aversion through Experience. *Journal of Behavioral and Experimental Finance*, *31*, 1–21. https://doi.org/10.1016/j.jbef.2021.100524.

Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2023). The Extent of Algorithm Aversion in Decision-Making Situations with varying Gravity. *PLoS ONE*, *18*(2), e0278751. https://doi.org/10.1371/journal.pone.0278751.

Fine, A., Le, S., & Miller, M. K. (2023). Content Analysis of Judges' Sentiments Toward Artificial Intelligence Risk Assessment Tools. *Criminology, Criminal Justice, Law & Society*, *24*(2), 31–46. https://doi.org/10.54555/CCJLS.8169.84869.

Fink, L., Newman, L., & Haran, U. (2024). Let Me Decide: Increasing User Autonomy Increases Recommendation Acceptance. *Computers in Human Behavior*, *156,* 108244. https://doi.org/10.1016/j.chb.2024.108244.

Firk, S., Gehrke, Y., & Wolff, M. (2024). Digital Anxiety in the Finance Function: Consequences and Mitigating Factors. *Journal of Management Accounting Research*, *36*(1), 95–118. https://doi.org/10.2308/JMAR-2021-056.

Fishbein, M. (1963). An Investigation of the Relationships between Beliefs about an Object and the Attitude toward that Object. *Human Relations*, *16*(3), 233–239. https://doi.org/10.1177/001872676301600302.

Fishbein, M., & Ajzen, I. (2010). *Predicting and Changing Behavior: The Reasoned Action Approach.* Predicting and changing behavior: The reasoned action approach. New York, NY, US: Psychology Press.

Flood, M. W., & Grimm, B. (2021). EntropyHub: An open-source toolkit for entropic time series analysis. *PLoS ONE*, *16*(11), e0259448. https://doi.org/10.1371/journal.pone.0259448.

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, *12*(2), 219–245. https://doi.org/10.1177/1077800405284363.

Fogg, B. J., & Tseng, H. (1999). The Elements of Computer Credibility. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems 1999*, 80–87. https://doi.org/10.1145/302979.303001.

Freisinger, E., Unfried, M., & Schneider, S. (2024). The AI-Augmented Crowd: How Human Crowdvoters Adopt AI (or not). *Journal of Product Innovation Management*, *41*(4), 865–889. https://doi.org/10.1111/jpim.12708.

Freistühler, S., Kempkes, J. A., Suprano, F., & Wömpener, A. (2019). Controller und Data Scientist in der Unternehmenspraxis. *Controlling: Zeitschrift für erfolgsorientierte Unternehmensführung*, *31*(3), 63–69. https://doi.org/10.15358/0935-0381-2019-3-63.

Frick, W. (2015). Here's Why People Trust Human Judgment Over Algorithms. https://hbr.org/2015/02/heres-why-people-trust-human-judgment-over-algorithms. Accessed 18.09.2024.

Friedman, N., & Ormiston, J. (2022). Blockchain as a Sustainability-Oriented Innovation?: Opportunities for and Resistance to Blockchain Technology as a Driver of Sustainability in Global Food Supply Chains. *Technological Forecasting and Social Change*, *175,* 121403. https://doi.org/10.1016/j.techfore.2021.121403.

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, *33*(2), 678–696. https://doi.org/10.1287/isre.2021.1079.

Fuller, R. M., & Dennis, A. R. (2009). Does Fit Matter? The Impact of Task-Technology Fit and Appropriation on Team Performance in Repeated Tasks. *Information Systems Research*, *20*(1), 2–17. https://doi.org/10.1287/isre.1070.0167.

Ganbold, O., Rose, A. M., Rose, J. M., & Rotaru, K. (2022). Increasing Reliance on Financial Advice with Avatars: The Effects of Competence and Complexity on Algorithm Aversion. *Journal of Information Systems*, *36*(1), 7–17. https://doi.org/10.2308/ISYS-2021-002.

Gardner, B., & Abraham, C. (2010). Going Green? Modeling the Impact of Environmental Concerns and Perceptions of Transportation Alternatives on Decisions to Drive. *Journal of Applied Social Psychology*, *40*(4), 831–849. https://doi.org/10.1111/j.1559-1816.2010.00600.x.

Garland, J., James, R., & Bradley, E. (2014). Model-free Quantification of Time-Series Predictability. *Physical Review E: Statistical, Nonlinear, and Soft matter physics*, *90*(5-1), 052910. https://doi.org/10.1103/PhysRevE.90.052910.

Geerts, G. L. (2011). A Design Science Research Methodology and its Application to Accounting Information Systems Research. *International Journal of Accounting Information Systems*, *12*(2), 142–151. https://doi.org/10.1016/j.accinf.2011.02.004.

Geerts, G. L., Graham, L. E., Mauldin, E. G., McCarthy, W. E., & Richardson, V. J. (2013). Integrating Information Technology into Accounting Research and Practice. *Accounting Horizons*, *27*(4), 815–840. https://doi.org/10.2308/acch-50573.

Germann, M., & Merkle, C. (2022). Algorithm Aversion in Delegated Investing. *Journal of Business Economics*, 1–37. https://doi.org/10.1007/s11573-022-01121-9.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151.

Glienke, M., Hartwich, N., & Antons, D. (2023). Working with AI: How Attitudes Shape Human-AI Collaboration. *ICIS 2023 Proceedings*. https://aisel.aisnet.org/icis2023/hti/hti/12.

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057.

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, *19*(2), 213. https://doi.org/10.2307/249689.

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

Goodwin, P., & Fildes, R. (1999). Judgmental Forecasts of Time Series affected by Special Events: does providing a statistical Forecast improve Accuracy? *Journal of*

