Journal of Management Control
https://doi.org/10.1007/s00187-025-00391-y

ORIGINAL PAPER

The effects of disclosing an algorithm's inner workings and analytic thinking on algorithm reliance

Alexander Pinzger[1]

Accepted: 28 April 2025

Abstract

This study draws on the idea that disclosing information about how algorithmic decision aids process data ('inner workings') may increase algorithm reliance. In this context, the decision-maker's thinking style is likely to play a crucial role, as analytic thinking may help to process algorithm information and therefore to foster algorithm reliance. In two between-subject experiments, I examine whether (the extent of) information provided about an algorithm and thinking style influence individuals' reliance on the algorithm. The findings suggest that disclosing an algorithm's inner workings significantly increases algorithm reliance up to a certain threshold, implying a concave relationship between information load and algorithm reliance. While analyses do not indicate an interaction effect between information disclosure and analytic thinking, a joint effect on algorithm reliance is shown. Furthermore, supplemental analyses of the study uncover the underlying mechanisms of algorithm reliance: Analytic thinking and other individual characteristics, such as a general faith in technology, are positively related to perceived advice source credibility that, in turn, is positively correlated with algorithm reliance. In this regard, the effect of advice source credibility appears to be stronger for analytic and weaker for intuitive thinkers. Additionally, I find that numerical differences such as the span between advices or advice direction correspond to varying levels of algorithm reliance. Thus, the study presents valuable insights for designing decision-making environments that may reinforce the utilization of data-driven approaches.

Keywords Advice source credibility · Algorithm reliance · Digitalization · Forecasting · Thinking style · Experimental study

JEL classification M41

✉ Alexander Pinzger
alexander.pinzger@uni-ulm.de

1 Institute of Management Accounting and Control, Ulm University, Helmholtzstraße 18, 89081 Ulm, Germany

Published online: 27 May 2025

1 Introduction

Data analytics have become crucial for most organizations due to their potential to improve process quality (Senoner et al., 2022), decision-making and performance (Mariani & Wamba, 2020) and enable new ways of analyzing (digital) information (Loebbecke & Picot, 2015; Vitale et al., 2020). Algorithmic decision aids are particularly beneficial for processing future-oriented information (Dietvorst et al., 2018; Knudsen, 2020), as they address the dynamics and complexity of the environment by integrating real-time data and unstructured information from external sources (Arnaboldi et al., 2017). In a management accounting context, algorithmic decision aids therefore play a key role in providing more accurate information for managerial decision-making (Guenther, 2013; Sprinkle, 2003).

However, research has shown that decision-makers tend to distrust and devaluate algorithms and are, thus, hesitant to use them (Önkal et al., 2017), particularly after the algorithm errs (Dietvorst et al., 2015) or when it presents a downward trend (Chen et al., 2022; Fehrenbacher et al., 2023). The literature points at the notable paradox of algorithm aversion, given that decision-makers often prefer human judgment over algorithms, even when the latter lead to superior decisions (Dietvorst et al., 2015; Elkins et al., 2013; Grove et al., 2000; Highhouse, 2008; Silver, 2012). Contrarily, some studies show algorithm appreciation in certain environments (e.g., Bogert et al., 2022), e.g., depending on task type (Logg et al., 2019) or task difficulty (Bogert et al., 2021). Against this background, a deeper understanding of how to mitigate irrational behavior and therefore to reinforce the use of algorithms appears crucial in order to improve decision-making.

An essential idea to promote algorithm reliance is disclosing detailed information on how an algorithm processes data. By sharing such 'inner workings' of an algorithm, the existing hesitation to use it and false expectations associated with it, may be reduced (Burton et al., 2020). Such a disclosure has the potential to reduce skepticism and build trust in an opaque algorithm (Adadi & Berrada, 2018; Jobin et al., 2019). However, it is decisive how much information is disclosed, given that too much information may overstrain decision-makers (e.g., Roetzel, 2019; Stocks & Harrell, 1995). In this regard, sharing relevant information (Hartmann & Weißenberger, 2023), in a sufficient amount (Turel & Kalhan, 2023), appears crucial to create a higher level of understanding and to overcome biases. As a result, decision-makers may perceive an algorithm as a more competent and credible advice source, leading to a higher tendency to use algorithmic decision aids (Alvarado-Valencia & Barrero, 2014; Fogg & Tseng, 1999).

However, such judgments are not necessarily a standard procedure when people make decisions. From the perspective of the dual-process theory from the field of psychology, the default decision-making situation is mainly intuition-based (Evans & Stanovich, 2013; Stanovich & West, 2008). A more analytic decision mode can be initiated in a way that rational, deliberate and therefore slower thinking overrides intuitive judgments that may be prone to errors (Evans, 2003). Besides reducing such biases, analytic thinking might contribute to a higher cognitive compatibility between an algorithm and a decision-maker (Burton et al., 2020; Vessey, 1991), implying an alignment of human thinking patterns and the algorithmic approach. In combination with the previously mentioned disclosure of an algorithm's inner workings, this study investigates whether different thinking styles have an effect on algorithm reliance.

Drawing on these ideas, the objective of this paper is to analyze these mechanisms that may mitigate existing biases and therefore enhance algorithm reliance. In this context, I conduct two between-subject experiments in which participants decide on sales predictions based on advice in a forecasting setting. The advice sources are a human expert (product manager) and a data scientist who relies on a self-designed forecasting algorithm. In experiment 1, the disclosure of an algorithm's 'inner workings' (disclosed vs. not disclosed) is manipulated by varying the availability of information and 'thinking style' (intuitive vs. analytic) with the help of a priming approach. In experimental accounting research, the psychological concept of priming is used to affect participant behavior in different ways (Clor-Proell & Nelson, 2007; Hammersley et al., 2010; Lambert & Agoglia, 2011; Thomas, 2016; Wolfe et al., 2020). In this study, priming is used to induce an analytic mindset and to reinforce intuitive thinking as performed in White (2005) and Huntsinger (2011). Experiment 2 provides more detailed insights on the 'inner workings' idea by manipulating three different levels of information availability ('low' vs. 'medium' vs. 'detailed'). After submitting their final decisions, participants indicate the extent to which they rely on both advice sources, out of which the assigned weight on algorithmic advice is the main variable of interest.

The findings partially support my hypotheses. The findings of experiment 2 suggest that disclosing relevant information about an algorithm's inner workings significantly increases algorithm reliance up to a certain threshold of information load, implying a concave relationship between information load and algorithm reliance. Supplemental analyses of experiment 1 identify a significant joint effect of disclosing an algorithm's inner workings and thinking style on algorithm reliance. Furthermore, the supplemental analyses show that there is an effect of analytic thinking on perceived advice source credibility that, in turn, has an increasing effect on algorithm reliance. Moreover, several individual characteristics, especially the general propensity of individuals to think rationally or intuitively, are related to the use of algorithms. In particular, highly intuitive decision-makers seem to make decisions rather independently of how credible they perceive the advice source to be. In addition, Experiment 2 shows that, as risk tendencies and forecast accuracy perceptions of participants matter as well, the span between advice values and the advice direction (i.e., which advice source forecasts a higher / lower value) result in different degrees of algorithm reliance in a way that algorithm reliance is increased for higher spans and lower forecast values.

This study contributes to the literature on algorithm aversion and appreciation as well as to the management accounting literature. Research on algorithm reliance is discussed in numerous fields such as computer science and information systems (Mahmud et al., 2022). Irrespective of the diverse causes for algorithm aversion (Burton et al., 2020), further empirical evidence on the promotion of algorithm reliance is needed to understand the complexity of decision-making involving data analytics. The paper addresses this by showing that disclosing an algorithm's inner workings enhances algorithm reliance. In addition to this, I build on findings from information overload studies (Hartmann & Weißenberger, 2023; Roetzel, 2019; Stocks &

Harrell, 1995) to demonstrate that information disclosure increases the use of algorithmic aids up to a certain threshold, implying a concave relation between information load and algorithm reliance. Moreover, analytic thinking is identified to be positively associated with the perceived credibility of an algorithmic advice source. This is the case without giving any direct performance-related information on the algorithm as proposed by Logg et al. (2019) in order to enhance the appreciation of algorithms. Therefore, I extend the credibility-related findings of Chen et al. (2022) in a way that perceived credibility is shown to be conditional on the specific type of advice source and individual characteristics. Furthermore, this study investigates possibilities to actively control irrational behavior rooted in cognitive biases in management accounting as called for by Wibbeke and Lachmann (2020). In this regard, my research aims at integrating psychological aspects, especially cognition and the dual-process theory, into the field of management accounting. While Fehrenbacher et al. (2018) already use this theory in the context of subjective performance evaluation, this study can be seen as a starting point in the forecasting domain. In this context, the decision-maker's thinking style and the disclosure of algorithm information have a joint effect on the use of algorithms. These are crucial findings to add to the still limited research on data analytics in a management accounting context (Rikhardsson & Yigitbasioglu, 2018). Taken together, new insights will be given in terms of understanding the black-box reasoning process of decision-makers encountering algorithmic decision aids (Tank & Farrell, 2022).

The structure of this paper is as follows: In Sect. 2, the theoretical background of algorithm reliance and the development of hypotheses is explained. Thereafter, the research method and experimental design choices of experiment 1 and experiment 2 (Sect. 3) are presented. Results as well as limitations are discussed in Sects. 4 and 5 before drawing final conclusions in Sect. 6.

2 Theoretical background and hypotheses development

2.1 Algorithmic decision aids and algorithm reliance

Data analytics encompass various algorithmic approaches that transform input data into decision-relevant information (Gillespie et al., 2014; Schneider et al., 2015).[1] The application of algorithmic decision aids aims at a more objective and therefore better decision-making (Glikson & Woolley, 2020). Algorithms surpass traditional decision support systems by processing vast amounts of data (Logg et al., 2019), where traditional decision support systems and human decision-makers reach their technical and cognitive limits (Al-Htaybat & von Alberti-Alhtaybat, 2017; Warren et al., 2015). By integrating new and relevant information – like social media data – into forecasts, algorithms enable the inclusion of indicators that were previously not available for analyses (Arnaboldi et al., 2017; Cui et al., 2018; Knudsen, 2020).

[1] Note that while I mainly use the term 'algorithm' as it is easier to explain in the scope of an experiment, the findings of this study generally refer to the broad concepts of data analytics.

Providing such information is a major objective of management accounting as it aims to support rational decision-making (Guenther, 2013; Sprinkle, 2003). While many economic models assume such rationality (Simon, 1986), human decision-makers often tend to intuitive and emotionally affected reasoning as it is usually very efficient (Farrell et al., 2014). However, intuitive reasoning may lead to decisions with lower economic value (Evans, 2008; Moreno et al., 2002) as important information might not be considered (Farrell et al., 2014). In other words, even though all humans are generally able to engage in rational decision-making, humans do not necessarily apply rational thinking effectively (Kahneman et al., 1982; Trotman et al., 2011).

Despite their potential to promote rational decision-making, algorithms are not readily embraced by users. Algorithm aversion, a term coined by Dietvorst et al. (2015), is a phenomenon that describes people's hesitation to rely on algorithms in decision-making processes. Since potential benefits of relying on algorithm-based information are knowingly ignored and a suboptimal decision basis is accepted, such aversion is a type of irrational behavior that can be reduced (Arnott, 2006). Findings related to algorithm aversion date back many decades. For example, Meehl (1954) and Dawes (1979) find a certain degree of hesitation of humans to use statistical predictions compared to human judgment. These findings reflect what is relevant today with regard to the usage of more sophisticated algorithms. From a management accounting perspective, a substantial area is forecasting as individuals tend to place greater weight on advice from human experts compared to statistical methods when making forecast adjustments (Önkal et al., 2009). This might be problematic since algorithms are able to include a more diverse set of data (Arnaboldi et al., 2017).

Challenging the literature on algorithm aversion, there are several contrasting findings that indicate an appreciation for algorithms. For example, Logg et al. (2019) show that in some rather specific tasks, algorithms are preferred over humans, as for example in forecasts of music popularity and romantic attraction. Therefore, task type and also task difficulty (Bogert et al., 2021), i.e., people tend to favor algorithms in more difficult tasks, affect the prevailing attitude towards algorithms. As Logg et al. (2019) point out, it is important to further understand algorithm reliance in decision-making in order to reduce the risk of missed potentials. This seems especially important, as studies show that the aversion towards algorithms remains a prevalent issue (e.g., Berger et al., 2021; Turel & Kalhan, 2023). Prior research has investigated the reliance on algorithms from different perspectives, indicating that algorithm trust and a corresponding reliance are highly complex and far from elucidated phenomena (Burton et al., 2020; Castelo et al., 2019; Dietvorst et al., 2015).

In this regard, it is crucial to investigate factors that cause algorithm reliance. While some authors find possibilities to increase algorithm reliance, e.g., by enabling decision-makers to modify the algorithmic suggestion (Dietvorst et al., 2018) or to integrate their own forecast into algorithmic processing (Kawaguchi, 2021), more insights are needed to understand algorithm reliance more thoroughly. In their review of the literature, Burton et al. (2020) identify two reasons which may serve as a starting point for enhancing algorithm reliance. First, people may have *false expectations* regarding an algorithm, implying that the decision aid is not perceived adequately and does not meet their expectations. In this regard, people seem to have unrealistically high expectations towards algorithms as they might even expect algorithms to make perfect predictions (Madhavan & Wiegmann, 2007). Awareness of algorithmic errors, therefore, is even more severe for algorithm reliance as they are perceived as systematic, while human inaccuracies tend to be perceived as natural (Dietvorst et al., 2018; Highhouse, 2008). Second, *cognitive compatibility* refers to the (perceived) difference between the human and the algorithmic cognitive approach, implying a different way of processing data. In general, individuals tend to prefer advice sources that they consider similar to themselves. As algorithms often seem opaque to humans (*black box*), an alignment between an algorithm and the human process is difficult to realize. These two reasons appear strongly connected, given that false expectations are likely to increase the perceived distance between human and algorithm. In the following sections, I will discuss how false expectations and cognitive differences may be addressed, resulting in the development of my hypotheses.[2]

2.2 Disclosing an algorithm's inner workings and perceived credibility

As outlined before, forecasting might benefit from additional information by the use of algorithms, particularly as forecasting is prone to diverse errors. These are caused, for instance, by irrational decisions due to missing knowledge in general (Brüggen et al., 2021), the high complexity of forecasts (Chen et al., 2022) or the forecaster's tendency of overconfidence (Logg et al., 2019). Using an algorithm that is able to perform more accurate predictions might contribute to overcoming these limitations (Teoh, 2018).

