

Not So Averse After All: Behavioral Mechanisms Underlying the Use of Algorithms in Managerial Forecasting

Peter Kotzian, Barbara E. Weißenberger, Sonja G. Prinz

Article - Version of Record

Suggested Citation:

Kotzian, P., Weißenberger, B. E., & Prinz, S. G. (2025). Not So Averse After All: Behavioral Mechanisms Underlying the Use of Algorithms in Managerial Forecasting. *International Journal of Human-Computer Interaction*, 42(7), 4907–4931. <https://doi.org/10.1080/10447318.2025.2543988>

Wissen, wo das Wissen ist.

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20260420-094939-7>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>



Not So Averse After All: Behavioral Mechanisms Underlying the Use of Algorithms in Managerial Forecasting

Peter Kotzian, Barbara E. Weißenberger & Sonja G. Prinz

To cite this article: Peter Kotzian, Barbara E. Weißenberger & Sonja G. Prinz (2026) Not So Averse After All: Behavioral Mechanisms Underlying the Use of Algorithms in Managerial Forecasting, *International Journal of Human-Computer Interaction*, 42:7, 4907-4931, DOI: [10.1080/10447318.2025.2543988](https://doi.org/10.1080/10447318.2025.2543988)

To link to this article: <https://doi.org/10.1080/10447318.2025.2543988>



© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 12 Aug 2025.



[Submit your article to this journal](#)



Article views: 1569



[View related articles](#)



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

Not So Averse After All: Behavioral Mechanisms Underlying the Use of Algorithms in Managerial Forecasting

Peter Kotzian , Barbara E. Weißenberger  and Sonja G. Prinz

Chair of Management Control and Accounting, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

ABSTRACT

Although algorithms potentially outperform human judgment in managerial decision-making and forecasting, practitioners frequently refrain from using them. Extant literature provides mixed evidence on algorithm aversion and suggests very different underlying behavioral mechanisms. We contribute to existing research in two ways. First, we replace the traditional experimental instrument of measuring algorithm aversion, which asks participants to abandon their personal forecast in favor of a second, usually superior algorithmic forecast, with a new approach, in which participants receive human-made and algorithmic information simultaneously to form their judgments. Second, we derive competing empirical implications from potential mechanisms, allowing for an evaluation of their relevance. Using two factorial survey experiments with a forecasting task, we find very limited algorithm aversion. Regarding mechanisms underlying algorithm aversion discussed in existing research, we find that selective attention is most relevant: Users are more likely to notice superior performance of a human than of an algorithm.

HIGHLIGHTS

- Algorithm aversion as measured in extant literature is largely due to psychological ownership effects concerning outputs produced by users.
- Algorithm's superior performance is perceived less than superior performance of a human.
- Users act on their perceptions of performance in the case of humans, but not in the case of algorithms.

KEYWORDS

Forecasting; algorithm aversion; algorithmic decision support; explainable AI



JEL CLASSIFICATION

M15; M41

1. Introduction

Driven by rapidly advancing development of artificial intelligence (AI)-based information technology, firms invest substantial resources in integrating algorithmic decision support into information systems (Chaturvedi et al., 2025; Lui et al., 2022; Rossmann & Wald, 2024; Schmidt et al., 2020), notably forecasting of business-relevant factors (Panda & Mohanty, 2023; Varghese Joseph et al., 2022; Shrivastava et al., 2023) and subsequent managerial decision-making and forecasting (KPMG, 2024). However, factual usage of algorithmic support in management lags behind its potential (Brink et al., 2024; Kaufmann et al., 2023; Monod et al., 2023; Patton et al., 2025). This phenomenon is typically attributed to “algorithm aversion” (see Dietvorst et al., 2015): Decision-makers are – allegedly – averse to decision support provided by algorithms even if it is superior to human expertise (Brink et al., 2024; Commerford et al., 2022; Fildes et al., 2019; Johnson et al., 2021; Longoni et al., 2019; Raisch & Krakowski, 2021).