Behavioral Decision Making, *12*(1), 37–53. https://doi.org/10.1002/(SICI)1099-0771(199903)12:1<37:AID-BDM319>3.0.CO;2-8.
Goretzki, L., & Messner, M. (2016). Coordination under Uncertainty. *Qualitative Research in Accounting & Management*, *13*(1), 92–126. https://doi.org/10.1108/QRAM-09-2015-0070.
Goretzki, L., Strauss, E., & Wiegmann, L. (2018). Exploring the Roles of Vernacular Accounting Systems in the Development of "Enabling" Global Accounting and Control Systems. *Contemporary Accounting Research*, *35*(4), 1888–1916. https://doi.org/10.1111/1911-3846.12357.
Grabski, S. V., Leech, S. A., & Schmidt, P. J. (2011). A Review of ERP Research: A Future Agenda for Accounting Information Systems. *Journal of Information Systems*, *25*(1), 37–78. https://doi.org/10.2308/jis.2011.25.1.37.
Granlund, M. (2001). Towards explaining Stability in and around Management Accounting Systems. *Management Accounting Research*, *12*(2), 141–166. https://doi.org/10.1006/mare.2000.0151.
Granlund, M. (2011). Extending AIS Research to Management Accounting and Control Issues: A Research Note. *International Journal of Accounting Information Systems*, *12*(1), 3–19. https://doi.org/10.1016/j.accinf.2010.11.001.
Granlund, M., & Malmi, T. (2002). Moderate Impact of ERPS on Management Accounting: a Lag or Permanent Outcome? *Management Accounting Research*, *13*(3), 299–321. https://www.sciencedirect.com/science/article/pii/S1044500502901897.
Gray, A. M., & Phillips, V. L. (1994). Turnover, Age and Length of Service: a Comparison of Nurses and Other Staff in the National Health Service. *Journal of Advanced Nursing*, *19*(4), 819–827. https://doi.org/10.1111/j.1365-2648.1994.tb01155.x.
Greiner, B., Grünwald, P., Lindner, T., Lintner, G., & Wiernsperger, M. (2025). Incentives, Framing, and Reliance on Algorithmic Advice: An Experimental Study. *Management Science,* mnsc.2022.02777. https://doi.org/10.1287/mnsc.2022.02777.
Hamoudia, M., Makridakis, S., & Spiliotis, E. (2023). *Forecasting with Artificial Intelligence*. Cham: Springer Nature Switzerland.
Hansen, S. C., Otley, D. T., & van der Stede, W. A. (2003). Practice Developments in Budgeting: An Overview and Research Perspective. *Journal of Management Accounting Research*, *15*(1), 95–116. https://doi.org/10.2308/jmar.2003.15.1.95.
Haque, A. B., Islam, A. K. M. N., & Mikalef, P. (2023). Explainable Artificial Intelligence (XAI) from a User Perspective: A Synthesis of Prior Literature and Problematizing Avenues for Future Research. *Technological Forecasting and Social Change*. https://doi.org/10.1016/j.techfore.2022.122120.
Hassani, H., Mashhad, L. M., Royer-Carenzi, M., Yeganegi, M. R., & Komendantova, N. (2025). White Noise and Its Misapplications: Impacts on Time Series Model Adequacy and Forecasting. *Forecasting*, *7*(1), 8. https://doi.org/10.3390/forecast7010008.
Haupt, M., Freidank, J., & Haas, A. (2024). Consumer Responses to Human-AI Collaboration at Organizational Frontlines: Strategies to escape Algorithm Aversion in Content Creation. *Review of Managerial Science*, 1–37. https://doi.org/10.1007/s11846-024-00748-y.
Hayes, A. F., & Little, T. D. (Eds.) (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach.* (Second edition). Methodology in the social sciences. New York, London: The Guilford Press.
Hebeler, C. (2005). Neugestaltung der Unternehmensplanung bei Henkel. *Controlling: Zeitschrift für erfolgsorientierte Unternehmensführung*, *17*(8-9), 515–522. https://doi.org/10.15358/0935-0381-2005-8-9-515.

Heidemann, A., Hülter, S. M., & Tekieli, M. (2024). Machine Learning with Real-World HR Data: Mitigating the Trade-Off Between Predictive Performance and Transparency. *The International Journal of Human Resource Management*, 1–24. https://doi.org/10.1080/09585192.2024.2335515.

Heinle, M. S., Hofmann, C., & Kunz, A. H. (2012). Identity, Incentives, and the Value of Information. *The Accounting Review*, *87*(4), 1309–1334. https://doi.org/10.2308/accr-50156.

Heßler, P. O., Pfeiffer, J., & Hafenbrädl, S. (2022). When Self-Humanization Leads to Algorithm Aversion. *Business & Information Systems Engineering*, *64*(3), 275–292. https://doi.org/10.1007/s12599-022-00754-y.

Hevner, Alan R, March, Salvatore T, & Sudha (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625.

Hiebl, M. R., & Pielsticker, D. I. (2020). Stand der Digitalisierung des Controllings in Familienunternehmen. *Controlling: Zeitschrift für erfolgsorientierte Unternehmensführung*, *32*(6), 73–74. https://doi.org/10.15358/0935-0381-2020-6-73.

High-Level Expert Group AI (EU) (2019). *Ethics guidelines for trustworthy AI*. Brussels.

Himmelstein, M., & Budescu, D. V. (2023). Preference for Human or Algorithmic Forecasting Advice does not predict if and how it is used. *Journal of Behavioral Decision Making*, *36*(1), e2285. https://doi.org/10.1002/bdm.2285.

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors that influence Trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570.

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture*, *2*(1). https://doi.org/10.9707/2307-0919.1014.

Hogarth, R. M., & Makridakis, S. (1981). Forecasting and Planning: An Evaluation. *Management Science*, *27*(2), 115–138. https://doi.org/10.1287/mnsc.27.2.115.

Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A Unified Foundation for Business Analytics. *Decision Support Systems*, *64*, 130–141. https://doi.org/10.1016/j.dss.2014.05.013.

Holtom, B. C., Mitchell, T. R., Lee, T. W., & Eberly, M. B. (2008). Turnover and Retention Research: A Glance at the Past, a Closer Review of the Present, and a Venture into the Future. *Academy of Management Annals*, *2*(1), 231–274. https://doi.org/10.1080/19416520802211552.

Horngren, C. T., Datar, S. M., & Rajan, M. V. (2018). *Horngren's Cost Accounting: A Managerial Emphasis.* (Sixteenth edition, global edition). Always learning. Harlow, England: Pearson Education Limited.

Horowitz, M. C., & Kahn, L. (2024). Bending the Automation Bias Curve: A Study of Human and AI-Based Decision Making in National Security Contexts. *International Studies Quarterly*, *68*(2), sqae020. https://doi.org/10.1093/isq/sqae020.

Hou, Y.-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), 477. https://doi.org/10.1145/3479864.

Hülter, S. M., Ertel, C., & Heidemann, A. (2024). Exploring the Individual Adoption of Human Resource Analytics: Behavioural Beliefs and the Role of Machine Learning Characteristics. *Technological Forecasting and Social Change*, *208*, 123709. https://doi.org/10.1016/j.techfore.2024.123709.

Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-022-05049-6.

Hunton, J. E. (2002). Blending Information and Communication Technology with Accounting Research. *Accounting Horizons*, *16*(1), 55–67. https://doi.org/10.2308/acch.2002.16.1.55.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice.* (2nd edition). Melbourne: OTexts.

Hyndman, R. J., & Koehler, A. B. (2006). Another Look at Measures of Forecast Accuracy. *International Journal of Forecasting*, *22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001.

Hyvönen, T., Järvinen, J., Pellinen, J., & Rahko, T. (2009). Institutional Logics, ICT and Stability of Management Accounting. *European Accounting Review*, *18*(2), 241–275. https://doi.org/10.1080/09638180802681511.

IBM (2025). IBM Cognos Analytics. https://www.ibm.com/de-de/products/cognos-analytics. Accessed 08.07.2025.

Iftikhar, R., & Khan, M. S. (2020). Social Media Big Data Analytics for Demand Forecasting. *Journal of Global Information Management*, *28*(1), 103–120. https://doi.org/10.4018/JGIM.2020010106.

Ioannidis, J. P. A., Cripps, S., & Tanner, M. A. (2022). Forecasting for COVID-19 Has Failed. *International Journal of Forecasting*, *38*(2), 423–438. https://doi.org/10.1016/j.ijforecast.2020.08.004.