Considering an algorithm to be an advice source in a forecasting setting, advice source credibility appears crucial, given that individuals rely more often on sources that they consider more credible (Alvarado-Valencia & Barrero, 2014; Petty & Cacioppo, 1986). In this regard, perceived competence and trustworthiness are the two main components that determine perceived credibility. In order for an advice source to be perceived as more competent and trustworthy, it should be perceived as knowledgeable and unbiased (Fogg & Tseng, 1999). For an algorithmic advice source, this is particularly difficult to achieve as many decision-makers lack experience with such decision-aids (Gillespie et al., 2023).

Therefore, disclosing ideas and steps in the development and the decision-making process of an algorithmic model ('*inner workings*') might help to demystify the previously opaque approach of algorithms. Especially for unexperienced users, basics such as the distinction between training and test data are crucial to open the frequently perceived *black box* (Gillespie et al., 2023; Miller, 2019). This might establish a starting point regarding algorithm understanding and experience that builds trust and therefore credibility towards algorithms as the literature on Explainable

[2] Although this paper investigates ideas how to promote algorithm reliance, it is important to note that the use of algorithms can induce certain risks. To mention a few, the promotion of algorithms might lead to over-reliance on algorithms in settings where even algorithms reach their limits (Krügel et al., 2022) or to egoistic and morally questionable decisions (Krügel et al., 2023). A decision maker's task should always be to question the limitations and possible consequences of trusting an algorithmic decision aid (Janssen et al., 2017). Nevertheless, in order to benefit from the numerous potentials by simultaneously controlling for negative effects, it is crucial to illuminate factors that drive algorithm reliance.

Artificial Intelligence (XAI) points out (Adadi & Berrada, 2018; Lipton, 2018). The aspect of transparency that involves the disclosure of central information seems to be the most prevailing idea in this regard (Jobin et al., 2019). In line with this, I suggest that a description of core principles and its operating approach, for example the presentation of inherent rule sets, is an essential part of making an algorithm more credible.[3]

As outlined in Sect. 2.1, humans tend to have false expectations that are in conflict with the credibility of algorithms. Correcting such false expectations might enhance the perceived credibility of the algorithmic advice source by making decision-makers aware of main assumptions and limitations. Gaining insights into how an algorithmic model processes information may help decision-makers recognize two important aspects. Firstly, considering *superiority*, decision-makers who believe that a human or their own approach is superior, e.g., due to overconfidence (Logg et al., 2019; Sieck & Arkes, 2005), might learn that many algorithms actually perform similar steps with the advantage of a comparatively higher processing power, resulting in a faster and potentially less erroneous outcome (Jussupow et al., 2020). Individuals therefore might identify situations where human decision-making may be prone to error or bias and consider using algorithms to augment or replace human judgment. Secondly, considering *task understanding*, decision-makers might learn about the possibility to incorporate new, potentially crucial aspects like additional data (Breuker et al., 2016) into their own analysis, resulting in a new understanding of impactful drivers. This is a precondition for (big) data analytics that, in most cases, are not possible to perform without some sort of algorithmic intervention (Bhimani & Willcocks, 2014).

Next to correcting false expectations, both aspects are essential considerations to achieve a higher alignment between human and an algorithmic approach that should mitigate the previously outlined cognitive differences (see Sect. 2.1). Eventually, an algorithm might be perceived more realistically, and therefore more understandable and credible. Dealing with the inner workings of an algorithm can help individuals to understand how algorithmic suggestions and results are developed and reflect on drivers that have an effect in a certain decision-making process, making it easier to justify the use of algorithms.[4] Thus, decision-makers can consider the algorithm a more credible source as it might be assessed (rationally) superior to human experts, resulting in a stronger algorithm acceptance.

The XAI literature finds several possibilities to make complex algorithms more explainable (Adadi & Berrada, 2018) to increase trust in and therefore the usage of algorithms. Interesting findings in this regard are that completeness and soundness

[3] Prior research already analyzed the effect of performance information of an algorithm on its use. Ensuring more realistic expectations like an algorithm's better performance *on average* (Dawes et al., 1989) should increase algorithm credibility and therefore mitigate the aversion. However, the existing literature points out that such performance information might even be considered a main driver of algorithm aversion as it reveals algorithmic errors (Dietvorst et al., 2015; Jung & Seiter, 2021) that, in turn, conflict with the unrealistically high expectations (Dietvorst et al., 2018). This is why I focus on a description referring to the general functioning of an algorithm.

[4] In a broader context, this might be a critical aspect in terms of accountability as well. Decision-makers often ought to explain their decisions to internal or external parties (Malmi & Brown, 2008; Messner, 2009) which is a particularly complicated challenge in case it involves (big) data analytics (Arnaboldi et al., 2017).

of explanations are important drivers of algorithm reliance while completeness tends to be more relevant than soundness (Kulesza et al., 2013). This suggests that giving rather comprehensive information is more promising than a seemingly sound but incomplete set of information. In line with this, Turel and Kalhan (2023) point out that sufficient information is needed to correct subconscious biases against an algorithmic decision aid.

Since this research focuses on rather broad algorithms, disclosing detailed yet basic background information about an algorithm may enhance general algorithm understanding, thereby reducing false expectations and cognitive differences. The correspondingly enhanced perceived credibility may result in an increasing effect on algorithm reliance. Therefore, I state the following hypothesis:

> H1: When inner workings are disclosed (not disclosed), the decision-maker's reliance on algorithms is higher (lower).

2.3 Thinking style

Since algorithms process information in an entirely analytic and rational way, a considerable difference exists compared to humans who consciously and subconsciously rely on intuition. Insights into how individuals reason are provided by the dual-process theory that is broadly discussed in the cognitive psychology literature (Evans, 2003, 2008, 2019; Kahneman & Frederick, 2002). According to this theory, an individual's reasoning process is two-fold: intuitive thinking is an instinctive form of universal cognition that is natural, automatic and heuristic (Gervais & Norenzayan, 2012; Stanovich & West, 2013); analytic thinking, in contrast, works completely sequential, logical, deductive and is therefore considered to trigger rational behavior (Evans, 2003).

There are individual differences in a way that some people tend to engage more in analytic thinking than others (Cacioppo & Petty, 1982; Frederick, 2005). Moreover, analytic thinking has limits due to its cognitive load (Evans, 2008), given that it is often time-consuming and demanding. Nevertheless, people are able to perform both thinking styles (Evans & St, 2003; Stanovich & West, 2008).

The dual-process theory suggests that intuitive thinking is active by default (Evans & Stanovich, 2013; Stanovich & West, 2008). However, it is possible to activate an analytic mindset in a sense that it may override parts of individual thinking (Evans, 2003). In case of activation, decision-makers may take more time to logically think about the superior solution to a problem and might deviate from their intuition. For some tasks, intuitive thinking is beneficial, i.e., faster, especially when task complexity is low, consequences are not too severe or the decision-maker's experience and, thus, intuition is of high value for the decision (Bargh & Ferguson, 2000; Dijksterhuis & Nordgren, 2006). However, major decisions with higher complexity are often susceptible for biases and can therefore be improved by analytic reasoning (Milkman et al., 2009).

Due to its rational nature, individuals engaging in analytic thinking will take their time to strive for the best solution and react less affective (Kahneman, 2011). Thus, an algorithm that processes more task-relevant data than a human should be considered rationally superior and therefore may be preferred. In contrast to that, intuitive thinking might be a main reason for decision-makers to discount algorithms, for example due to the tendency of individuals to distrust technology in the first place (Davis, 1989) so that they intuitively fall back to their own judgment or rather use a human advice source. An intuitive preference for human judgment might occur due to the higher similarity between oneself and another human being, as, to a certain degree, decision-makers tend to strive for decisions that are based on familiarity (Simon, 1955). When sufficient information about an algorithm is given, this is strongly connected to the cognitive compatibility idea that was previously introduced: Individuals engaging in analytic thinking should feel less distant from an algorithm, resulting in a higher compatibility between the algorithmic approach and the individual way of processing.

Moreover, individuals cannot only reflect on their own biases and preconceptions (White, 2005), but also on the opportunities that an algorithmic decision aid might provide in this context. Analytic thinking therefore might support decision-makers to avoid making irrational judgments about algorithms. This idea is closely related to the effect of disclosing an algorithm's inner workings as individuals can only analytically consider an algorithm's potential advantage if it is not entirely opaque. Contrarily, the findings of Woolley and Risen (2018) suggest that people tend to prefer their intuitive judgment even if other, more objective information is provided. Thus, disclosed algorithm information can be processed more effectively by individuals engaging in analytic thinking, potentially resulting in an increased use of the algorithm. This implies a positive interaction effect of analytic thinking on the relation between information disclosure and algorithm reliance.

However, it is also possible that this potential interaction effect is weakened when individuals who prefer intuitive thinking rely on algorithmic advice in order to compensate for their own lack of analytic thinking. Disclosing how an algorithm works may help to identify the complementarity of the approaches chosen by the intuitively thinking individual and the algorithm. Given the ambiguity inherent in the relationship between thinking style and disclosing an algorithm's inner workings, I formulate the following non-directional hypothesis.

> H2: The effect of disclosing an algorithm's inner workings on algorithm reliance depends on whether analytic or intuitive thinking is activated.

In order to address the hypotheses, I conducted two experiments. In the next section, the experimental design and procedures of these experiment will be discussed.

3 Experimental design

3.1 Experiment 1

3.1.1 Participants

I conducted an online experiment with 300 participants recruited from *Prolific* (www.prolific.com).[5] Due to several attention check questions that will be outlined in Sect. 3.1.5, the main sample includes 196 participants on whom main analyses are conducted. Essential information about the task and context was presented in the experiment and no special knowledge or experience in management or accounting was needed to understand the task itself. Focusing on the general population appears appropriate since, for example, students often show a higher degree of algorithmic knowledge (Alexander et al., 2018; Önkal et al., 2009). Furthermore, as Mahmud et al. (2022) point out, algorithm reliance studies should be conducted with samples that rather represent the real decision environment in terms of a more diverse population. Therefore, the only precondition for all participants was that they were German native speakers, since the text-intensive experiment was conducted in German language. The demographic data shows a diverse set of participants in terms of age (m = 32.4, sd = 10.3), gender (male: 115; female: 75; diverse: 5; not specified: 1) and profession (employed: 109; self-employed: 19; student: 52; other: 16). Median completion time of the experiment was close to 14 min for which the participants received £ 2.25, resulting in an average hourly wage of £ 9.64, thus, exceeding the hourly wage of £ 9.00 recommended by Prolific.[6]

3.1.2 Overview and experimental task

In the experiment, participants assumed the role of a decision-maker with high responsibility for a fictional company. They had to make a final product demand estimation after they had received several information about the decision-making setting. This information involved a product manager (human expert) and his estimation as well as a data scientist who initiated an algorithm that generates another forecasting estimation. Since there was no specific information about financial data, participants were expected to use the product manager's output, the algorithm's output or something in between, so that the extent of algorithm reliance could be observed. In order to test the hypotheses, I manipulated thinking style by conducting a word completion task at the beginning of the experiment and information disclosure of the algorithm within

[5] Prolific is an online crowdfunding platform, comparable to Amazon Mechanical Turk. Comparing different online platforms, Prolific is considered to provide the highest data quality for research purposes (Peer et al., 2017, 2022). In terms of effort and intrinsic motivation, online experiments might even be considered advantageous over lab experiments (Farrell et al., 2017). Facing challenges like less controllable environments, Prolific has comparably strict rules and guidelines, such as a comprehensive check and attention check policy, the possibility to reject certain participants due to misbehavior or restrictions to certain devices. (Palan & Schitter, 2018; www.prolific.com).

[6] Prolific points out that £ 9.00 or more is a fair payment and leads to higher data quality (https://researcher-help.prolific.com/en/article/2273bd).

the experimental information. The experimental task was adapted from Chen et al. (2022) who use the task in an algorithm reliance context as well. After the decision, participants answered several questions for additional analyses (post-experimental questionnaire).

The experiment proceeded as follows. At first, a brief overview over the study was presented to the participants, followed by a word completion task that will be explained in the next section. After that, a general introduction to the case was given where participants were told that they assume the role of a CEO of *InnoToy*, an innovative toy company. Further, they were informed about a new product that would soon be produced. For this reason, a demand forecast had to be finally determined by the CEO. Case information included an explanation that the forecast decision was crucial for personnel and production planning in order to meet market demands. In their role as CEO, participants were supposed to make an informed decision. Input for their decision could be obtained from two advice sources: A product manager and a data scientist who utilizes an algorithm. Both internal sources of information were presented as equally experienced employees whose salary is dependent on the precision of their forecast. The product manager's estimation was set to 70,000 units and the algorithm estimation to 50,000 units.

3.1.3 Independent variables

Regarding the experimental design, I used a 2 × 2 between-subjects format to manipulate the independent variables 'inner workings' (with regard to the algorithm) and 'thinking style'.

The 'inner workings disclosed' condition provided additional information about the algorithmic decision aid. More precisely, information about its operating approach and basic principles about its development were given to enable the decision-maker to develop a thorough understanding of how the algorithm works and why it does so. Participants, for example, received information about a content analysis that classified text from social media and news articles to be negative or positive in order to estimate the sales number more accurately. Moreover, the distinction between a *training set* and a *validation set* as a main development principle was explained thoroughly. To avoid a potential information overload, the given information was limited to this substantial information about the algorithm processing basics (Hartmann & Weißenberger, 2023). The 'inner workings not disclosed' condition provided no deeper information about the algorithm, solely stating that the algorithm incorporates data from several sources in order to generate the data scientist's forecast estimations.