However, in more recent literature on usage of algorithmic support, algorithm aversion is no robust finding. It is situational and subject to change, especially with the rapid proliferation of algorithms in many areas of personal and professional life (Davenport, 2018; Harrell, 2016; Köchling et al., 2021). There is even evidence of over-reliance on algorithms (Keding & Meissner, 2021), as well as instances

CONTACT Peter Kotzian  kotzian@hhu.de  Chair of Management Control and Accounting, Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

of over-trust in inferior algorithmic output or a clear preference for algorithms, a phenomenon labeled “algorithm appreciation” (Aschauer et al., 2024; Choung et al., 2024; Logg et al., 2019; Schecter et al., 2023). Empirical results indicate that both, algorithm aversion and appreciation, depend on specific conditions, e.g., the character of the task at hand (Castelo et al., 2019) or whether users want to delegate unpleasant decisions (e.g., Maasland & Weißmüller, 2022). Algorithm aversion is found in some business settings, e.g., in forecasting (Himmelstein & Budescu, 2023), but not in others, e.g., recruiting (Castelo et al., 2019; Gazit et al., 2023; Kaufmann et al., 2023). Finally, algorithm aversion depends on the size of the stakes (Filiz et al., 2023), on whether poor advice from the algorithm was received in the past (Prahla & van Swol, 2017; Saragih & Morrison, 2022), and on whether supplementary explanations for algorithmic output are provided (e.g., Ning et al., 2024; Zhang et al., 2022).

As a result and despite substantial research, the phenomenon of algorithm aversion is not well understood. Much existing research focuses on the role of trust (Choung et al., 2023; Nakashima et al., 2024) as a factor for algorithm aversion, but on the other hand, the relevance of trust for algorithm usage was questioned by recent research (e.g., Schaap et al., 2024). Taken together, this indicates that the behavioral mechanisms underlying the (non-)usage of algorithmic advice are unclear. Indeed, recent reviews on algorithm usage emphasize that a theoretical understanding of the mechanisms underlying the phenomenon is largely missing (e.g., Baines et al., 2024; Jussupow et al., 2020; Oomen et al., 2024). This also makes it difficult for firms to develop effective ways to encourage usage of superior algorithmic advice by managers, where appropriate.

Our study specifically investigates into competing behavioral mechanisms, deriving empirical implications from each. Using a practitioner sample and a forecasting task, we conduct two experimental studies in which we study algorithm usage in a realistic management setting introduced by Fildes et al. (2019). Participants receive two forecasts from different sources: One provided by a human, a second one provided either by another human or by an algorithm. Participants are then asked to use this input as a basis for their own forecast which is passed on to a superior manager. This setting represents a typical situation in business planning, where pre-processed information from different sources has to be integrated (Mayer et al., 2023).

Such an experimental design differs from the standard approach to capturing algorithm aversion, used initially by Dietvorst et al. (2015) and going back to Meehl (1954) who was the first to observe the phenomenon that people – in his case clinician psychologists – rather rely on their own expertise and intuition for predicting human behavior compared to even superior statistical (“actuarial”) decision support. In this approach, participants first produce their own forecast based on “raw” (i.e., unprocessed) data and only then they receive decision support from a second algorithmic source. As a result, they at least partially have to abandon their personal forecast in favor of an algorithmic one.

We argue, that the standard approach disadvantages algorithms as a source of a forecast. First, it induces psychological ownership, as participants have invested effort in producing their forecast (Haesebrouck, 2021; Lim & O’Connor, 1995; Pierce et al., 2001). Second, by producing a forecast from raw data, actors gain, or at least feel to gain, more knowledge about the subject of the forecast. Expertise and knowledge are known to reduce the acceptability of (algorithmic) advice (Bayer et al., 2022; Kaufmann et al., 2023). Finally, a forecast, for which participants know exactly how it was made, competes with a forecast about which participants know nothing. Given that the intransparency of algorithms is often criticized and seen as a barrier to their use (Adadi & Berrada, 2018; Ning et al., 2024; Schmidt et al., 2020; Shin, 2021), this measurement approach is prone to seemingly increase algorithm aversion, while actually capturing aversion to intransparency. Moreover, even though there is much research on algorithmic intransparency, there is little research comparing how users cope with intransparency of human advice compared to algorithmic advice. Our research design avoids these problems and allows us to distinguish between aversion due to an ownership effect, aversion concerning the algorithm per se and aversion to intransparency.