Jans, M., Aysolmaz, B., Corten, M., Joshi, A., & van Peteghem, M. (2023). Digitalization in Accounting – Warmly Embraced or Coldly Ignored? *Accounting, Auditing & Accountability Journal*, *36*(9), 61–85. https://doi.org/10.1108/AAAJ-11-2020-4998.

Järvenpää, M. (2007). Making Business Partners: A Case Study on how Management Accounting Culture was Changed. *European Accounting Review*, *16*(1), 99–142. https://doi.org/10.1080/09638180701265903.

Jauernig, J., Uhl, M., & Walkowitz, G. (2022). People Prefer Moral Discretion to Algorithms: Algorithm Aversion Beyond Intransparency. *Philosophy & Technology*, *35*(1), 1–25. https://doi.org/10.1007/s13347-021-00495-y.

Jordan, S., & Messner, M. (2020). The Use of Forecast Accuracy Indicators to Improve Planning Quality: Insights from a Case Study. *European Accounting Review*, *29*(2), 337–359. https://doi.org/10.1080/09638180.2019.1577150.

Jung, M., & Seiter, M. (2021). Towards a better Understanding on Mitigating Algorithm Aversion in Forecasting: an Experimental Study. *Journal of Management Control*, *32*(4), 495–516. https://doi.org/10.1007/s00187-021-00326-3.

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why Are We Averse Towards Algorithms? A Comprehensive Literature Review on Algorithm Aversion. *ECIS Proceedings 2020*, 168. https://aisel.aisnet.org/ecis2020_rp/168.

Jussupow, E., Benbasat, I., & Heinzl, A. (2024). An Integrative Perspective on Algorithm Aversion and Appreciation in Decision-Making. *MIS Quarterly*, *48*(4), 1575–1590. https://doi.org/10.25300/MISQ/2024/18512.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–292. https://doi.org/10.2307/1914185.

Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information Technology Adoption Across Time: A Cross-Sectional Comparison of Pre-Adoption and Post-Adoption Beliefs. *MIS Quarterly*, *23*(2), 183–213. https://doi.org/10.2307/249751.

Kasurinen, T. (2002). Exploring Management Accounting Change: the Case of Balanced Scorecard Implementation. *Management Accounting Research*, *13*(3), 323–343. https://doi.org/10.1006/mare.2002.0191.

Kawaguchi, K. (2021). When Will Workers Follow an Algorithm? A Field Experiment with a Retail Business. *Management Science*, *67*(3), 1670–1695. https://doi.org/10.1287/mnsc.2020.3599.

Keding, C., & Meissner, P. (2021). Managerial Overreliance on AI-Augmented Decision-Making Processes: How the Use of AI-Based Advisory Systems Shapes Choice Behavior in R&D Investment Decisions. *Technological Forecasting and Social Change*, *171*, 120970. https://doi.org/10.1016/j.techfore.2021.120970.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals*, *14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174.

Khashei, M., & Bijari, M. (2011). A novel Hybridization of Artificial Neural Networks and ARIMA Models for Time Series Forecasting. *Applied Soft Computing*, *11*(2), 2664–2675. https://doi.org/10.1016/j.asoc.2010.10.015.

Kilimci, Z. H., Akyuz, A. O., Uysal, M., Akyokus, S., Uysal, M. O., Atak Bulbul, B., & Ekmis, M. A. (2019). An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain. *Complexity*, *2019*(1), 9067367, 9067367. https://doi.org/10.1155/2019/9067367.

Kim, A., Yang, M., & Zhang, J. (2023a). When Algorithms Err Differential Impact of Early vs. Late Errors on Users' Reliance on Algorithms. *ACM Transactions on Computer-Human Interaction*, *30*(1), 1–36. https://doi.org/10.1145/3557889.

Kim, D., Song, Y., Kim, S., Lee, S., Wu, Y., Shin, J., & Lee, D. (2023b). How Should the Results of Artificial Intelligence be Explained to Users? - Research on Consumer Preferences in User-Centered Explainable Artificial Intelligence. *Technological Forecasting and Social Change*, *188*, 122343. https://doi.org/10.1016/j.techfore.2023.122343.

Kinney, W. R. (1978). ARIMA and Regression in Analytical Review: An Empirical Test. *The Accounting Review*, *53*(1), 48–60. http://www.jstor.org/stable/245725.

Knauer, T., Nikiforow, N., & Wagener, S. (2020). Determinants of Information System Quality and Data Quality in Management Accounting. *Journal of Management Control*, *31*(1-2), 97–121. https://doi.org/10.1007/s00187-020-00296-y.

Kokina, J., Gilleran, R., Blanchette, S., & Stoddard, D. (2021). Accountant as Digital Innovator: Roles and Competencies in the Age of Automation. *Accounting Horizons*, *35*(1), 153–184. https://doi.org/10.2308/HORIZONS-19-145.

Korhonen, T., Selos, E., Laine, T., & Suomala, P. (2021). Exploring the Programmability of Management Accounting Work for Increasing Automation: an Interventionist Case Study. *Accounting, Auditing & Accountability Journal*, *34*(2), 253–280. https://doi.org/10.1108/AAAJ-12-2016-2809.

Kostić, N., & Sedej, T. (2022). Blockchain Technology, Inter-Organizational Relationships, and Management Accounting: A Synthesis and a Research Agenda. *Accounting Horizons*, *36*(2), 123–141. https://doi.org/10.2308/HORIZONS-19-147.

Kumar, A., Patel, T., Benjamin, A. S., & Steyvers, M. (2021). Explaining Algorithm Aversion with Metacognitive Bandits. *43rd Annual Meeting of the Cognitive Science Society: Comparative Cognition: Animal Minds, CogSci 2021*. https://experts.illinois.edu/en/publications/explaining-algorithm-aversion-with-metacognitive-bandits.

Küpper, H.-U., Friedl, G., Hofmann, C., Hofmann, Y., & Pedell, B. (2024). *Controlling: Konzeption - Aufgaben - Instrumente.* (7th ed.). Stuttgart: Schäffer-Poeschel.

Kusche, I. (2024). Possible Harms of Artificial Intelligence and the EU AI Act: Fundamental Rights and Risk. *Journal of Risk Research*, 1–14. https://doi.org/10.1080/13669877.2024.2350720.

La Barbera, F., & Ajzen, I. (2020). Control Interactions in the Theory of Planned Behavior: Rethinking the Role of Subjective Norm. *Europe's Journal of Psychology*, *16*(3), 401–417. https://doi.org/10.5964/ejop.v16i3.2056.

Lacroux, A., & Martin-Lacroux, C. (2022). Should I Trust the Artificial Intelligence to Recruit? Recruiters' Perceptions and Behavior When Faced With Algorithm-Based Recommendation Systems During Resume Screening. *Frontiers in Psychology*, *13*(895997). https://doi.org/10.3389/fpsyg.2022.895997.

Lamot, K., & Paulussen, S. (2020). Six Uses of Analytics: Digital Editors' Perceptions of Audience Analytics in the Newsroom. *Journalism Practice*, *14*(3), 358–373. https://doi.org/10.1080/17512786.2019.1617043.