Regarding the 'thinking style' manipulation, the concept of priming was used. Priming can be understood as an implicit stimulus that affects the subsequent way of thinking (Clor-Proell & Nelson, 2007; Higgins et al., 1985). The analytic thinking condition primed participants for analytic reasoning with the help of a word completion task,[7] previously performed in Huntsinger (2011) and White (2005). The

[7] I acknowledge that alternative ways to elicit analytic thinking, such as directly induce thinking style through explicit instructions as in Wolfe et al. (2020), could have been employed. However, such approaches are applicable in cases where participants have to perform a task that involves problem-

intuitive thinking condition primed for the intuitive counter-part instead. The word completion task included ten target words that are strongly associated with either analytic thinking or intuitive thinking, e.g., the German translations of ‘analytic’, ‘rational’, ‘systematic’ or ‘intuitive’, ‘spontaneous’, ‘effortless’. One letter was missing in each case so that participants had to type in the complete word. Next to the target words, five identical neutral words were added to both treatments in order for the target concepts to be less obvious. The fifteen words in total appeared in random order and were followed by one example that provided feedback in case participants did not enter the complete word correctly.

3.1.4 Dependent variable

To measure algorithm reliance as the variable of main interest, there are several possibilities, including the weight on algorithmic advice, the forecast decision itself and the perceived credibility of the algorithm. To capture how much weight participants placed on the algorithmic advice, they were asked to report the extent to which they included the advice in their decision immediately after making it. This question was presented as a two-fold slider from 0 to 100% for both advice sources that prevented inconsistent answers by allowing the sum of both weights to be 100% only. Due to its relative character, the reported weight on algorithmic advice reflects best what this study aims to analyze. Therefore, I use this conceptualization – that is a self-reported perception of algorithm reliance – as main dependent variable. To measure absolute algorithm reliance, it is possible to use the forecast decision itself. But there might be flaws in terms of interpreting absolute numbers (Chen et al., 2022) so that the participants’ decisions might not mirror their actual preference towards the algorithmic advice.[8] Furthermore, perceived credibility of the advice source as one driver in this context was also measured by seven-point Likert-type scale questions, similar to

solving or analysis (Dane & Pratt, 2009). This is not the case in this experiment. Additionally, a limitation of such direct methods is the high difficulty to determine whether participants actually used the desired thinking style (Dane & Pratt, 2009). Furthermore, experimental participants might tend to pay more attention to questions than to instructions (Peer et al., 2017). Thus, an indirect task might even be more appropriate to trigger a subconscious thinking style than a direct instruction that might be overlooked or even ignored. Still, the priming effects proved difficult to determine in this experiment as the results in Sect. 4.1.1 show.

[8] In this context, two aspects should be noted: First, Chen et al. (2022) investigate if advice valence (good news vs. bad news) has an effect on algorithm reliance. In order to manipulate advice valence, a baseline of 60,000 units is set so that 50,000 units equal bad news and 70,000 equal good news. Results indicate that there might be lower algorithm reliance if an algorithm presents bad news. In order to adapt the experimental task to this paper, firstly, a baseline is not necessary and, secondly, algorithm output is set to 50,000 units for every participant in case participants perceive this to be bad news. This results in a scenario where an aversion towards algorithms is more likely to be present. Second, case information in the experiment gives no indications that any chosen number lower than 50,000 or higher than 70,000 units makes sense or is useful to the fictional company. Instead, given information solely points out different reasons why the demand forecast should be as precise as possible. Still, three participants decided to choose 30,000, 40,000 and 80,000 units, respectively. They reported ‘own experience’ or ‘convincing investors’ as other factors that had influence on their decision. Therefore, absolute numbers might reflect more confounding factors than, for example, a relative weight that participants assign to the available advice sources.

but more detailed than in Chen et al. (2022).[9] In contrast to Chen et al. (2022), I do not use perceived credibility as dependent variable as it does not directly reflect the actual use of an algorithmic advice source. However, findings related to perceived credibility will be presented as a part of supplemental analyses in Sect. 4.2.2.

3.1.5 Exclusion of participants

In order to assure high data quality, a two-fold attention check was set in place. After reading the case information, participants had to answer attention check question one (Q1). Q1 was a multiple-choice question with three possible answers that asked for the type of product involved in the case. 299 participants answered correctly. Attention check question two (Q2) was set in place after participants had read the information about the data scientist and the algorithm. Q2 was a multiple-choice question with three possible answers that were all true. Q2 asked which kinds of data were used by the algorithmic models. On average, 2.45 out of 3 correct crosses were set, resulting in 196 participants who have answered both questions correctly. Reasons for the high number of incorrect answers are difficult to determine. Although Q1 was answered almost entirely correctly and intrinsic motivation of Prolific workers tends to be high,[10] it seems possible that either participants were not sufficiently attentive or Q2 had limitations. Potential limitations are: First, one possible answer was paraphrased slightly different in the previously given text.[11] Second, the question was set in place as part of the 'inner workings' manipulation, meaning that information complexity around the relevant information differed for participants. Third, participants might have become skeptical and insecure as they might not have expected all answers to be true and, therefore, tended to choose only one or two options. As it cannot be ruled out that the participants were not sufficiently attentive, Q1 and Q2 are used to exclude 104 participants who answered this question incorrectly, resulting in the main sample of 196 participants. A comparison of timestamps suggests that these participants were clearly engaged with the task as the ones in the 'inner workings disclosed' condition took significantly more time to read through the manipulated information on algorithm's inner workings than those in the 'inner workings not disclosed' condition (153.34 vs. 91.47 seconds, $p=0.008$).

[9] While Chen et al. (2022) focus solely on the credibility of the advice source to measure algorithm reliance, I asked participants not only to report their used weighting, but also for a more detailed assessment on credibility, i.e., if they find the source 'trustworthy' and 'competent'.

[10] Farrell et al. (2017) show that online workers have comparably high intrinsic motivation and tend to show a high effort level. This can also be observed in this experiment with regard to the final question participants could answer. They were asked whether they have guesses about the purpose of the study. Around half of the participants answered this question, even though it was optional and apparently the last question of the study.

[11] The case information referred to 'similar products' whereas the attention check question used 'other products' as a possible answer.

3.2 Experiment 2

3.2.1 Rationales for experiment 2

As pointed out in Sect. 2.2, the XAI literature finds several possibilities to make complex algorithms more explainable to increase trust in and, correspondingly, the reliance on algorithms (Adadi & Berrada, 2018). Considering information disclosure to be one of these possibilities, experiment 1 focuses on comparing the effects of disclosing information about the algorithm's inner workings with providing no additional information about the algorithm.

However, instead of this dichotomous conceptualization, information disclosure can be considered as a more continuous construct with an optimal amount of information. Various research streams consider partial information disclosure to be more promising for decision quality than full information disclosure. For example, Stocks and Harrell (1995) find that especially individuals struggle with a high amount of different information and tend to use it less thoroughly and less consistently in an evaluation setting involving financial distress. The feedback and optimal information disclosure literature argues that a high degree of information disclosure is suboptimal for stock markets as it reduces individual information acquisition endeavors (Gao & Liang, 2013).

These findings relate to the idea that there is an inverted U-curve for the effect of information environment complexity on the level of information processing (Schroder et al., 1967) that is interesting from an information science (Hwang & Lin, 1999) as well as from an accounting perspective (Libby, 1981; Stocks & Harrell, 1995). Especially in the context of the digital transformation, the high degree of information availability can easily result in an information overload of decision-makers (Roetzel, 2019). In this regard, the relevance of information may be important as well. While irrelevant information might lead to lower decision quality, an increasing amount of relevant information has a consistently positive, albeit diminishing, effect on decision quality (Hartmann & Weißenberger, 2023; Iselin, 1996). Thus, it might be essential to consider a level of information disclosure that does not overwhelm decision-makers and actually results in a better understanding of algorithmic decision aids.

In order to take these ideas into account, a second experiment, described in the following sections, investigates different levels of information. As a further specification of disclosing an algorithm's inner workings, experiment 2 offers more detailed insights, with the potential to provide additional support for H1.

3.2.2 Participants

I conducted another online experiment with 90 participants recruited from *Prolific*. Equivalent to the first experiment, the only precondition for all participants was that they were German native speakers. The demographic data of the participants is comparable to experiment 1 in terms of age (m = 32.94, sd = 11.69) and gender (male: 58; female: 31; diverse: 1). Compared to experiment 1, the proportion of students is lower in the experimental population (employed: 53; self-employed: 7; student: 19;

other: 11). Median completion time of the experiment was close to 9 min for which the participants received £ 1.60, resulting in an average hourly wage of £ 10.67.

In this experiment, it is not necessary to exclude any participants from the data analysis, as both attention check questions were answered entirely correctly.[12]

3.2.3 Experimental task

Regarding the experimental setting, the second experiment was mainly based on the first experiment, with a specific focus on the idea to further differentiate levels of information about the algorithm in a similar forecasting decision. In addition, a few changes were made in order to gain further insights:

Due to several design choices, experiment 1 provided data for only one decision per participant and did not randomize between human and algorithm advice direction. Preventing confounding effects, the second experimental setting included forecast decisions for four different regions in order to examine whether the results are robust across slightly different sets of numbers. This might be crucial since differences between human and algorithm advice may be considered as bad news or good news according to Chen et al. (2022). The forecast values of the algorithm (product manager) were alternated with a different factor in each case so that the range between the two values remained the same in relative terms: 50,000 (70,000); 35,000 (25,000); 35,000 (49,000); 91,000 (65,000) for regions 1–4.

3.2.4 Independent and dependent variables

As mentioned before, the second experiment focused on the disclosure of algorithm information in a more nuanced way than the first experiment. Priming thinking style was not included in the second experiment due to possible confounding factors that will be outlined in Sect. 4.1 and 5. Therefore, I relied on a 3×1 between-subjects format to manipulate the independent variable 'information level' with regard to an algorithm's inner workings, sharing a 'detailed', 'medium' and 'low' level of insights with the participants.

To implement considerable differences between the manipulations, the following adjustments compared to experiment 1 were made:[13]

- In the 'detailed' treatment, information about the development of the algorithm was presented including highly detailed examples. Furthermore, deliberate re-

[12] The first attention check question was taken from the first experiment. Due to the potential issues of the second attention check question (see Sect. 3.1.5), the question was replaced with another common way to check for inattentive participants – a seven-point Likert-type format question where participants were instructed to check a certain field.

[13] Comparing the first and second experiment with each other, the 'inner workings disclosed' condition can be situated in between the 'detailed' and the 'medium' treatment, while the 'inner workings not disclosed' condition is slightly above the 'low' treatment in terms of level of information. The design choices ensure that the 'detailed' and 'medium' condition can both be seen as treatments in which the inner workings are disclosed to the participants – involving different loads of information – while the 'low' condition is comparable with the 'inner workings not disclosed' condition.

dundancies and short rephrasing summaries were integrated to ensure a certain learning effect.
- The 'medium' treatment also informed participants about the most important inner workings of the algorithm, but relied on much shorter explanations and did not include any examples or repetitions.
- While the product manager and the data scientist were presented equally here as well, the 'low' treatment solely mentioned that the data scientist makes an assessment with the help of a self-developed algorithm without sharing any further information.

The manipulation of information load across the three treatments was consistent, with the number of words being varied by roughly the same amount in each condition. Following the ideas of the first experiment, participants were asked to report the weight they assigned to the algorithmic advice for each of the four decisions. The separate decisions enable additional analyses, while the average weight on algorithmic advice across all decisions is used as the main dependent variable.

4 Results

4.1 Main analyses

4.1.1 Test of hypothesis 1

H1 predicts a main effect of disclosing an algorithm's inner workings on the weight decision-makers place on the algorithmic advice source. To formally test this hypothesis, I conduct several statistical analyses.[14]

An ANOVA, as commonly used for comparable 2×2 factorial designs (Kirk, 2009), is presented in Table 1. Results show a slight tendency of disclosing an algorithm's inner workings on the weight on algorithmic advice (m = 56.84 vs. 55.67, $p = 0.597$). This difference is insignificant. The expected direction cannot be identified in the full sample ($n = 300$) either, due to almost identical mean values of the dependent variable (m = 55.00 vs. 55.41, $p = 0.819$).

Analyses of experiment 2

In order to better understand the results and to investigate Hypothesis H1 with greater detail, I conducted another experiment focusing specifically on the effect of sharing different levels of information with regard to an algorithm's inner workings. The second experiment takes into account that there might be varying effects due to the potential non-linearity between the amount of information and algorithm reliance and, simultaneously, leaves out the priming part as a potential source of interference. The focus on the inner workings concept and the slight changes in the experimental design help to emphasize the effect of disclosing an algorithm's inner workings on the weight on algorithmic advice. In order to check whether the manipulation has worked

[14] All tests are two-tailed due to a potential bidirectionality of the effect of disclosing an algorithm's inner workings on algorithm reliance.

Table 1 Effect of disclosing an algorithm's inner workings and thinking style on weight on algorithmic advice (ANOVA) in the main sample in experiment 1 (n = 196)

Panel A: *Means and standard deviations for weight on algorithmic advice*

Thinking style	Inner workings	Mean	Std. Deviation	N
Intuitive thinking	Not disclosed	56.04	13.841	48
	Disclosed	56.80	18.180	49
	Total	56.42	16.099	97
Analytic thinking	Not disclosed	55.24	13.605	41
	Disclosed	56.88	16.145	58
	Total	56.20	15.094	99
Total	Not disclosed	55.67	13.660	89
	Disclosed	56.84	17.024	107
	Total	56.31	15.560	196

Dependent variable: weight on algorithmic advice – Participants had to adjust two sliders (weight on product manager forecast / weight on data scientist forecast) so that they add up to 100%.

Panel B: *ANOVA*

Factor	Type III Sum of Squares	df	Mean Square	F	p-value	Partial Eta Squared
Inner workings	68.906	1	68.906	0.281	0.597	0.001
Thinking style	6.158	1	6.158	0.025	0.874	0.000
Inner workings x Thinking style	9.369	1	9.369	0.038	0.845	0.000
Error	47129.592	192	245.467			

ANOVA results indicate that neither the disclosure of the algorithm's inner workings ($p = 0.597$, $\eta^2 = 0.001$) nor thinking style ($p = 0.874$, $\eta^2 = 0.000$) have a significant effect on the weight on algorithmic advice. The interaction effect is also non-significant ($p = 0.845$, $\eta^2 = 0.000$), suggesting that thinking style does not moderate the effect of disclosure.

properly, participants had to answer several questions with regard to the algorithm information they received. Asking if algorithm information was disclosed (m = 6.07 vs. 5.91 vs. 2.76), whether plenty of information was given (m = 6.17 vs. 6.09 vs. 2.93), how complete the information was perceived (m = 5.59 vs. 5.19 vs. 3.10) and how sound the information was perceived (m = 5.86 vs. 5.34 vs. 3.31) revealed significant differences between the 'detailed', 'medium' and 'low' treatment.[15]

Since the 'medium' and the 'detailed' treatment both disclosed the most relevant inner workings of the algorithm, an ANOVA using the weight on algorithmic advice might provide further support for H1. As can be seen in Table 2, Panel A and B, the effect of disclosing an algorithm's inner workings has a highly significant effect on algorithm usage (m = 58.51 vs. 59.94 vs. 47.17, $p = 0.005$).[16] More specifically,

[15] The questions had a seven-point Likert-type format. Differences between the 'low' and 'detailed' level as well as between the 'low' and 'medium' level are highly significant ($p < 0.001$).

[16] This effect is not only highly significant, but also notably large (Cohen, 1988). Compared to other offline and online experiments in the algorithm reliance domain, such as Dietvorst et al. (2015) that report effect sizes of η^2 between 0.024 and 0.069, or Logg et al. (2019) with η^2 between 0.01 and 0.08 in large sample sizes of $n > 300$ participants. With $\eta^2 = 0.113$, $p < 0.01$ and a sample size of $n = 90$, experiment 2 has

Table 2 Effect of level of information on weight on algorithmic advice (ANOVA) in experiment 2

Panel A: *Means and standard deviations for weight on algorithmic advice*

Information level	Mean	Std. Deviation	N
Low	47.17	16.404	29
Medium	59.94	16.282	32
High	58.51	15.751	29
Total	55.36	16.962	90

Dependent variable: weight on algorithmic advice – Participants had to adjust two sliders (weight on product manager forecast / weight on data scientist forecast) so that they add up to 100%.

Panel B: *ANOVA*

	Sum of Squares	df	Mean Square	F	p-value	Partial Eta Squared
Information level	2905.932	2	1452.966	5.568	0.005	0.113
Error	22700.721	87	260.928			
Total	25606.653	89				

ANOVA results confirm a significant main effect of information level on the weight on algorithmic advice ($p=0.005$, $\eta^2 = 0.113$).

Panel C: *Post-hoc Tukey HSD test*

	(I) Information level	(J) Information level	Mean Difference (I-J)	Std. Error	p-value	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Low	Medium	-12.772	4.141	0.008	-22.648	-2.897
		Detailed	-11.344	4.242	0.024	-21.459	-1.229
	Medium	Low	12.772	4.141	0.008	2.897	22.648
		Detailed	1.428	4.141	0.937	-8.447	11.303
	Detailed	Low	11.344	4.242	0.024	1.229	21.459
		Medium	-1.428	4.141	0.937	-11.303	8.447

Based on observed means.

The error term is Mean Square (Error) = 260.928.

The Tukey HSD test further clarifies the effect of information level, showing that the significant differences stem from the contrast between the 'low' condition and both the 'medium' ($p=0.008$) and 'detailed' ($p=0.024$) conditions.

this effect can be traced back to the differences between the 'low' and the 'medium' level ($m=47.17$ vs. 59.94, $p=0.008$) as well as between the 'low' and the 'detailed' level ($m=47.17$ vs. 58.51, $p=0.024$) with the help of a post-hoc Tukey HSD test (Table 2, Panel C). These results are robust for several ways to measure the dependent variable.[17]

uncovered a comparably large and highly significant effect of disclosing an algorithm's inner workings on algorithm reliance.

[17] The results are consistent for all of the single regional decisions except for region 2. Region 2 decisions do not produce significant differences between the three groups ($m=49.45$ vs. 58.91 vs. 53.31, $p=0.188$). This might be due to the fact that region 2 has the narrowest absolute span between both advices, possibly affecting the participants' perception of advice source differences. This indicates that the different sets of numbers might have different effects on algorithm reliance. I elaborate on this in supplemental analyses (Sect. 4.2.4).

While the disclosure of relevant inner workings results in a significantly increased algorithm reliance, the findings indicate that there is a tendency of the 'detailed' treatment to lead to marginally lower algorithm reliance, implying a concave relationship. This suggests that there might be an optimum of information load as too much information might lead to cognitive challenges (information overload) or even increase distrust in some algorithmic mechanisms. Still, the difference between the 'detailed' and the 'medium' treatment is insignificant ($p=0.937$).

Thus, the results of experiment 2 provide support for H1 and indicate that the relationship between information load and algorithm reliance may indeed be of a concave nature.

To disentangle these divergent findings, I further analyze the data from experiment 1. Possible explanations might be connected to the priming techniques used in the first experiment, as the robust and significant results of experiment 2 indicate. Therefore, the following considerations present a manipulation check and rather explorative analyses based on a restricted sample of experiment 1.

Manipulation check of experiment 1

In order to verify possible limitations of the priming approach, the manipulation of 'thinking style' requires further attention in the form of a manipulation check. The mean score in the word completion task is 14.43 out of 15 words. Most mistakes clearly stem from spelling mistakes. Accordingly, there are no indications that participants did not process the target words as expected, which is a precondition for the priming approach. In order to analyze if the priming was effective, I asked several questions regarding the participants' decision manner. After the decision, participants had to answer six seven-point Likert-type scale questions. They had to assess their own decision in terms of different criteria. I asked to which degree they made their decision in a rational, random or slow manner by using different wording, for example: "I made my decision in a slow manner." and "I made my decision in a fast manner.". Differences between the intuitively and analytically primed groups are statistically not significant in the main sample. Tendency-wise, the priming approach even had counterintuitive effects on participant behavior, e.g., a potentially more rational approach of intuitively primed vs. analytically primed subjects ("I made my decision in a conscientious manner.", $m=5.92$ vs. 5.76, $p=0.227$). This suggests that the priming approach did not work properly in the main sample which might explain parts of the unexpected findings.

Robustness checks and explorative analyses based on a restricted sample

Due to the manipulation check results, I perform further analyses on a restricted sample with 132 participants, applying even stricter quality filters. Next to the attention check criterion, time is a crucial aspect as the experiment involves priming techniques. In order to ensure that the priming approach works appropriately, data collection should be following the temporary activation of concepts as soon as possible (Bargh & Chartrand, 2014). These thoughts are supported by a meta-analysis

Also, I calculated the actual weight on algorithm advice participants used in their decisions (e.g., 55,000 in region 1 equals a 75% actual weight on algorithm advice). Using the average of these actual decision weights in an ANOVA, results remain consistent except for the tendency that the 'detailed' treatment actually results in a slightly higher algorithm reliance than the 'medium' treatment ($m=47.46$ vs. 58.89 vs. 59.14, $p<0.001$).

on priming by Weingarten et al. (2016), stating that especially direct perception-behavior effects as present in this study are highly time-sensitive. In effective priming studies, the intended measurement is performed very shortly after the priming occurred (e.g., Albarracin & Hart, 2011; Chartrand & Bargh, 1996; Fitzsimons et al., 2008; Weingarten et al., 2016). Therefore, participants who took a break or were exceptionally slow might affect results in an unexpected way. Besides, participants who were exceptionally fast might not have read information about the case and, especially, about the algorithm's inner workings carefully. Ensuring the 'inner workings' manipulation to work as expected, an exclusion of 'speed-runners' therefore seems logical. Since median completion time of the 196 remaining participants was ≈ 14 minutes (sd ≈ 8), I choose a strict inclusion boundary of 10 to 18 minutes for further tests.[18] Validating these boundaries with the previously mentioned manipulation check, a t-test in the restricted sample reveals that participants in the analytic thinking condition reported a slightly lower decision speed than those in the intuitive thinking condition ("I made my decision in a slow manner.", m = 3.28 vs. 2.95, p = 0.097). This self-assessment can be confirmed by analyzing experimental time stamps: Participants in the analytic thinking condition took significantly more time to submit their decision than participants in the intuitive thinking condition (m = 35.04 vs. 30.45 seconds, p = 0.084). Furthermore, there is evidence that the priming did not work as intended beyond the defined boundaries: First, for participations over 18 minutes, a non-significant and even counter-intuitive effect between the analytic thinking and intuitive thinking treatments can be found in terms of self-reported decision time (m = 3.47 vs. 3.78, p = 0.226). Second, for participations under 10 minutes, analytically primed participants reported significantly lower decision times than intuitively primed participants (m = 2.63 vs. 3.40, p = 0.037), contradicting the idea of analytic thinking.[19] The real decision times show similar but slightly significant tendencies. Therefore, the priming approach appears effective in the restricted sample only.[20]

[18] Dropping response time outliers dependent on standard deviation is a common method (Ratcliff, 1993; Vankov, 2023), especially in time-sensitive studies. I chose a rather strict inclusion criterion of 0.5 standard deviations around the median completion time as both the treatments in this study are specifically sensitive to completion time outliers. For the inner workings manipulation, it is especially important that participants take their time to read and incorporate the information (lower boundary) and for the thinking style manipulation to make decisions shortly after the priming was performed (upper boundary). Another argument for the lower boundary is that a certain level of thoroughness is required in order to (subconsciously) perceive the priming concepts.

[19] Another possible confound with regard to thinking style might be time pressure. Inducing time pressure might lead to a shift to intuitive thinking, as Evans & Curtis-Holmes (2005) show. While there was no direct time pressure implemented in this experiment, some participants might feel a certain degree of time pressure to finish the experiment quickly in order to maximize their own outcome.

[20] The priming approach was also ineffective in the full sample as there are insignificant differences in decision times (m = 34.01 vs. 33.76 s, $p = 0.463$) and a contrary effect for the self-reported decision speed (m = 3.08 vs. 3.32, $p = 0.086$) due to the inconsistent participant behavior in the lower and upper time segment. In general, the effects of priming are highly subtle and therefore rather difficult to measure (Bargh & Chartrand, 2014). For this reason, other experimental research already regards the priming approach to be successfully implemented if, for example, participants consider words used for priming to be related to the targeted concept (Andrejkow et al., 2022). For the task used in this research, White (2005) and Huntsinger (2011) ensured that this relation between target words and target concepts is given.

Further analyzing the restricted sample, a t-test between the two experimental treatments 'inner workings disclosed' and 'inner workings not disclosed' shows an slightly significant difference in weight on algorithmic advice (m = 59.36 vs. 55.29, $p = 0.146$). Table 3 repeats the ANOVA analyses in the restricted sample ($n = 132$). Panel B reveals an insignificant, but clearer effect of disclosing an algorithm's

Table 3 Effect of disclosing an algorithm's inner workings and thinking style on weight on algorithmic advice (ANOVA) in the restricted sample ($n = 132$) in experiment 1

Panel A: *Means and standard deviations for weight on algorithmic advice*

Thinking style	Inner workings	Mean	Std. Deviation	N
Intuitive thinking	Not disclosed	53.53	13.356	32
	Disclosed	57.70	19.366	33
	Total	55.65	16.683	65
Analytic thinking	Not disclosed	57.10	14.372	31
	Disclosed	60.89	16.899	36
	Total	59.13	15.779	67
Total	Not disclosed	55.29	13.870	63
	Disclosed	59.36	18.057	69
	Total	57.42	16.263	132

Dependent variable: weight on algorithmic advice – Participants had to adjust two sliders (weight on product manager forecast / weight on data scientist forecast) so that they add up to 100%.

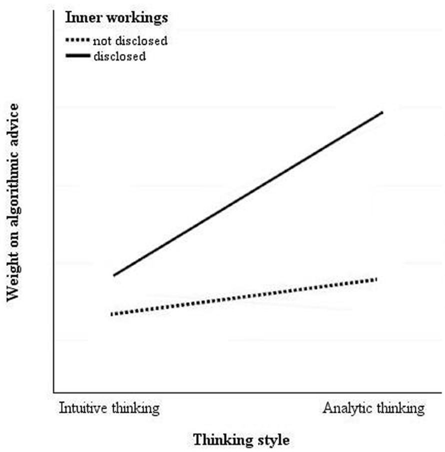
Panel B: *ANOVA*

Factor	Type III Sum of Squares	df	Mean Square	F	p-value	Partial Eta Squared
Inner workings	520.830	1	520.830	1.977	0.162	0.015
Thinking style	375.553	1	375.553	1.425	0.235	0.011
Inner workings x Thinking style	1.148	1	1.148	0.004	0.948	0.000
Error	33723.204	128	263.463			

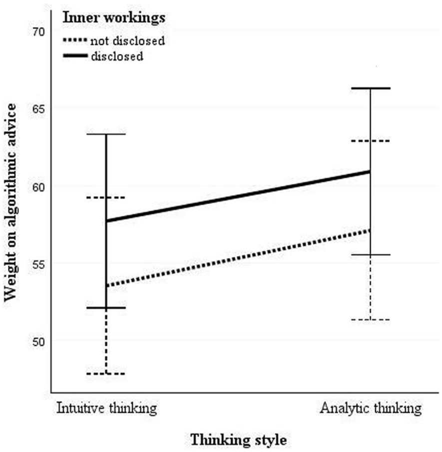
ANOVA results remain non-significant. However, the p-values for both, a main effect of inner workings ($p = 0.162$, $\eta^2 = 0.015$) and of thinking style ($p = 0.235$, $\eta^2 = 0.011$), indicate a slight shift towards potential effects compared to the main sample. The interaction effect remains negligible ($p = 0.948$, $\eta^2 = 0.000$), suggesting that there is no meaningful moderation by thinking style.