Langer, M., & König, C. J. (2021). Introducing a Multi-stakeholder Perspective on Opacity, Transparency and Strategies to reduce Opacity in Algorithm-based Human Resource Management. *Human Resource Management Review,* 100881. https://doi.org/10.1016/j.hrmr.2021.100881.

Langer, M., König, C. J., Back, C., & Hemsing, V. (2023). Trust in Artificial Intelligence: Comparing Trust Processes Between Human and Automated Trustees in Light of Unfair Bias. *Journal of Business and Psychology*, *38*, 493–508. https://doi.org/10.1007/s10869-022-09829-9.

Lavia López, O., & Hiebl, M. R. W. (2015). Management Accounting in Small and Medium-Sized Enterprises: Current Knowledge and Avenues for Further Research. *Journal of Management Accounting Research*, *27*(1), 81–119. https://doi.org/10.2308/jmar-50915.

Lee, C., & Cha, K. (2023). FAT-CAT—Explainability and Augmentation for an AI System: A Case Study on AI Recruitment-System Adoption. *International Journal of Human-Computer Studies*, *171,* 102976. https://doi.org/10.1016/j.ijhcs.2022.102976.

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

Lee, S., & Peng, T.-Q. (2024). Understanding Audience Behavior with Digital Traces: Past, Present, and Future. *Digital Journalism*, *12*(8), 1140–1164. https://doi.org/10.1080/21670811.2023.2254821.

Liu, N. T. Y., Kirshner, S. N., & Lim, E. T. (2023). Is Algorithm Aversion WEIRD? A cross-country Comparison of individual-Differences and Algorithm Aversion. *Journal of Retailing and Consumer Services*, *72,* 103259. https://doi.org/10.1016/j.jretconser.2023.103259.

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. https://doi.org/10.1016/j.obhdp.2018.12.005.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013.

Longoni, C., Cian, L., & Kyung, E. J. (2023). Algorithmic Transference: People Overgeneralize Failures of AI in the Government. *Journal of Marketing Research*, *60*(1), 170–188. https://doi.org/10.1177/00222437221110139.

Lynch, J. G. (1999). Theory and External Validity. *Journal of the Academy of Marketing Science*, *27*(3), 367–376. https://doi.org/10.1177/0092070399273007.

Lynch, J. J. G. (1982). On the External Validity of Experiments in Consumer Research. *Journal of Consumer Research*, *9*(3), 225. https://doi.org/10.1086/208919.

Maasland, C., & Weißmüller, K. S. (2022). Blame the Machine? Insights From an Experiment on Algorithm Aversion and Blame Avoidance in Computer-Aided Human Resource Management. *Frontiers in Psychology*, *13,* 779028. https://doi.org/10.3389/fpsyg.2022.779028.

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and Differences Between Human–Human and Human–Automation Trust: An Integrative Review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277–301. https://doi.org/10.1080/14639220500337708.

Magni, F., Park, J., & Chao, M. M. (2024). Humans as Creativity Gatekeepers: Are We Biased Against AI Creativity? *Journal of Business and Psychology*, *39*(3), 643–656. https://doi.org/10.1007/s10869-023-09910-x.

Mahmud, H., Islam, A. N., Ahmed, S. I., & Smolander, K. (2022). What Influences Algorithmic Decision-Making? A Systematic Literature Review on Algorithm Aversion. *Technological Forecasting and Social Change*, *175,* 121390. https://doi.org/10.1016/j.techfore.2021.121390.

Mahmud, H., Islam, A. N., & Mitra, R. K. (2023). What Drives Managers Towards Algorithm Aversion and How to Overcome it? Mitigating the Impact of Innovation Resistance through Technology Readiness. *Technological Forecasting and Social Change*, *193,* 122641. https://doi.org/10.1016/j.techfore.2023.122641.

Makridakis, S., Hibon, M., Lusk, E., & Belhadjali, M. (1987). Confidence Intervals. *International Journal of Forecasting*, *3*(3-4), 489–508. https://doi.org/10.1016/0169-2070(87)90045-8.

Makridakis, S., Hyndman, R. J., & Petropoulos, F. (2020a). Forecasting in Social Settings: The State of the Art. *International Journal of Forecasting*, *36*(1), 15–28. https://doi.org/10.1016/j.ijforecast.2019.05.011.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020b). The M4 Competition: 100,000 Time Series and 61 Forecasting Methods. *International Journal of Forecasting*, *36*(1), 54–74. https://doi.org/10.1016/j.ijforecast.2019.04.014.

Makridakis, S., & Taleb, N. (2009). Decision Making and Planning under low Levels of Predictability. *International Journal of Forecasting*, *25*(4), 716–733. https://doi.org/10.1016/j.ijforecast.2009.05.013.

Malmi, T., & Brown, D. A. (2008). Management Control Systems as a Package—Opportunities, Challenges and Research Directions. *Management Accounting Research*, *19*(4), 287–300. https://doi.org/10.1016/j.mar.2008.09.003.

Manokhin, V. (2023). Unlocking Predictive Power: Harnessing Permutation Entropy for Superior Time-Series Forecasting. *Medium,* 16.10.2023. https://valeman.medium.com/unlocking-predictive-power-harnessing-permutation-entropy-for-superior-time-series-forecasting-7772b52f4ec4. Accessed 02.07.2025.

Margherita, A. (2022). Human Resources Analytics: A Systematization of Research Topics and Directions for Future Research. *Human Resource Management Review*, *32*(2), 100795. https://doi.org/10.1016/j.hrmr.2020.100795.

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

Martins, A., Bianchi de Aguiar, M. T., Sambento, M., & Branco, M. C. (2025). Business Intelligence System Adoption and the Leveraging of Reporting Process Capabilities. *Journal of Accounting & Organizational Change*, *21*(3), 506–534. https://doi.org/10.1108/JAOC-11-2023-0204.

Marton, J. (2025). Can Research in Accounting be Relevant for Practice? *Accounting in Europe*, 1–12. https://doi.org/10.1080/17449480.2025.2527599.

Mauldin, E. G., & Ruchala, L. V. (1999). Towards a Meta-Theory of Accounting Information Systems. *Accounting, Organizations and Society*, *24*(4), 317–331. https://doi.org/10.1016/S0361-3682(99)00006-9.

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.

McCarthy, W. E. (2012). Accounting Craftspeople Versus Accounting Seers: Exploring the Relevance and Innovation Gaps in Academic Accounting Research. *Accounting Horizons*, *26*(4), 833–843. https://doi.org/10.2308/acch-10313.

McCartney, S., & Fu, N. (2022). Bridging the Gap: Why, How and When HR Analytics can Impact Organizational Performance. *Management Decision*, *60*(13), 25–47. https://doi.org/10.1108/MD-12-2020-1581.

Meehl, P. E. (1954). *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. Minneapolis: University of Minnesota Press.

Meijerink, J., & Bondarouk, T. (2023). The Duality of Algorithmic Management: Toward a Research Agenda on HRM Algorithms, Autonomy and Value Creation. *Human Resource Management Review*, *33*(1), 100876. https://doi.org/10.1016/j.hrmr.2021.100876.