Panel C: *Predictions vs. results of weight on algorithmic advice*

Predictions



Results



While theoretical predictions expected a main effect of inner workings and an interaction effect with thinking style, the data in the restricted sample indicate a tendency toward a main effect of both manipulated variables, as illustrated on the right side with 95% confidence intervals.

inner workings on the weight on algorithmic advice compared to the main sample (m = 59.36 vs. 55.29, $p = 0.162$).[21] Another frequently used method to test for differences in experimental conditions is contrast coding (Buckless & Ravenscroft, 1990; Guggenmos et al., 2018). By using a [-1, 1, -1, 1] contrast (see also Table 4), the main effect of disclosing an algorithm's inner workings (H1) is, consistent with the ANOVA results, slightly significant (t = 1.417, $p = 0.160$).

4.1.2 Test of hypothesis 2

H2 addresses the potential interaction effect of analytic thinking on the relationship between disclosing an algorithm's inner workings on algorithm reliance. In order to address H2, I first refer to the ANOVA results (see Table 1, Panel B). They indicate that no significant interaction effect of both independent variables on the weight on algorithmic advice can be found in the main sample ($p = 0.845$). Considering the graphical presentation of the ANOVA results in the restricted sample (Table 3, Panel C), a potential direct effect of both manipulated variables – inner workings and thinking style – seems plausible. While this tendency will be examined in supplemental analyses, a significant interaction effect cannot be identified. A further analysis using a [1, 1, 1, -3] contrast, which tests whether the mean for the inner workings disclosed / analytic thinking condition differs from the average of all other cell means, also yields an insignificant result (t = -1.471, $p = 0.146$). The results are reported in Table 4. The lack of significant evidence suggests that any potential interaction effect is weak or inconsistent.[22] However, the following section provides analyses of a possible joint effect of disclosing an algorithm's inner workings and thinking style that might be present after all.

4.2 Supplemental analyses

4.2.1 Joint effect of disclosing an algorithm's inner workings and thinking style

In order to gain a deeper understanding of the manipulated concepts and their effect on algorithm reliance, I perform additional analyses with the restricted sample.[23]

[21] Despite the insignificant results, the analyses of the restricted sample show more support for the theoretical predictions. However, the effect size measure partial eta squared ($\eta^2 \sim 0.01$) is considered very small (Cohen, 1988). While some priming experiments have found significant results with comparable effect sizes, such as experiment 3 by Wolfe et al. (2020) with $\eta^2 = 0.013$, the observed effect remains marginal. Bayesian statistics following Derks et al. (2025) and Jeffreys (1961) yield a Bayes factor of $BF_{10} = 0.065$, providing evidence against a meaningful effect. Moreover, post-hoc power analyses indicate that, given the current sample size ($n = 132$), only effects of $\eta^2 \geq 0.045$ would have been detectable at a significance level of 10%, with a power of 80%. This suggests that the absence of a statistically significant effect could be due to the limited power of the study rather than the absence of a true effect. These considerations highlight potential limitations of the priming approach within this experiment. This deduction is further supported by comparable analyses for experiment 2 (see footnote 16).

[22] There is also no significant effect in the full sample (t = 0.142, $p = 0.888$).

[23] As the priming approach appears effective only in the restricted sample, the supplemental analyses of experiment 1 (Sect. 4.2.1 to 4.2.3) are based on this sample. Any qualitative differences between the samples will be indicated.

Table 4 Effect of disclosing an algorithm's inner workings and thinking style on weight on algorithmic advice (contrast analysis) in the restricted sample (n = 132) in experiment 1

Panel A: *Contrast Coefficients*

Contrast	Condition: Inner workings not disclosed / intuitive thinking	Inner workings disclosed / intuitive thinking	Inner workings not disclosed / analytic thinking	Inner workings disclosed / analytic thinking
a	-1	1	-1	1
b	1	1	1	-3
c	-1	0	0	1
d	-2	1	-1	2

Dependent variable: weight on algorithmic advice – Participants had to adjust two sliders (weight on product manager forecast / weight on data scientist forecast) so that they add up to 100%.

Contrast a: The [-1, 1, -1, 1] contrast tests for a main effect of disclosing an algorithm's inner workings (H1).

Contrast b: The [1, 1, 1, -3] contrast tests for a difference between the inner workings disclosed / analytic thinking condition and the average of all other cell means (H2).

Contrast c: The [-1, 0, 0, 1] contrast tests for a difference between the inner workings not disclosed / intuitive thinking and the inner workings disclosed / analytic thinking condition.

Contrast d: The [-2, 1, -1, 2] contrast tests for the order that both manipulations (analytic thinking and disclosing inner workings) have a stronger effect on the use of algorithms than only manipulating one of both concepts.

Panel B: *Contrast Tests*

	Contrast	Value of Contrast	Std. Error	t	df	p-value	95% Confidence Interval Lower	95% Confidence Interval Upper
Weight on algorithmic advice	a	7.96	5.616	1.417	119.579	0.160	-3.16	19.08
	b	-14.34	9.747	-1.471	59.311	0.146	-33.84	5.16
	c	7.36	3.675	2.002	65.150	0.049	0.02	14.70
	d	15.32	8.489	1.804	103.178	0.074	-1.52	32.15

H1: Contrast a supports the trends observed in the previous ANOVA (Table 3), indicating a tendency toward two main effects.

H2: In contrast, Contrast b remains insignificant, further suggesting the absence of a meaningful interaction effect.

Supplemental analyses: Contrast c and d suggest a positive joint effect of disclosing an algorithm's inner workings and analytic thinking on algorithm reliance.

Since the graphical representation of the results and the mean comparisons in the restricted sample of experiment 1 indicate a possible main effect for both investigated variables (Table 3, Panel A and C),[24] another contrast might help to pinpoint the relations present in experiment 1. A [-1, 0, 0, 1] contrast tests for a difference between the inner workings not disclosed / intuitive thinking and the inner workings disclosed / analytic thinking condition while simultaneously taking no notice of the other conditions. This approach is useful because it compares the situation where both variables are manipulated towards enhancing algorithm reliance versus the situation where both variables are set to the 'standard' initial situation involving a low availability of algorithm information and intuitive thinking presumed as the default approach. The contrast is statistically significant at the 5% level ($t=2.002$, $p=0.049$), implying a joint effect of both variables on the weight on algorithmic advice.

In addition, as I suggest that thinking style reinforces the effect of disclosing an algorithm's inner workings on algorithm reliance, a certain outcome order of the four experimental groups may be interesting to look at. A contrast test with a [-2, 1, -1, 2] specification allows to test for this order that is also apparent in the mean comparison. Assuming the 'standard' situation described above to be the starting point, the manipulation of one of both variables considered results in a weaker but still positive effect on the use of algorithms. The contrast is statistically significant at the 10% level ($t=1.804$, $p=0.074$).[25]

Comparing these findings with previous results (see Table 3), I suggest that disclosing an algo-rithm's inner workings and analytic thinking taken together have a significant and positive joint effect on the use of algorithms. Considering the separate effects of information disclo-sure and analytic thinking, the potential lack of statistical impact in ANOVA approaches might stem from several factors like a general tendency of weighting both advice sources equally (50:50), the subtle influence of priming and other experimental design choices that will be further discussed in section 5. Another explanation for the somewhat ambiguous findings concerning thinking style might stem from complex underlying mechanisms that drive a decision-makers tendency to rely on algorithms. These mechanisms will be examined in the next sections.

4.2.2 Control variables and underlying mechanisms

As outlined in Sect. 2.2, algorithm reliance might be heavily driven by perceived advice source credibility, i.e., perceived competence and perceived trustworthiness (Chen et al., 2022; Petty & Cacioppo, 1986). Thus, I intend to identify mechanisms that may underlie the reliance on an algorithmic advice with the help of several additional data gathered during the experiment. This is why I asked participants after their forecast decision to which degree they find the product manager, the data scien-

[24] A mean comparison of the experimental groups shows: The highest difference in terms of weight on algorithmic advice is between the inner workings disclosed / analytic thinking and the inner workings not disclosed / intuitive thinking condition ($m=60.89$ vs. 53.53, $sd=16.899$ vs. 13.356).

[25] Due to the almost identical mean values of weight on algorithmic advice, the contrast analyses results are insignificant in the main sample as well as in the full sample.

tist and the algorithm itself competent and trustworthy, resulting in six seven-point Likert-type scale questions analyzed in the following section.

Chen et al. (2022) consider perceived competence of product manager and data scientist to be the major driver in the given task. This is because the two advice sources should rather differ in competence than in general trust as they have similar incentives to aim for accuracy and therefore may be perceived equally trustworthy (Chen et al., 2022). This assumption can be validated based on a regression analysis (model 1) that investigates the effects of credibility elements on algorithm reliance (see Table 5). Perceived product manager (data scientist) competence has a strongly significant negative (positive) effect on algorithm usage (B = -14.595 (11.849), $p<0.001$), while trustworthiness is not significantly related to algorithm reliance. This experiment offers an even more nuanced perspective on advice source cred-

Table 5 Regression analysis of advice source credibility on weight on algorithmic advice in the restricted sample (n = 132) in experiment 1

Panel A: *Model Summary*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.635	0.403	0.374	12.867

Panel B: *ANOVA*

Model		Sum of Squares	df	Mean Square	F	p-value
1	Regression	13950.909	6	2325.152	14.044	0.000
	Residual	20695.174	125	165.561		
	Total	34646.083	131			

The regression model 1 examines the impact of advice source credibility measures on the weight on algorithmic advice and is highly significant (F = 14.044, p < 0.001).

Panel C: *Coefficients*

Model		Unstandardized Coefficients		Standardized Coefficients	t	p-value
		B	Std. Error	Beta		
1	(Constant)	59.154	10.764		5.495	0.000
	Product manager competence	-14.595	2.812	-0.621	-5.190	0.000
	Product manager trustworthiness	-1.577	2.442	-0.075	-0.646	0.520
	Data scientist competence	11.849	2.966	0.538	3.995	0.000
	Data scientist trustworthiness	-0.503	2.425	-0.026	-0.208	0.836
	Algorithm competence	-0.522	1.445	-0.036	-0.361	0.718
	Algorithm trustworthiness	5.016	1.335	0.347	3.756	0.000

Dependent variable: Weight on algorithmic advice – Participants had to adjust two sliders (weight on product manager forecast / weight on data scientist forecast) so that they add up to 100%.
Independent variables: Participants were asked to which degree they find the product manager, the data scientist and the algorithm itself competent and trustworthy.

ibility, since I not only asked for an assessment of the product manager and the data scientist, but for the algorithm itself as well. An interesting finding in this context is that for the evaluation of the algorithm itself, trustworthiness seems to replace competence as main driver for the ultimate algorithm reliance. Even if it does not seem common to assess an algorithm with humane criteria like competence or trustworthiness, the results raise the question of how detailed an advice source usage should be analyzed to learn more about underlying decision-making processes. More specifically, the distinction between the algorithm itself and the person responsible for it (e.g., data scientist, programmer) seems to be relevant for the perceived credibility and therefore the usage of an algorithmic decision aid. The fact that, in this experiment, a human employee is emphasized to be the person responsible for an algorithmic output, might have a severe effect on a general tendency of algorithm appreciation or at least a more neutral algorithm perception.

To gain a deeper understanding on the concepts of competence and trustworthiness in this context, I perform regression analyses on data scientist competence, product manager competence and algorithm trustworthiness. This is consistent with the previously mentioned mechanisms (see Table 5), as these sub-concepts of advice source credibility have a significant effect on the usage of the algorithmic advice.[26] All possible control variables are added to these regressions in order to test for effects that stem from individual participant characteristics.[27] While the regression analysis on product manager competence is not significant, the ones on data scientist competence (model 2) and algorithm trustworthiness (model 3) presented in Table 6 are significant. They are presented in Table 6.

The regression on data scientist competence (model 2) shows a significant interaction of the manipulated variables inner workings and thinking style ($B=0.525$, $p=0.031$). This indicates that, as opposed to the findings presented in previous analyses, there might be a moderating effect after all. This effect refers to the data scientist competence that, in turn, affects the weight on algorithmic advice (model 1). However, the main effect coefficients of inner workings and thinking style are negative which is a contradiction to theory as they indicate a negative effect of disclosing an algorithm's inner workings and analytic thinking on the perceived competence of the data scientist. Along with the interaction effect, this might represent the positive joint effect of both manipulated variables that is identified in Sect. 4.2.1.

Alternatively, the finding above could point to a moderated mediation effect. In order to analyze this idea, I use the *PROCESS* module for SPSS (Hayes, 2018). I rely on *Model 7* of Hayes' *PROCESS* module with a bootstrap sample of 5,000 replications to test for a mediated moderation. Prior studies indicate that bootstrapping is

[26] In the full sample, product manager trustworthiness ($B = -2.854$, $p=0.061$) and data scientist trustworthiness ($B=2.280$, $p=0.079$) show slight effects with p-values of less than 10%. Due to the limitations of the full sample, I focus on the restricted sample for the following analyses.

[27] Various additional questions were asked in the experiment next to the usual demographics. Participants had to answer the 4-item Faith in General Technology and the 3-item Trusting Stance questionnaires (McKnight et al., 2011) as well as indicate their prior knowledge in terms of basic statistics and artificial intelligence, machine learning etc. in order to rule out alternative explanations. All questions were implemented with a seven-point Likert-type format. Furthermore, participants had to assess their risk preference on a scale from 0 to 10.

the recommended method for this type of analysis (Preacher & Hayes, 2008). The analysis confirms the regression results of model 2 (see Table 6) as the interaction effect of disclosing an algorithm's inner workings and analytic thinking on data scientist competence is significant ($p=0.003$). Data scientist competence, in turn, has a significant effect on the weight of algorithmic advice at the 10%-level ($p=0.093$). However, the confidence intervals for the indirect effect include zero, suggesting that there is no significant moderated mediation effect (-0.100; 7.795).[28] Nevertheless, the analysis shows some clear tendencies paving the way for a deeper understanding of the connections between algorithm information, thinking style and different shapes of perceived credibility in future studies.

Coming back to regression model 2, other factors like the Rational Experiential Inventory (REI)[29] rationality score (B = 0.151, $p=0.057$) and Faith in General Technology[30] score (B = 0.345, $p<0.001$) have a significant positive influence on perceived data scientist competence. Rational decision-makers and people generally in favor of technology, thus, seem to find a data scientist operating with algorithms rather competent.

The regression on algorithm trustworthiness (model 3) shows a few similar tendencies with regard to Faith in General Technology (B = 0.312, $p=0.02$) and the related concept of Trusting Stance (B = 0.209, $p=0.029$). Since algorithm trustworthiness is strongly connected to technology trustworthiness, the significant effect of the mentioned scales is not surprising. Also, a higher prior knowledge in basic statistics might help to gain trust in algorithms. According to the regression, there is a corresponding effect on algorithm trustworthiness that is significant (B = 0.146, $p=0.086$). Another factor that is relevant in many cases is risk preference. The regression analysis shows that a higher risk preference has a negative effect on algorithm trustworthiness (B = -0.104, $p=0.04$). A potential explanation for this somewhat counterintuitive finding is that, for example, decision-makers who are more willing to take risks might hope for the human expert to be experienced and more adapted to the business setting. In contrast, risk-averse decision-makers might consider algorithms to be better able to smooth out outliers, spikes etc. in predictions.[31]

[28] While the main sample yields qualitatively similar results, analyzing the full sample delivers a significant interaction effect of inner workings and thinking style on data scientist competence ($p=0.037$), a significant effect of data scientist competence on weight of algorithmic advice ($p=0.000$) and zero is not included in the relevant confidence interval (0.094; 3.691). Due to the limitations of the full sample, I further focus on the restricted sample.

[29] Detailed information about the REI can be found in Sect. 4.2.3.

[30] According to McKnight et al. (2011), the Faith in General Technology scale focuses on the level of trust individuals have in technology or information systems as a whole, rather than in a specific kind of technology. It aims to measure the general disposition of individuals to trust technology.

[31] Regarding the reported effects of model 2 und 3, the results based on the main sample and the full sample are qualitatively similar. Risk preference is an exception, as it is the only factor that is significant in the main sample but insignificant in the full sample.

Table 6 Regression analyses on advice source credibility in the restricted sample (n=132) in experiment 1

Panel A: *Model Summary*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
2	0.544	0.296	0.225	0.651
3	0.463	0.214	0.135	1.046

Panel B: *ANOVA*

Model	Sum of Squares	df	Mean Square	F	p-value	
2	Regression	21.177	12	1.765	4.169	0.000
	Residual	50.368	119	0.423		
	Total	71.545	131			
3	Regression	35.498	12	2.958	2.703	0.003
	Residual	130.222	119	1.094		
	Total	165.720	131			

Regression model 2 examines the impact of the manipulated variables and possible covariates on data scientist competence and is highly significant ($F=4.169$, $p<0.001$). Regression model 3 examines the impact of the manipulated variables and possible covariates on algorithm trustworthiness and is highly significant ($F=2.703$, $p=0.003$).

Panel C: *Coefficients*

Model		Unstandardized Coefficients		t	p-value
		B	Std. Error		
2	(Constant)	3.710	0.746	4.974	0.000
	Inner workings	-0.273	0.169	-1.615	0.109
	Thinking style	-0.162	0.170	-0.952	0.343
	Inner workings x Thinking style	0.525	0.240	2.187	0.031
	REI rationality score	0.151	0.078	1.920	0.057
	REI experientiality score	-0.019	0.082	-0.235	0.815
	Faith in General Technology score	0.345	0.083	4.177	0.000
	Trusting Stance score	0.053	0.059	0.906	0.367
	Gender	-0.051	0.116	-0.442	0.659
	Age	0.004	0.007	0.531	0.596
	Risk preference	-0.039	0.031	-1.255	0.212
	Prior knowledge in basic statistics	-0.013	0.052	-0.249	0.804

Table 6 (continued)

	Prior knowledge in AI, machine learning, algorithms	-0.022	0.055	-0.400	0.690
3	(Constant)	1.762	1.199	1.470	0.144
	Inner workings	-0.074	0.271	-0.273	0.785
	Thinking style	-0.189	0.273	-0.692	0.490
	Inner workings x Thinking style	0.087	0.386	0.227	0.821
	REI rationality score	0.156	0.126	1.236	0.219
	REI experientiality score	0.075	0.131	0.569	0.570
	Faith in General Technology score	0.312	0.133	2.349	0.020
	Trusting Stance score	0.209	0.095	2.206	0.029
	Gender	-0.067	0.186	-0.358	0.721
	Age	0.000	0.011	-0.043	0.966
	Risk preference	-0.104	0.050	-2.072	0.040
	Prior knowledge in basic statistics	0.146	0.084	1.731	0.086
	Prior knowledge in AI, machine learning, algorithms	-0.036	0.088	-0.416	0.679

Dependent variables: Data scientist competence (model 2), algorithm trustworthiness (model 3).

Independent variables: Manipulated variables as well as possible covariates.

4.2.3 Rational experiential inventory

As outlined in Sect. 2.3, people's approach to make decisions differs depending on the thinking style currently activated. While humans are generally able to engage in intuitive and analytic thinking (Evans, 2003; Stanovich & West, 2008), individual characteristics are one influential factor of thinking style (Frederick, 2005). The following section therefore deals with the so-called REI. After participants made their forecast decision and answered several corresponding questions, they had to complete the 40-item REI according to Epstein et al. (1996) and Pacini and Epstein (1999). The REI measures people's characteristics in terms of rationality (20 items) and experientiality (20 items) which are used as synonyms for analytic and intuitive thinking. These two subscales exist as it is theoretically possible that individuals are capable of and comfortable with both thinking styles. While the priming manipulation in this study aimed at a temporarily activated focus on one of both thinking styles (*state*), the REI refers to the general tendency of the participants (*trait* or *characteristic*). For people with analytic thinking tendencies, for example, priming effects might still enhance analytic thinking, but not as strongly as for intuitive thinkers. Therefore, the REI is helpful for additional analyses of thinking style-related findings.

Reflecting upon t-tests, perceived competence (m = 6.32 vs. 5.92, $p = 0.002$) and perceived trustworthiness (m = 6.10 vs. 5.75, $p = 0.016$) towards the data scientist are significantly higher for participants whose REI rationality score is higher than average compared with their counterpart. While this is generally in line with the theory that more rational people tend to be less skeptical towards algorithmic decision aids, another finding is more surprising: Perceived competence (m = 6.00 vs. 5.72, $p = 0.019$) and perceived trustworthiness (m = 6.01 vs. 5.70, $p = 0.020$) towards the product manager are also significantly higher for participants whose REI rationality is higher than average compared with their counterpart.

This finding might result from the tendency of analytic thinkers to scrutinize the given information, i.e., advice sources, more thoroughly. Both advice sources– product manager and data scientist – were depicted relatively similar in the experiment in order to avoid a different perception of ability. Thus, the analytic thinkers might have noticed the similarity of the advice sources and assessed both more trustworthy and competent in general. Note that this might be a finding that blurs the original intention to compare the standalone effect of human expert vs. algorithm – an aspect that will be further discussed in Sect. 5.

Beyond that, it is crucial to analyze the REI experientiality subscale as well. This is because the two thinking styles might be considered as overlapping systems that can still be addressed independently from each other (Evans, 2003; Pacini & Epstein, 1999). Consistent with the previous considerations, there are no significant differences between perceived competence (m = 6.13 vs. 6.15, $p = 0.865$) and perceived trustworthiness (m = 5.97 vs. 5.91, $p = 0.694$) towards the data scientist for participants whose REI experientiality is higher than average and their counterpart. The same holds true for perceived competence (m = 5.88 vs. 5.87, $p = 0.952$) and perceived trustworthiness (m = 5.92 vs. 5.82, $p = 0.469$) towards the product manager.

Contrarily, comparing participants with higher-than-average experientiality with their counterpart, it is notable that there is a significant difference regarding the weight on algorithmic advice (54.72 vs. 59.96, $p=0.062$). This means that more intuitively inclined participants clearly prefer the human advice over the algorithmic one. Interestingly, this seems to be independent from perceived competence and trustworthiness towards the advice source. Thus, it seems likely that convincing highly intuitive decision-makers is an even bigger challenge as they might make decisions regardless of their own credibility perceptions towards advice sources. Reflecting the theoretical assumptions in Sect. 2, highly intuitive decision-makers are more prone to biases and misconceptions (Milkman et al., 2009), for example due to missing familiarity (Simon, 1955) and therefore tend to discount algorithms considerably. Against this background, the findings provide more detailed insights adding to existing literature on the credibility of an algorithmic advice source (Chen et al., 2022) as thinking style characteristics have previously not been analyzed in this regard.

4.2.4 Good news vs. bad news

A further analysis is concerned with the idea, and the existing corresponding findings from previous research, that a downward trend (bad news) results in a lower algorithm reliance (Chen et al., 2022; Fehrenbacher et al., 2023). This possibly affects the results in experiment 1 where the algorithmic forecast suggestion is lower than the human forecast. The original idea of good news vs. bad news is connected to a general trend that affects peoples' decisions as they are more likely to question the credibility of information for a worsening situation (Ditto et al., 1998; Ditto & Lopez, 1992). This is not directly covered by both experiments in this study since no baseline trend is given. However, experiment 2 involves different advice directions (i.e., which advice source forecasts a higher / lower value). In order to evaluate if there are similar effects for lower algorithm-based forecast predictions as 'bad news' and higher algorithm-based forecast predictions as 'good news', I run a paired samples t-test that reveals contradictory results (see Table 7). Interestingly, the decisions in regions with the algorithm forecast being lower than the product manager forecast ('bad news' in regions 1 and 3) result in a significantly higher algorithm reliance than the ones in regions 2 and 4 (m=58.11 vs. 52.63, $p<0.001$). These results, however, may not be overstated since the observations are not independent from each other and sequence effects cannot be ruled out. Furthermore, the 'good and bad news' in this experiment come along with various absolute spans between both advices that may affect the results.[32] Taking the actual participants' decisions (actual weight on algorithm advice) into account, the idea that absolute spans and numbers matter can be confirmed. Comparing their means with the self-reported weights, they show a

[32] Answering the question "My decision has been influenced by other aspects", one participant pointed out that their weighting decision was indeed affected by the size of the span. As one of the 'good news' forecast span is the highest in the experiment (region 4), this might be a possible explanation for the surprising finding. Additionally, the other 'good news' forecast (region 2) has the smallest span. In this case, participants might argue that slight differences between both advices do not have a large impact on the success of the firm. This might result in participants tending more to the middle, being more algorithm-averse than average (weight on algorithm advice: $m_2=54.06$ vs. $m_{total}=55.36$; actual weight on algorithm advice: $m_2=50.06$ vs. $m_{total}=55.29$).

Table 7 Analyses of 'good news' and 'bad news' (paired samples t-test) in experiment 2

	Mean					N	Std. Deviation	Std. Error Mean	
'Good news'	52.62					90	18.997	2.002	
'Bad news'	58.11					90	18.774	1.979	
	Paired Differences					t	df	Significance	
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				One-Sided p	Two-Sided p
				Lower	Upper				
'Good news' & 'Bad news'	-5.48	16.607	1,750	-8.961	-2,005	-3.132	89	0.001	0.002

Dependent variable: Weight on algorithmic advice – Participants had to adjust two sliders (weight on product manager forecast / weight on data scientist forecast) so that they add up to 100%.
'Good news' and 'bad news' in experiment 2 are defined in the way that the algorithmic prediction is either higher ('good news') or lower ('bad news') than the human advice. The paired samples t-test reveals a significant positive effect ($t = -3.132$, $p = 0.002$) of 'bad news' vs. 'good news' on the weight on algorithmic advice.

similar but slightly stronger tendency for regions 1–4 ($m = 61.47$ vs. 50.06 vs. 61.67 vs. 47.97).

There might be two explanations resulting in this surprising tendency as it is possible that participants perceive an algorithmic advice differently from the human advice depending on their risk preferences (e.g., Germann & Merkle, 2023). First, in case an algorithm assumes a considerably higher forecast than a human ('good news' as present in region 4), participants might be alarmed more severely and fall back to their intuitive judgment, preferring the more familiar human prediction. From the participants' view, the high deviation between human and algorithm forecast might even be perceived as 'bad news', indicating that one or both forecasts are flawed. Second, following a very similar idea, participants might have a general preference towards the lower prediction. One risk-averse participant actually reported that their decision was entirely based on the lowest advice in each region. In line with this, the participants' decisions for three of the four regions were below the respective average forecast ($m_1 = 57{,}005.56$ vs. 60,000, $m_2 = 30{,}005.56$ vs. 30,000, $m_3 = 40{,}366.67$ vs. 42,000, $m_4 = 77{,}472.22$ vs. 78,000), automatically discounting the algorithmic forecast more strongly in the 'good news' conditions.

Taken together, 'good news' and 'bad news' in different shapes (as in Chen et al. (2022) and as in this study) seem to be an important driver of algorithm reliance. In this experimental setting, the absolute span between different forecast values matters as well as the risk to overstate the final forecast, regardless of the advice source. It is crucial to consider that such differences are likely to affect the ultimate reliance on algorithms in real decision-making situations.

5 Discussion

The experiments in this paper yield several results that need to be discussed in order to broaden our understanding of algorithm reliance. In particular, the more detailed approach in experiment 2 suggests that the disclosure of different levels of information about the algorithm significantly influences reliance on an algorithmic advice. Specifically, participants show an increased algorithm reliance when detailed information about the algorithm's inner workings is provided, compared to when less information is disclosed. Additionally, analyses indicate that giving relevant information per se is more important than disclosing a very detailed level of information, as algorithm reliance increases with information load only up to a certain limit, implying a concave relation between information load and algorithm reliance.