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), 2545–2562. https://doi.org/10.1080/09585192.2021.1925326.

Merton, R. K., & Kendall, P. L. (1946). The Focused Interview. *American Journal of Sociology*, *51*(6), 541–557. https://doi.org/10.1086/219886.

Microsoft (2025). Microsoft Fabric. https://www.microsoft.com/de-de/microsoft-fabric. Accessed 08.07.2025.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big Data Analytics and Firm Performance: Findings from a Mixed-Method Approach. *Journal of Business Research*, *98*, 261–276. https://doi.org/10.1016/j.jbusres.2019.01.044.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook.* (2nd ed.). Thousand Oaks, Calif.: Sage.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook.* (3rd ed.). Los Angeles, London, New Delhi, Singapore, Washington DC: Sage.

Mok, L., Nanda, S., & Anderson, A. (2023). People Perceive Algorithmic Assessments as Less Fair and Trustworthy Than Identical Human Assessments. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2), 1–26. https://doi.org/10.1145/3610100.

Möller, K., Schäffer, U., & Verbeeten, F. (2020). Digitalization in Management Accounting and Control: an Editorial. *Journal of Management Control*, *31*(1), 1–8. https://doi.org/10.1007/s00187-020-00300-5.

Morrell, K., Loan-Clarke, J., & Wilkinson, A. (2001). Unweaving Leaving: the Use of Models in the Management of Employee Turnover. *International Journal of Management Reviews*, *3*(3), 219–244. https://doi.org/10.1111/1468-2370.00065.

Moser, D. V. (2012). Is Accounting Research Stagnant? *Accounting Horizons*, *26*(4), 845–850. https://doi.org/10.2308/acch-10312.

Mula, C., Zybura, N., & Hipp, T. (2024). From Digitalized Start-up to Scale-up: Opening the Black Box of Scaling in Digitalized Firms towards a Scaling Process

Framework. *Technological Forecasting and Social Change*, *202,* 123275. https://doi.org/10.1016/j.techfore.2024.123275.

Namdari, A., & Li, Z. (2019). A Review of Entropy Measures for Uncertainty Quantification of Stochastic Processes. *Advances in Mechanical Engineering*, *11*(6), 1-14. https://doi.org/10.1177/1687814019857350.

Neely, A., Bourne, M., & Adams, C. (2003). Better Budgeting or Beyond Budgeting? *Measuring Business Excellence*, *7*(3), 22–28. https://doi.org/10.1108/13683040310496471.

Neumann, M., Niessen, A. S. M., Hurks, P. P. M., & Meijer, R. R. (2023a). Holistic and Mechanical Combination in Psychological Assessment: Why Algorithms are Underutilized and What is Needed to Increase their Use. *International Journal of Selection and Assessment*, *31*(2), 267–285. https://doi.org/10.1111/ijsa.12416.

Neumann, M., Niessen, A. S. M., & Meijer, R. R. (2023b). Predicting Decision-Makers' Algorithm Use. *Computers in Human Behavior*, *145*(107759). https://doi.org/10.1016/j.chb.2023.107759.

Neumann, M., Niessen, A. S. M., Tendeiro, J. N., & Meijer, R. R. (2022a). The Autonomy-Validity Dilemma in Mechanical Prediction Procedures: The Quest for a Compromise. *Journal of Behavioral Decision Making*, *35*(4), e2270. https://doi.org/10.1002/bdm.2270.

Neumann, O., Guirguis, K., & Steiner, R. (2022b). Exploring Artificial Intelligence Adoption in Public Organizations: a Comparative Case Study. *Public Management Review*, 1–28. https://doi.org/10.1080/14719037.2022.2048685.

Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When Eliminating Bias isn't Fair: Algorithmic Reductionism and Procedural Justice in Human Resource Decisions. *Organizational Behavior and Human Decision Processes*, *160*, 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008.

Nickell, E. B., Schwebke, J., & Goldwater, P. (2023). An Introductory Audit Data Analytics Case Study: Using Microsoft Power BI and Benford's Law to Detect Accounting Irregularities. *Journal of Accounting Education*, *64,* 100855. https://doi.org/10.1016/j.jaccedu.2023.100855.

Nielsen, S. (2018). Reflections on the Applicability of Business Analytics for Management Accounting – and Future Perspectives for the Accountant. *Journal of Accounting & Organizational Change*, *14*(2), 167–187. https://doi.org/10.1108/JAOC-11-2014-0056.

Nielsen, S. (2022). Management Accounting and the Concepts of Exploratory Data Analysis and Unsupervised Machine Learning: a Literature Study and Future Directions. *Journal of Accounting & Organizational Change*, *18*(5), 811–853. https://doi.org/10.1108/JAOC-08-2020-0107.

Niszczota, P., & Kaszás, D. (2020). Robo-Investment Aversion. *PLoS ONE*, *15*(9), 1–19, e0239277. https://doi.org/10.1371/journal.pone.0239277.

Ochmann, J., Zilker, S., Michels, L., Tiefenbeck, V., & Laumer, S. (2021). The Influence of Algorithm Aversion and Anthropomorphic Agent Design on the Acceptance of AI-Based Job Recommendations. *ICIS 2020 Proceedings*, 1–17.

Oesterreich, T. D., & Teuteberg, F. (2019). The Role of Business Analytics in the Controllers and Management Accountants' Competence Profiles. *Journal of Accounting & Organizational Change*, *15*(2), 330–356. https://doi.org/10.1108/JAOC-10-2018-0097.

Oliva, R., & Watson, N. (2009). Managing Functional Biases in Organizational Forecasts: A Case Study of Consensus Forecasting in Supply Chain Planning.

Production and Operations Management, *18*(2), 138–151. https://doi.org/10.1111/j.1937-5956.2009.01003.x.

Omrani, N., Rivieccio, G., Fiore, U., Schiavone, F., & Agreda, S. G. (2022). To Trust or not to Trust? An Assessment of Trust in AI-based Systems: Concerns, Ethics and Contexts. *Technological Forecasting and Social Change*, *181,* 121763. https://doi.org/10.1016/j.techfore.2022.121763.

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

Oomen, T., Gonçalves, J., & Mols, A. (2024). Rage Against the Artificial Intelligence?: Understanding Contextuality of Algorithm Aversion and Appreciation. *International Journal of Communication*, *18*, 609–633. https://ijoc.org/index.php/ijoc/article/view/20809/4454.

Otley, D. (1999). Performance Management: a Framework for Management Control Systems Research. *Management Accounting Research*, *10*(4), 363–382. https://doi.org/10.1006/mare.1999.0115.

Pabinger, D., Feldbauer-Durstmüller, B., Duller, C., & Hiebl, M. R. W. (2021). Digitalisierung im Controlling: Verdrängt der Data Scientist den Controller? *Controlling: Zeitschrift für erfolgsorientierte Unternehmensführung*, *33*(5), 43–45. https://doi.org/10.15358/0935-0381-2021-5-43.