Contrary to my initial predictions, an interaction effect between disclosing an algorithm's inner workings and thinking style cannot be identified. However, additional analyses reveal more detailed findings: First, Sect. 4.2.1 indicates that while priming analytic thinking has no significant effect on algorithm reliance, analytic thinking combined with disclosing an algorithm's inner workings increases algorithm reliance. Second, it may be argued that disclosing an algorithm's inner workings is connected to make an algorithm more credible (Chen et al., 2022; Önkal et al., 2009). However, results in Sect. 4.2.2 show that the credibility of a data scientist and an algorithm and even more the use of algorithms are complex functions that are driven by individual characteristics. While the propensity to engage in analytic thinking and other factors like trust in technology are positively related to advice source credibility, the perceived competence of the algorithm developer as well as the perceived trustworthiness of the algorithm itself have an increasing effect on algorithm reliance. There are slight indications of an underlying moderated mediation effect between disclosing an algorithm's inner workings (independent variable), perceived competence of the advice source (mediator), thinking style (moderator between independent variable and mediator) and algorithm reliance (dependent variable), but the overall effect is not significant. Thus, to understand and ultimately mitigate biases in terms of algorithm reliance, it is important to consider several details of the setting, including decision-makers traits and the perceived type of the advice source. Furthermore, Sect. 4.2.3 elucidates the importance of individual characteristics in terms of thinking styles. It might be presumed that the effect of advice source credibility is rather weak for highly intuitive thinkers, resulting in a poor algorithm reliance and an even more complex task to mitigate this.

In interpreting the findings, a number of limitations should be kept in mind. A potential confounding effect of experiment 1 is that participants might have perceived the advice sources to be optimistic vs. conservative, additionally to human vs. algorithmic. The second experiment accounts for these potential framing effects by varying advice direction, gathering more observations per subject. Exploring the impact of different numerical scenarios on algorithm reliance, experiment 2 reveals effects influenced by participants' risk preferences and perceptions of forecast accuracy. The findings emphasize the importance of considering both the level of information disclosure and contextual factors, such as the actual forecast results, in understanding algorithm reliance for decision-making contexts.

One possible explanation for the lack of strong effect in experiment 1 is the way the experimental design framed the decision context. It is important to consider the participants' perception of what an algorithm actually is (Logg et al., 2019). While the 'inner workings' manipulation accounts for this, both the data scientist and product manager were presented as human advisors. This may have led some participants to interpret the algorithmic advice as human-augmented rather than purely algorithmic. As a result, the decision frame might have shifted towards choosing between two human experts with different forecasting approaches rather than evaluating algorithmic versus human advice. This interpretation aligns with prior research showing that the perceived role of human involvement in algorithmic decision-making can significantly shape trust and reliance: A *human-in-the-loop* decision-making might bias trust in the algorithmic advice source positively (Burton et al., 2020; Dietvorst et al., 2018), contributing to an overly balanced weighting tendency in experiment 1. An analysis of weighting tendencies shows that the predominant weighting choice among participants was the 50:50 ratio: 107 out of 196 participants in the main sample and 73 out of 132 participants in the restricted sample selected this option. Only four participants chose to entirely rely on the algorithm, and no participant selected a weight below 20%. These results indicate a limited level of variance in the data, making it more challenging to derive significant conclusions. Even though experiment 2 did not change this general approach. the more diverse decision situation helped to mitigate this issue as the 50:50 weight was only used in approximately 37% of all decisions, resulting in statistically significant results.[33]

In terms of methodical choices, the potential priming effects in the experiment require further attention. While priming was employed to manipulate participants' perceptions, it might not have been strong enough to override the influence of individual characteristics on their decisions. As indicated in Sect. 4.2.3, individual characteristics have a significant effect on participants' perceived competence and trustworthiness of the advice source. Furthermore, due to the nature of the experimental design, it is not possible to entirely verify the effectiveness of the priming manipulation. The psychology literature acknowledges that priming effects can vary, sometimes aligning with expectations and sometimes deviating, irrespective of the underlying theory (Bargh & Chartrand, 2014).

The participants should also be considered. The sample in experiment 1 included a rather high share of students (~40%), a group that generally exhibits higher levels of algorithm reliance (Alexander et al., 2018; Önkal et al., 2009). This characteristic of the sample might have influenced the overall results. However, it is worth noting that age and gender do not show significant effects on the weight on algorithmic advice or on perceived competence and trustworthiness measures. Therefore, it can be assumed that result biases stemming from participant demographics might not exist.

[33] Interesting in this regard is a different behavior for regions where the midpoint of algorithm and human judgment is more obvious (e.g., 50,000 vs. 70,000) compared to regions where a certain effort is needed to actually determine the midpoint (e.g., 65,000 vs. 91,000). Regions 1 and 2 with a more obvious midpoint result much more often in the 50:50 ratio (~47%) than regions 3 and 4 with a less obvious midpoint (~27%). This is particularly surprising as I refer to the self-reported weights participants assigned to the advice sources, not to the absolute decisions they made.

Concerning the use of the Prolific platform for participant recruitment, it is important to consider that online experiments present challenges in identifying and controlling various factors, such as participants taking breaks or engaging in other activities that could influence their decision-making process. This issue appears particularly crucial in the context of priming experiments, in which maintaining strict control over participants' attention and focus is crucial for reliable results. Yet, by excluding potentially inattentive participants based on attention checks, performing robustness checks and excluding time-wise outliers in the restricted sample of experiment 1, I try to mitigate this issue. The second experiment shows not only much better attention check results but also a way smaller standard deviation of completion times (m = 10.26 vs. 15.74, sd = 4.26 vs. 7.63), although the experimental treatments differed more strongly in terms of text length. Additionally, in experiment 2, no participant reported interruptions or technical issues so that there are no indications to distrust the results.[34]

6 Conclusion

The objective of this study was to analyze mechanisms that could promote algorithm reliance in a forecasting context. My findings suggest that disclosing an algorithm's inner workings– aiming to open the often-perceived *black box*– increases algorithm reliance in a non-linear way, but it is crucial to consider an individual's thinking style as well as other characteristics, uncovering a highly complex decision situation to be aware of. Furthermore, decisions differ depending on the specific prognoses, as for example the absolute span between advice sources and the direction of the difference.

This study contributes to the literature, as it is one of the first to examine specific mechanisms to increase algorithm reliance in a management accounting context. By integrating insights from cognitive psychology, i.e., thinking style, elaborating on advice source credibility and explicitly analyzing the impact of disclosing specific kind of information regarding algorithms (inner workings), empirical evidence to enhance the use of algorithmic decision aids is presented. In particular, the findings suggest that providing relevant information is crucial, while additional information above a certain threshold does not increase algorithm reliance significantly. In addition, this study extends findings from Chen et al. (2022) who show that the perceived competence of the advice source is a driver of algorithm reliance. I find that thinking style characteristics affects people's reaction to credibility and that perceived trustworthiness of an algorithmic advice source is more important than assumed. Furthermore, this study emphasizes the role of AI in accelerating decision-making processes and shows that disclosing information about an algorithm's inner workings can enhance the use of algorithms. In this sense, providing a sufficient amount of relevant information seems more promising than overwhelming decision-makers with redundant information. From a management accounting perspective, these find-

[34] One participant mentioned a brief interruption due to network problems which was quickly resolved. As participants had no incentives to lie in this question and extensive completion times were rare (two participants took longer than 20 min), the data can be considered valid.

ings are relevant to gain more knowledge about the decision-making process involving algorithmic decision aids (Tank & Farrell, 2022). Results show that enabling a successful integration of data analytics requires considerations relating to the whole context, including given information and people's characteristics. Analyzing thinking style, this study provides insights to complement ideas of Glikson and Woolley (2020) who consider cognitive and emotional trust in data analytics to be crucial. Eventually, decision-makers' thinking style characteristics are shown to be relevant for credibility perceptions that drive the use of data analytics.

The practical implications emerging from this study refer to the design of decision-making environments that aim to mitigate biases. Particularly, by sharing specific information about the algorithm, decision-makers can be encouraged to overcome biases and embrace data-driven decision modes. What seems essential is to take decision-makers characteristics into consideration and trying to align the human and the algorithm way of thinking. As outlined in Sect. 5, emphasizing the human (expert) involvement in the modelling process of an algorithm might also enhance trust towards data analytics.

Future research may focus on a more in-depth understanding of the effectiveness of disclosing an algorithm's inner workings to comprehend how various types of information and presentation formats impact algorithm reliance. Researchers should focus on exploring the disclosure of different forms of information and alternative presentation methods, for example, as shown by Perkhofer et al. (2020) in a big data setting, to gain a deeper understanding of how the presentation of information affects algorithm reliance. In this regard, the limitations outlined in the previous discussion should be addressed, especially by identifying (experimental) settings where individual decisions are more diverse. In general, further elaboration of the complexity of decision situations involving algorithmic decision aids should be addressed more thoroughly by future studies, especially regarding various nuances of advice sources and advice output itself. As various technologies continue to play an increasingly important role in decision-making, investigating algorithm reliance is crucial to harness the benefits of data-driven decision-making across diverse domains.

Funding Open Access funding enabled and organized by Projekt DEAL.
No funds, grants, or other support was received.

Data availability Experimental proceedings and experimental data are available from the author upon request.

Declarations

Competing interests The author has no competing interests to declare that are relevant to the content of this article.

References

Adadi, A., & Berrada, M. (2018). Peeking inside the Black-Box: A survey on explainable artificial intelligence (XAI). *Ieee Access: Practical Innovations, Open Solutions*, *6*, 52138–52160. https://doi.org/10.1109/access.2018.2870052

Al-Htaybat, K., & von Alberti-Alhtaybat, L. (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

Albarracin, D., & Hart, W. (2011). Positive mood + action = negative mood + inaction: Effects of general action and inaction concepts on decisions and performance as a function of affect. *Emotion*, *11*(4), 951–957. https://doi.org/10.1037/a0024130

Alexander, V., Blinder, C., & Zak, P. J. (2018). Why trust an algorithm? Performance, cognition, and neurophysiology. *Computers in Human Behavior*, *89*, 279–288. https://doi.org/10.1016/j.chb.2018.07.026

Alvarado-Valencia, J. A., & Barrero, L. H. (2014). Reliance, trust and heuristics in judgmental forecasting. *Computers in Human Behavior*, *36*, 102–113. https://doi.org/10.1016/j.chb.2014.03.047

Andrejkow, J., Berger, L., & Guo, L. (2022). Conscious and nonconscious goal pursuit in multidimensional tasks. *Accounting Organizations and Society*, *103*, 101376. https://doi.org/10.1016/j.aos.2022.101376

Arnaboldi, M., Busco, C., & Cuganesan, S. (2017). Accounting, accountability, social media and big data: Revolution or hype? *Accounting Auditing & Accountability Journal*, *30*(4), 762–776. https://doi.org/10.1108/AAAJ-03-2017-2880

Arnott, D. (2006). Cognitive biases and decision support systems development: A design science approach. *Information Systems Journal*, *16*(1), 55–78. https://doi.org/10.1111/j.1365-2575.2006.00208.x

Bargh, J. A., & Chartrand, T. L. (2014). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2. publ, pp. 311–344). Cambridge University Press.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*(6), 925–945. https://doi.org/10.1037/0033-2909.126.6.925

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

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

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

Bogert, E., Lauharatanahirun, N., & Schecter, A. (2022). Human preferences toward algorithmic advice in a word association task. *Scientific Reports*, *12*(1), 14501. https://doi.org/10.1038/s41598-022-18638-2

Breuker, D., Matzner, M., Delfmann, P., & Becker, J. (2016). Comprehensible predictive models for business processes. *MIS Quarterly*, *40*(4), 1009–1034. https://doi.org/10.25300/MISQ/2016/40.4.10

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

Buckless, F. A., & Ravenscroft, S. P. (1990). Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review*, *65*(4), 933–945.