Paialunga, P. (2023). Time Series Complexity Analysis Using Entropy - TDS Archive - Medium. *TDS Archive,* 04.09.2023. https://medium.com/data-science/time-series-complexity-analysis-using-entropy-ec49a4aaff11. Accessed 02.07.2025.

Park, H., Ahn, D., Hosanagar, K., & Lee, J. Human-AI Interaction in Human Resource Management: Understanding Why Employees Resist Algorithmic Evaluation at Workplaces and How to Mitigate Burdens. *Proceedings of the CHI 2021 Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3411764.3445304.

Park, J., Feng, Y., & Jeong, S.-P. (2024). Developing an Advanced Prediction Model for New Employee Turnover Intention Utilizing Machine Learning Techniques. *Scientific Reports*, *14*(1), 1221. https://doi.org/10.1038/s41598-023-50593-4.

Parlier, J., & Lee, L. (2023). Inventory Analytics: A Teaching Case using Excel and Alteryx. *Journal of Accounting Education*, *63*(100848). https://doi.org/10.1016/j.jaccedu.2023.100848.

Paulhus, D. L. (1984). Two-Component Models of Socially Desirable Responding. *Journal of Personality and Social Psychology*, *46*(3), 598–609. https://doi.org/10.1037/0022-3514.46.3.598.

Perryer, C., Jordan, C., Firns, I., & Travaglione, A. (2010). Predicting Turnover Intentions. *Management Research Review*, *33*(9), 911–923. https://doi.org/10.1108/01409171011070323.

Peters, M. D., Wieder, B., Sutton, S. G., & Wakefield, J. (2016). Business Intelligence Systems Use in Performance Measurement Capabilities: Implications for Enhanced Competitive Advantage. *International Journal of Accounting Information Systems*, *21*, 1–17. https://doi.org/10.1016/j.accinf.2016.03.001.

Petrie, K., Seeliger, A., & Tischler, R. (2024). The Future of BI & Analytics – Adopting GenAI for Analytics. *BARC - Data Decisions. Built on BARC.,* 2024. https://barc.com/research/bi-analytics-genai-trends-lessons-best-practices/. Accessed 02.07.2025.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., Baets, S. de, Dokumentov, A., Ellison, J., Fiszeder, P., Franses, P. H., Frazier, D. T., Gilliland, M., Gönül, M. S., Goodwin, P., Grossi, L., Grushka-Cockayne, Y., Guidolin, M., Guidolin, M., Gunter, U., Guo, X., Guseo, R., Harvey, N., Hendry, D. F., Hollyman, R., Januschowski, T., Jeon, J., Jose, V. R. R., Kang, Y., Koehler, A. B., Kolassa, S., Kourentzes, N., Leva, S., Li, F., Litsiou, K., Makridakis, S., Martin, G. M., Martinez, A. B., Meeran, S., Modis, T., Nikolopoulos, K., Önkal, D., Paccagnini, A., Panagiotelis, A., Panapakidis, I., Pavía, J. M., Pedio, M., Pedregal, D. J., Pinson, P., Ramos, P., Rapach, D. E., Reade, J. J., Rostami-Tabar, B., Rubaszek, M., Sermpinis, G., Shang, H. L., Spiliotis, E., Syntetos, A. A., Talagala, P. D., Talagala, T. S., Tashman, L., Thomakos, D., Thorarinsdottir, T., Todini, E., Trapero Arenas, J. R., Wang, X., Winkler, R. L., Yusupova, A., & Ziel, F. (2022). Forecasting: Theory and Practice. *International Journal of Forecasting*, *38*(3), 705–871. https://doi.org/10.1016/j.ijforecast.2021.11.001.

Pezzo, M. V., Nash, B., Vieux, P., & Foster-Grammer, H. W. (2021). Effect of Having, but Not Consulting, a Computerized Diagnostic Aid. *Medical Decision Making*. https://doi.org/10.1177/0272989X211011160.

Pfister, J. A., & Lukka, K. (2019). Interrelation of Controls for Autonomous Motivation: A Field Study of Productivity Gains Through Pressure-Induced Process Innovation. *The Accounting Review*, *94*(3), 345–371. https://doi.org/10.2308/accr-52266.

Plattfaut, R., & Borghoff, V. (2022). Robotic Process Automation: A Literature-Based Research Agenda. *Journal of Information Systems*, *36*(2), 173–191. https://doi.org/10.2308/ISYS-2020-033.

Polimeni, R. S., & Burke, J. A. (2021). Integrating Emerging Accounting Digital Technologies and Analytics into an Undergraduate Accounting Curriculum—A Case Study. *Journal of Emerging Technologies in Accounting*, *18*(1), 159–173. https://doi.org/10.2308/JETA-2020-042.

Prahl, A., & van Swol, L. (2017). Understanding Algorithm Aversion: When is Advice from Automation Discounted? *Journal of Forecasting*, *36*(6), 691–702. https://doi.org/10.1002/for.2464.

Prikshat, V., Islam, M., Patel, P., Malik, A., Budhwar, P., & Gupta, S. (2023a). AI-Augmented HRM: Literature Review and a Proposed Multilevel Framework for Future Research. *Technological Forecasting and Social Change*, *193,* 122645. https://doi.org/10.1016/j.techfore.2023.122645.

Prikshat, V., Malik, A., & Budhwar, P. (2023b). AI-Augmented HRM: Antecedents, Assimilation and Multilevel Consequences. *Human Resource Management Review*, *33*(1), 100860. https://doi.org/10.1016/j.hrmr.2021.100860.

Proeger, T., & Meub, L. (2014). Overconfidence as a Social Bias: Experimental Evidence. *Economics Letters*, *122*(2), 203–207. https://doi.org/10.1016/j.econlet.2013.11.027.

Proksch, S., Schühle, J., Streeb, E., Weymann, F., Luther, T., & Kimmerle, J. (2024). The Impact of Text Topic and Assumed Human vs. AI Authorship on Competence and Quality Assessment. *Frontiers in Artificial Intelligence*, *7,* 1412710. https://doi.org/10.3389/frai.2024.1412710.

Pumplun, L., Tauchert, C., & Heidt, M. (2019). A New Organizational Chassis for Artificial Intelligence - Exploring Organizational Readiness factors. *ECIS Proceedings 2019*. https://aisel.aisnet.org/ecis2019_rp/106.

Qasim, A., Issa, H., El Refae, G. A., & Sannella, A. J. (2020). A Model to Integrate Data Analytics in the Undergraduate Accounting Curriculum. *Journal of Emerging Technologies in Accounting*, *17*(2), 31–44. https://doi.org/10.2308/JETA-2020-001.

Quintão, C., Andrade, P., & Almeida, F. (2020). How to Improve the Validity and Reliability of a Case Study Approach? *Journal of Interdisciplinary Studies in Education*, *9*(2), 273–284. https://doi.org/10.32674/jise.v9i2.2026.

Rabinovitch, H., Budescu, D. V., & Meyer, Y. B. (2024). Algorithms in Selection Decisions: Effective, but Unappreciated. *Journal of Behavioral Decision Making*, *37*(2), e2368. https://doi.org/10.1002/bdm.2368.