Burton, J. W., Stein, M., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, *33*(2), 220–239. https://doi.org/10.1002/bdm.2155

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

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

Chartrand, T. L., & Bargh, J. A. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology*, *71*(3), 464–478. https://doi.org/10.1037/0022-3514.71.3.464

Chen, C. X., Hudgins, R., & Wright, W. F. (2022). 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

Clor-Proell, S., & Nelson, M. W. (2007). Accounting standards, implementation guidance, and Example-Based reasoning. *Journal of Accounting Research*, *45*(4), 699–730. https://doi.org/10.1111/j.1475-679X.2007.00248.x

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

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

Dane, E., & Pratt, M. G. (2009). Conceptualizing and measuring intuition: A review of recent trends. In G. P. Hodgkinson & J. K. Ford (Eds.), *International Review of Industrial and Organizational Psychology* (1st ed., pp. 1–40). Wiley. https://doi.org/10.1002/9780470745267

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319. 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

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science (New York N Y)*, *243*(4899), 1668–1674. https://doi.org/10.1126/science.2648573

Derks, K., de Swart, J., Wagenmakers, E. J., & Wetzels, R. (2025). The bayesian approach to audit evidence: Quantifying statistical evidence using the Bayes factor. *AUDITING: A Journal of Practice & Theory*, *44*(1), 55–71. https://doi.org/10.2308/AJPT-2021-086

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

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *1*(2), 95–109. https://doi.org/10.1111/j.1745-6916.2006.00007.x

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*(4), 568. https://doi.org/10.1037/0022-3514.63.4.568

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, *75*(1), 53. https://doi.org/10.1037/0022-3514.75.1.53

Elkins, A. C., Dunbar, N. E., Adame, B., & Nunamaker, J. F. (2013). Are users threatened by credibility assessment systems?? *Journal of Management Information Systems*, *29*(4), 249–262. https://doi.org/10.2753/MIS0742-1222290409

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of Personality and Social Psychology*, *71*(2), 390–405. https://doi.org/10.1037/0022-3514.71.2.390

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, J. St. B. T. (2019). Reflections on reflection: The nature and function of type 2 processes in dual-process theories of reasoning. *Thinking & Reasoning*, *25*(4), 383–415. https://doi.org/10.1080/13546783.2019.1623071

Evans, J. St. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. Thinking & Reasoning, *11*(4), 382–389. https://doi.org/10.1080/13546780542000005

Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-Process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685
Farrell, A. M., Goh, J. O., & White, B. J. (2014). The effect of Performance-Based incentive contracts on system 1 and system 2 processing in affective decision contexts: fMRI and behavioral evidence. *The Accounting Review*, *89*(6), 1979–2010. https://doi.org/10.2308/accr-50852
Farrell, A. M., Grenier, J. H., & Leiby, J. (2017). Scoundrels or stars?? Theory and evidence on the quality of workers in online labor markets. *The Accounting Review*, *92*(1), 93–114. https://doi.org/10.2308/accr-51447
Fehrenbacher, D. D., Schulz, A. K. D., & Rotaru, K. (2018). The moderating role of decision mode in subjective performance evaluation. *Management Accounting Research*, *41*, 1–10. https://doi.org/10.1016/j.mar.2018.03.001
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
Fitzsimons, G. M., Chartrand, T. L., & Fitzsimons, G. J. (2008). Automatic effects of brand exposure on motivated behavior: How Apple makes you think different. *Journal of Consumer Research*, *35*(1), 21–35. https://doi.org/10.1086/527269
Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. *CHI '99: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems the CHI Is the Limit*, 80–87. https://doi.org/10.1145/302979.303001
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732
Gao, P., & Liang, P. J. (2013). Informational feedback, adverse selection, and optimal disclosure policy. *Journal of Accounting Research*, *51*(5), 1133–1158. https://doi.org/10.1111/1475-679X.12019
Germann, M., & Merkle, C. (2023). Algorithm aversion in delegated investing. *Journal of Business Economics*, *93*(9), 1691–1727. https://doi.org/10.1007/s11573-022-01121-9
Gervais, W. M., & Norenzayan, A. (2012). Analytic thinking promotes religious disbelief. *Science (New York N Y)*, *336*(6080), 493–496. https://doi.org/10.1126/science.1215647
Gillespie, T., Boczkowski, P. J., & Foot, K. A. (2014). In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society*. MIT Press.
Gillespie, N., Lockey, S., Curtis, C., Pool, J., & Akbari, A. (2023). *Trust in artificial intelligence: A global study*. The University of Queensland. https://doi.org/10.14264/00d3c94. KPMG Australia.
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
Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, *12*(1), 19–30. https://doi.org/10.1037/1040-3590.12.1.19
Guenther, T. W. (2013). Conceptualisations of 'controlling' in German-speaking countries: Analysis and comparison with Anglo-American management control frameworks. *Journal of Management Control*, *23*(4), 269–290. https://doi.org/10.1007/s00187-012-0166-7
Guggenmos, R. D., Piercey, M. D., & Agoglia, C. P. (2018). Custom contrast testing: Current trends and a new approach. *The Accounting Review*, *93*(5), 223–244. https://doi.org/10.2308/accr-52005
Hammersley, J. S., Bamber, E. M., & Carpenter, T. D. (2010). The influence of Documentation specificity and priming on auditors' fraud risk assessments and evidence evaluation decisions. *The Accounting Review*, *85*(2), 547–571. https://doi.org/10.2308/accr.2010.85.2.547
Hartmann, M., & Weißenberger, B. E. (2023). Information overload research in accounting: A systematic review of the literature. *Management Review Quarterly*. https://doi.org/10.1007/s11301-023-00343-7
Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis, second edition: A Regression-Based approach*. Guilford.
Higgins, E. T., Bargh, J. A., & Lombardi, W. J. (1985). Nature of priming effects on categorization. *Journal of Experimental Psychology: Learning Memory and Cognition*, *11*(1), 59–69. https://doi.org/10.1037/0278-7393.11.1.59
Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, *1*(3), 333–342. https://doi.org/10.1111/j.1754-9434.2008.00058.x
Huntsinger, J. R. (2011). Mood and trust in intuition interactively orchestrate correspondence between implicit and explicit attitudes. *Personality and Social Psychology Bulletin*, *37*(9), 1245–1258. https://doi.org/10.1177/0146167211410069

Hwang, M. I., & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science*, *25*(3), 213–218. https://doi.org/10.1177/016555159902500305

Iselin, E. R. (1996). Accounting information and the quality of financial managerial decisions. *Journal of Information Science*, *22*(2), 147–153. https://doi.org/10.1177/016555159602200208

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, *70*, 338–345. https://doi.org/10.1016/j.jbusres.2016.08.007

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jung, M., & Seiter, M. (2021). Towards a better Understanding on mitigating algorithm aversion in forecasting: An experimental study. *Journal of Management Control*. 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. *Proceedings of the 28th European Conference on Information Systems (ECIS)*.

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

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases: Heuristics and Biases* (1. edition). Cambridge University Press.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. W. Griffin, D. Kahneman, T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (1. edition, pp. 49–81). Cambridge Univ. Press. https://doi.org/10.1017/CBO9780511808098.004

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

Kirk, R. E. (2009). Experimental design. *The SAGE handbook of quantitative methods in psychology* (pp. 23–45). SAGE Publications Ltd. https://doi.org/10.4135/9780857020994.n2

Knudsen, D. R. (2020). Elusive boundaries, power relations, and knowledge production: A systematic review of the literature on digitalization in accounting. *International Journal of Accounting Information Systems*, *36*, 100441. https://doi.org/10.1016/j.accinf.2019.100441

Krügel, S., Ostermaier, A., & Uhl, M. (2022). Zombies in the loop? Humans trust untrustworthy AI-Advisors for ethical decisions. *Philosophy & Technology*, *35*(1), 1–37. https://doi.org/10.1007/s13347-022-00511-9

Krügel, S., Ostermaier, A., & Uhl, M. (2023). Algorithms as partners in crime: A lesson in ethics by design. *Computers in Human Behavior*, *138*, 107483. https://doi.org/10.1016/j.chb.2022.107483

Kulesza, T., Stumpf, S., Burnett, M., Yang, S., Kwan, I., & Wong, W. K. (2013). Too much, too little, or just right? Ways explanations impact end users' mental models. *2013 IEEE Symposium on Visual Languages and Human Centric Computing*, *3-10*. https://doi.org/10.1109/VLHCC.2013.6645235

Lambert, T. A., & Agoglia, C. P. (2011). Closing the loop: Review process factors affecting audit staff Follow-Through. *Journal of Accounting Research*, *49*(5), 1275–1306. https://doi.org/10.1111/j.1475-679X.2011.00423.x

Libby, R. (1981). *Accounting and human information processing: Theory and applications*. Prentice-Hall.

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31–57. https://doi.org/10.1145/3236386.3241340

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, *24*(3), 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

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

Madhavan, P., & Wiegmann, D. A. (2007). Effects of information source, pedigree, and reliability on operator interaction with decision support systems. *Human Factors*, *49*(5), 773–785. https://doi.org/10.1518/001872007X230154

Mahmud, H., Islam, A. K. M. 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

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

Mariani, M. M., & Wamba, S. F. (2020). Exploring how consumer goods companies Innovate in the digital age: The role of big data analytics companies. *Journal of Business Research*, *121*, 338–352. https://doi.org/10.1016/j.jbusres.2020.09.012

McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology. *ACM Transactions on Management Information Systems*, *2*(2), 1–25. https://doi.org/10.1145/1985347.1985353

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press. https://doi.org/10.1037/11281-000

Messner, M. (2009). The limits of accountability. *Accounting Organizations and Society*, *34*(8), 918–938. https://doi.org/10.1016/j.aos.2009.07.003

Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How can decision making be improved? *Perspectives on Psychological Science*, *4*(4), 379–383. https://doi.org/10.1111/j.1745-6924.2009.01142.x

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*, 1–38. https://doi.org/10.1016/j.artint.2018.07.007

Moreno, K., Kida, T., & Smith, J. F. (2002). The impact of affective reactions on risky decision making in accounting contexts. *Journal of Accounting Research*, *40*(5), 1331–1349. https://doi.org/10.1111/1475-679X.t01-1-00056

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

Önkal, D., Sinan Gönül, M., Goodwin, P., Thomson, M., & Öz, E. (2017). Evaluating expert advice in forecasting: Users' reactions to presumed vs. Experienced credibility. *International Journal of Forecasting*, *33*(1), 280–297. https://doi.org/10.1016/j.ijforecast.2015.12.009

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, *76*(6), 972–987. https://doi.org/10.1037/0022-3514.76.6.972

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Perkhofer, L., Walchshofer, C., & Hofer, P. (2020). Does design matter when visualizing big data? An empirical study to investigate the effect of visualization type and interaction use. *Journal of Management Control*, *31*(1–2), 55–95. https://doi.org/10.1007/s00187-020-00294-0

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Communication and persuasion* (pp. 1–24). Springer. https://doi.org/10.1007/978-1-4612-4964-1_1

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

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

Roetzel, P. G. (2019). Information overload in the information age: A review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research*, *12*(2), 479–522. https://doi.org/10.1007/s40685-018-0069-z

Schneider, G. P., Dai, J., Janvrin, D. J., Ajayi, K., & Raschke, R. L. (2015). Infer, predict, and assure: Accounting opportunities in data analytics. *Accounting Horizons*, *29*(3), 719–742. https://doi.org/10.2308/acch-51140

Schroder, H. M., Driver, M. J., & Streufert, S. (1967). *Human information processing: Individuals and groups functioning in complex social situations*. Holt, Rinehart and Winston.

Senoner, J., Netland, T., & Feuerriegel, S. (2022). Using explainable artificial intelligence to improve process quality: Evidence from semiconductor manufacturing. *Management Science*, *68*(8), 5704–5723. https://doi.org/10.1287/mnsc.2021.4190

Sieck, W. R., & Arkes, H. R. (2005). The recalcitrance of overconfidence and its contribution to decision aid neglect. *Journal of Behavioral Decision Making*, *18*(1), 29–53. https://doi.org/10.1002/bdm.486
Silver, N. (2012). *The Signal and the Noise: The Art and Science of Prediction*. Penguin UK.
Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99. https://doi.org/10.2307/1884852
Simon, H. A. (1986). Rationality in psychology and economics. *The Journal of Business*, *59*(4), 209–224. https://doi.org/10.1086/296363
Sprinkle, G. B. (2003). Perspectives on experimental research in managerial accounting. *Accounting Organizations and Society*, *28*(2–3), 287–318. https://doi.org/10.1016/S0361-3682(01)00058-7
Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*(4), 672–695. https://doi.org/10.1037/0022-3514.94.4.672
Stanovich, K. E., & West, R. F. (2013). Evolutionary versus instrumental goals: How evolutionary psychology misconceives human rationality. *Evolution and the psychology of thinking: The debate* (pp. 171–230). Psychology.
Stocks, M. H., & Harrell, A. (1995). The impact of an increase in accounting information level on the judgment quality of individuals and groups. *Accounting Organizations and Society*, *20*(7–8), 685–700. https://doi.org/10.1016/0361-3682(95)00012-X
Tank, A. K., & Farrell, A. M. (2022). Is neuroaccounting taking a place on the stage?? A review of the influence of neuroscience on accounting research. *European Accounting Review*, *31*(1), 173–207. https://doi.org/10.1080/09638180.2020.1866634
Teoh, S. H. (2018). The promise and challenges of new datasets for accounting research. *Accounting Organizations and Society*, *68–69*, 109–117. https://doi.org/10.1016/j.aos.2018.03.008
Thomas, T. F. (2016). Motivating revisions of management accounting systems: An examination of organizational goals and accounting feedback. *Accounting Organizations and Society*, *53*, 1–16. https://doi.org/10.1016/j.aos.2016.07.001
Trotman, K. T., Tan, H. C., & Ang, N. (2011). Fifty-year overview of judgment and decision-making research in accounting. *Accounting & Finance*, *51*(1), 278–360. https://doi.org/10.1111/j.1467-629X.2010.00398.x
Turel, O., & Kalhan, S. (2023). Prejudiced against the machine?? Implicit associations and the transience of algorithm aversion: MIS quarterly. *MIS Quarterly*, *47*(4), 1369–1394. https://doi.org/10.25300/MISQ/2022/17961
Vankov, I. I. (2023). The hazards of dealing with response time outliers. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1220281
Vessey, I. (1991). Cognitive fit: A Theory-Based analysis of the graphs versus tables literature. *Decision Sciences*, *22*(2), 219–240. https://doi.org/10.1111/j.1540-5915.1991.tb00344.x
Vitale, G., Cupertino, S., & Riccaboni, A. (2020). Big data and management control systems change: The case of an agricultural SME. *Journal of Management Control*, *31*(1–2), 123–152. https://doi.org/10.1007/s00187-020-00298-w
Warren, J. D., Moffitt, K. C., & Byrnes, P. (2015). How big data will change accounting. *Accounting Horizons*, *29*(2), 397–407. https://doi.org/10.2308/acch-51069
Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, *142*(5), 472–497. https://doi.org/10.1037/bul0000030
White, R. J. (2005). *A priming/temperament model of system 1 and system 2 decision making processes*. (Unpublished doctoral dissertation). Ohio State University, Columbus.
Wibbeke, L. M., & Lachmann, M. (2020). Psychology in management accounting and control research: An overview of the recent literature. *Journal of Management Control*, *31*(3), 275–328. https://doi.org/10.1007/s00187-020-00302-3
Wolfe, C. J., Christensen, B. E., & Vandervelde, S. D. (2020). Intuition versus analytical thinking and impairment testing. *Contemporary Accounting Research*, *37*(3), 1598–1621. https://doi.org/10.1111/1911-3846.12568
Woolley, K., & Risen, J. L. (2018). Closing your eyes to follow your heart: Avoiding information to protect a strong intuitive preference. *Journal of Personality and Social Psychology*, *114*(2), 230–245. https://doi.org/10.1037/pspa0000100