Reich, T., Kaju, A., & Maglio, S. J. (2023). How to Overcome Algorithm Aversion: Learning from Mistakes. *Journal of Consumer Psychology*, *33*(2), 285–302. https://doi.org/10.1002/jcpy.1313.

Remneland Wikhamn, B., Styhre, A., & Wikhamn, W. (2023). HRM Work and Open Innovation: Evidence from a Case Study. *The International Journal of Human Resource Management*, *34*(10), 1940–1972. https://doi.org/10.1080/09585192.2022.2054285.

Renier, L. A., Schmid Mast, M., & Bekbergenova, A. (2021). To Err is Human, not Algorithmic – Robust Reactions to Erring Algorithms. *Computers in Human Behavior*, *124*(106879). https://doi.org/10.1016/j.chb.2021.106879.

Retraction Watch (2025). James Hunton – Retraction Watch. https://retractionwatch.com/category/by-author/james-hunton/. Accessed 26.08.2025.

Rikhardsson, P., & Yigitbasioglu, O. (2018). Business Intelligence & Analytics in Management Accounting Research: Status and Future Focus. *International Journal of Accounting Information Systems*, *29*, 37–58. https://doi.org/10.1016/j.accinf.2018.03.001.

Riordan, C. M., Gatewood, R. D., & Bill, J. B. (1997). Corporate Image: Employee Reactions and Implications for Managing Corporate Social Performance. *Journal of Business Ethics*, *16*(4), 401–412. https://doi.org/10.1023/A:1017989205184.

Rogers, E. M. (2003). *Diffusion of Innovations.* (Fifth edition). Social science. New York, London, Toronto, Sydney: Free Press.

Rom, A., & Rohde, C. (2007). Management Accounting and Integrated Information Systems: A Literature Review. *International Journal of Accounting Information Systems*, *8*(1), 40–68. https://doi.org/10.1016/j.accinf.2006.12.003.

Rombaut, E., & Guerry, M.-A. (2018). Predicting Voluntary Turnover Through Human Resources Database Analysis. *Management Research Review*, *41*(1), 96–112. https://doi.org/10.1108/MRR-04-2017-0098.

Rubenstein, A. L., Eberly, M. B., Lee, T. W., & Mitchell, T. R. (2018). Surveying the Forest: A Meta-Analysis, Moderator Investigation, and Future-Oriented Discussion of the Antecedents of Voluntary Employee Turnover. *Personnel Psychology*, *71*(1), 23–65. https://doi.org/10.1111/peps.12226.

Salvino, L. W., Cawley, R., Grebogi, C., & Yorke, J. A. (1995). Predictability in Time Series. *Physics Letters A*, *209*(5-6), 327–332. https://doi.org/10.1016/0375-9601(95)00871-3.

SAP (2025). SAP Analytics Cloud. https://pages.community.sap.com/topics/cloud-analytics. Accessed 08.07.2025.

Saragih, M., & Morrison, B. W. (2021). The Effect of Past Algorithmic Performance and Decision Significance on Algorithmic Advice Acceptance. *International Journal of Human-Computer Interaction*, *38*(13), 1228–1237. https://doi.org/10.1080/10447318.2021.1990518.

Schäffer, U., & Weber, J. (2016). Die Digitalisierung wird das Controlling radikal verändern. *Controlling & Management Review*, *60*(6), 6–17. https://doi.org/10.1007/s12176-016-0093-9.

Schepers, J., & Wetzels, M. (2007). A Meta-Analysis of the Technology Acceptance Model: Investigating Subjective Norm and Moderation Effects. *Information & Management*, *44*(1), 90–103. https://doi.org/10.1016/j.im.2006.10.007.

Schmidt, P., Biessmann, F., & Teubner, T. (2020a). Transparency and Trust in Artificial Intelligence Systems. *Journal of Decision Systems*, *29*(4), 260–278. https://doi.org/10.1080/12460125.2020.1819094.

Schmidt, P. J., Riley, J., & Swanson Church, K. (2020b). Investigating Accountants' Resistance to Move Beyond Excel and Adopt New Data Analytics Technology. *Accounting Horizons*, *34*(4), 165–180. https://doi.org/10.2308/HORIZONS-19-154.

Schneider, S., & Freisinger, E. (2022). Overcoming Algorithm Aversion: The Power of Task-Procedure-Fit. *Academy of Management Proceedings*, *2022*(1), 15716. https://doi.org/10.5465/AMBPP.2022.122.

Schuessler, E. S., Lohmeyer, N., & Ashwin, S. (2023). "We Can't Compete on Human Rights": Creating Market-Protected Spaces to Institutionalize the Emerging Logic of Responsible Management. *Academy of Management Journal*, *66*(4), 1071–1101. https://doi.org/10.5465/amj.2020.1614.

Schulte Steinberg, A. L., & Hohenberger, C. (2023). Can AI close the Gender Gap in the Job Market? Individuals' Preferences for AI Evaluations. *Computers in Human Behavior Reports*, *10*(100287). https://doi.org/10.1016/j.chbr.2023.100287.

Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples.* (4th edition). SpringerLink Bücher. Cham: Springer International Publishing.

Sinclair-Desgagné, B. (2024). On the (Non-) Reliance on Algorithms—A Decision-Theoretic Account. *Journal of Mathematical Psychology*, *119,* 102844. https://doi.org/10.1016/j.jmp.2024.102844.

Sivabalan, P., Booth, P., Malmi, T., & Brown, D. A. (2009). An Exploratory Study of Operational Reasons to Budget. *Accounting & Finance*, *49*(4), 849–871. https://doi.org/10.1111/j.1467-629X.2009.00305.x.

Snow, T. (2021). From Satisficing to Artificing: The Evolution of Administrative Decision-Making in the Age of the Algorithm. *Data & Policy*, *3,* e3. https://doi.org/10.1017/dap.2020.25.

Sohrabpour, V., Oghazi, P., Toorajipour, R., & Nazarpour, A. (2021). Export Sales Forecasting Using Artificial Intelligence. *Technological Forecasting and Social Change*, *163,* 120480. https://doi.org/10.1016/j.techfore.2020.120480.

Soto-Ferrari, M. (2025). SOT-FER: A Multi-Tier Entropy-Based Time Series Forecasting Framework with an Application to Manufacturing. *Computers & Industrial Engineering*, *204,* 111071. https://doi.org/10.1016/j.cie.2025.111071.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google Effects on Memory: Cognitive Consequences of Having Information at our Fingertips. *Science*, *333*(6043), 776–778. https://doi.org/10.1126/science.1207745.

Spraakman, G., Sanchez-Rodriguez, C., & Tuck-Riggs, C. A. (2021). Data Analytics by Management Accountants. *Qualitative Research in Accounting & Management*, *18*(1), 127–147. https://doi.org/10.1108/QRAM-11-2019-0122.

Sprinkle, G. B., Williamson, M. G., & Upton, D. R. (2008). The Effort and Risk-Taking Effects of Budget-Based Contracts. *Accounting, Organizations and Society*, *33*(4-5), 436–452. https://doi.org/10.1016/j.aos.2007.11.001.

Steiner, H., & Welker, P. (2016). Wird der Controller zum Data Scientist? *Controlling & Management Review*, *60*(S1), 68–73. https://doi.org/10.1007/s12176-016-0003-1.

Sutton, S. G. (2006). Enterprise Systems and the Re-shaping of Accounting Systems: A call for Research. *International Journal of Accounting Information Systems*, *7*(1), 1–6. https://doi.org/10.1016/j.accinf.2006.02.002.

Sutton, S. G., Arnold, V., Collier, P., & Leech, S. A. (2021). Leveraging the Synergies between Design Science and Behavioral Science Research Methods. *International Journal of Accounting Information Systems*, *43*(100536). https://doi.org/10.1016/j.accinf.2021.100536.

Sutton, S. G., Arnold, V., & Holt, M. (2023). An Extension of the Theory of Technology Dominance: Capturing the Underlying Causal Complexity. *International Journal of Accounting Information Systems*, *50,* 100626. https://doi.org/10.1016/j.accinf.2023.100626.

Tandoc, E. C., & Thomas, R. J. (2015). The Ethics of Web Analytics. *Digital Journalism*, *3*(2), 243–258. https://doi.org/10.1080/21670811.2014.909122.

Tse, T. T. K., Hanaki, N., & Mao, B. (2024). Beware the Performance of an Algorithm before Relying on it: Evidence from a Stock Price Forecasting Experiment. *Journal of Economic Psychology*, *102,* 102727. https://doi.org/10.1016/j.joep.2024.102727.

Tucker, B., & Parker, L. (2014). In our Ivory Towers? The Research-Practice Gap in Management Accounting. *Accounting and Business Research*, *44*(2), 104–143. https://doi.org/10.1080/00014788.2013.798234.

Tucker, B. P., & Lowe, A. D. (2014). Practitioners are from Mars; Academics are from Venus? *Accounting, Auditing & Accountability Journal*, *27*(3), 394–425. https://doi.org/10.1108/AAAJ-01-2012-00932.

Turel, O., & Kalhan, S. (2023). Prejudiced Against the Machine? Implicit Associations and the Transience of Algorithm Aversion. *MIS Quarterly*. https://doi.org/10.25300/MISQ/2022/17961.

Tursunbayeva, A., Pagliari, C., Di Lauro, S., & Antonelli, G. (2022). The Ethics of People Analytics: Risks, Opportunities and Recommendations. *Personnel Review*, *51*(3), 900–921. https://doi.org/10.1108/PR-12-2019-0680.

van den Broek, E., Sergeeva, A., & Huysman Vrije, M. (2021). When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. *MIS Quarterly*, *45*(3), 1557–1580. https://doi.org/10.25300/MISQ/2021/16559.

van Deursen, A. J., Helsper, E. J., & Eynon, R. (2016). Development and Validation of the Internet Skills Scale (ISS). *Information, Communication & Society*, *19*(6), 804–823. https://doi.org/10.1080/1369118X.2015.1078834.

van Donselaar, K. H., Gaur, V., van Woensel, T., Broekmeulen, R. A. C. M., & Fransoo, J. C. (2010). Ordering Behavior in Retail Stores and Implications for Automated Replenishment. *Management Science*, *56*(5), 766–784. https://doi.org/10.1287/mnsc.1090.1141.

van Slooten, A. C., Dirks, P. M., & Firk, S. (2024). Digitalization and Management Accountants' Role Conflict and Ambiguity: A Double-Edged Sword for the Profession. *The British Accounting Review*(101460). https://doi.org/10.1016/j.bar.2024.101460.

Vargas, R. (2016). *Adoption Factors Impacting Human Resource Analytics among Human Resource Professionals.* Adoption factors impacting human resource analytics among human resource professionals., 76. US: ProQuest Information & Learning.

Vargas, R., Yurova, Y. V., Ruppel, C. P., Tworoger, L. C., & Greenwood, R. (2018). Individual Adoption of HR Analytics: A Fine Grained View of the Early Stages Leading to Adoption. *The International Journal of Human Resource Management*, *29*(22), 3046–3067. https://doi.org/10.1080/09585192.2018.1446181.

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big Data in Accounting: An Overview. *Accounting Horizons*, *29*(2), 381–396. https://doi.org/10.2308/acch-51071.

Venkatesh, Thong, & Xu (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540.

Venkatesh, V., Thong, J., & Xu, X. (2016). Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead. *Journal of the Association for Information Systems*, *17*(5), 328–376. https://doi.org/10.17705/1jais.00428.

Wang, R., & Tan, J. (2021). Exploring the Coupling and Forecasting of Financial Development, Technological Innovation, and Economic Growth. *Technological Forecasting and Social Change*, *163,* 120466. https://doi.org/10.1016/j.techfore.2020.120466.

Weaver, R. K. (1986). The Politics of Blame Avoidance. *Journal of Public Policy*, *6*(4), 371–398. https://doi.org/10.1017/S0143814X00004219.

Werner, M. (2017). Financial Process Mining - Accounting Data Structure Dependent Control Flow Inference. *International Journal of Accounting Information Systems*, *25*, 57–80. https://doi.org/10.1016/j.accinf.2017.03.004.

Youssef, M. A. E.-A., & Mahama, H. (2021). Does Business Intelligence Mediate the Relationship Between ERP and Management Accounting Practices? *Journal of Accounting & Organizational Change*, *17*(5), 686–703. https://doi.org/10.1108/JAOC-02-2020-0026.

Yzer, M., & van den Putte, B. (2014). Control Perceptions Moderate Attitudinal and Normative Effects on Intention to Quit Smoking. *Psychology of Addictive Behaviors*, *28*(4), 1153–1161. https://doi.org/10.1037/a0037924.

Zachariah, R. A., Sharma, S., & Kumar, V. (2023). Systematic Review of Passenger Demand Forecasting in Aviation Industry. *Multimedia Tools and Applications*, *82*(30), 1–37. https://doi.org/10.1007/s11042-023-15552-1.

Zerbino, P., Stefanini, A., & Aloini, D. (2021). Process Science in Action: A Literature Review on Process Mining in Business Management. *Technological Forecasting and Social Change*, *172,* 121021. https://doi.org/10.1016/j.techfore.2021.121021.

Zhang, C., Cho, S., & Vasarhelyi, M. (2022). Explainable Artificial Intelligence (XAI) in Auditing. *International Journal of Accounting Information Systems*, *46*, 100572.

Zhang, K. (2021). How Information Processing Style Shapes People's Algorithm Adoption. *Social Behavior and Personality*, *49*(8), 1–13. https://doi.org/10.2224/sbp.10417.

Zhang, P. G. (2003). Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing*, *50*, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0.

Zorn, M., Vandesype, N., Vercruyssen, V., Büter, J., Nobleaux, M., Stumpf, M., & Vonholder, M. (2025). Predictive Analytics Market Study: How Finance Leaders already leverage Predictive Analytics. *Deloitte,* 2025. https://www.deloitte.com/ch/en/services/consulting/research/predictive-analytics-market-study.html. Accessed 02.07.2025.

(removed in online document)