In terms of mechanisms, we first inquire into whether trust, often considered a crucial antecedent of usage (Ning et al., 2024; Schmidt et al., 2020; Siau & Wang, 2018; Venkatesh et al., 2016), mediates the relationship between origin and usage of a forecast. We measure unspecific, general trust in the forecast, but also specific trust in various aspects of the forecast, including the processes of producing the

forecast. We find trust to be neither systematically affected by origin nor the past performance of the source of a given forecast, and in addition, trust does not affect usage.

As a second mechanism, we investigate whether lack of transparency, which arises from a so-called black box algorithm, affects algorithm usage, as extant literature argues (Ning et al., 2024; Schmidt et al., 2020; Shin, 2021). As before, we find no support for the view of intransparency being a specific problem of algorithms. Providing an explanation of a forecast has a positive effect on usage, but this effect is weak and moreover very similar for humans as well as for algorithms as forecast sources.

Third, we examine a mechanism linking the use of algorithms to participants' presumptions about their superiors' attitudes. Specifically, we investigate whether a superior's aversion to algorithms hinders usage. Existing research indicates that users care about the impression others have of them if they use an algorithm (e.g., Efendić et al., 2024; Shaffer et al., 2013). If this were the case, participants who believe that their superiors disapprove of the use of algorithms, would not use algorithmic decision support regardless of their personal preference. While there is evidence that participants – just as employees in firms – presume their superiors to be more skeptical of algorithmic decision support, this presumption proves to be not relevant for participants' usage of algorithms in our experiment.

Finally, as a fourth mechanism, we inquire into the role of algorithm performance for usage. Here we find that algorithm aversion originates by participants tending to ignore the objective performance of a source because of its origin. While superior performance of a human forecaster is recognized as such and advice from this source appropriately used as decision support, this is not the case for algorithms. Participants do not perceive the superior algorithm to be superior in the first place. Instead, to be perceived as superior, an algorithm has to outperform an alternative source to a much stronger degree than a human alternative.

Our article's first contribution is to unravel behavioral mechanisms underlying algorithm aversion by systematically testing several competing mechanisms, something which is lacking in extant literature. Our second contribution consists in introducing a new mechanism which links algorithm usage to the organizational environment, via the presumed acceptance of algorithms by superior managers. Ultimately, we find aversion to using algorithmic decision support is driven by selective attention to performance but also by participants acting upon their subjective beliefs about performance rather than the objective performance of the source providing advice: In both aspects, participants treat algorithms and humans differently. This mechanism was not yet covered in existing research.

Our third contribution is methodological: We measure algorithm usage in a way which eliminates effects of psychological ownership arising from having to discard a personally produced output in favor of an output coming from a different source. Our results, using this approach, suggest that algorithm aversion at least today is much weaker than hitherto presumed, indicating that part of the reported algorithm aversion is a measurement effect, because psychological ownership effects might be misinterpreted as algorithm aversion.

Our fourth and final contribution relates to business practice. Our findings suggest that perceptions of algorithmic performance are crucial. Simply providing objectively high performing algorithms in management practice might therefore be insufficient to ensure usage, as users tend to ignore this fact: Firms cannot assume that users automatically become aware of an algorithm's performance.

In terms of an outlook for future research, the limited role of transparency and trust together with selective attention puts the focus on other potential barriers for the use of objectively superior algorithmic decision support. For instance, issues of responsibility and blame-shifting towards human sources in risky decision situations may constitute greater barriers to usage of algorithmic advice. This, as well as the role of different types of explanations reducing the black box nature of algorithms should be inquired further.

Our article is organized as follows. In [Section 2](#), we provide the theoretical background of our research and derive hypotheses. In [Section 3](#), we present the experimental setting and the research design used. [Sections 4](#) and [5](#) describe the results of two full factorial survey experiments we conducted. In [Section 6](#), we discuss our results and draw conclusions.

2. Theoretical background

2.1. Measuring aversion to algorithms

Since the seminal studies of Dietvorst et al. (2015) and Dietvorst et al. (2018), algorithm aversion is understood as the non-usage of algorithms or algorithmic advice despite the fact that doing so is beneficial, as the algorithm outperforms the user or other sources of advice. A highly illustrative example of aversion is manager's reluctance of using of algorithm-based decision support in forecasting tasks (Mahmud et al., 2022). First, forecasts are an immensely relevant element in management accounting and managerial decision-making (e.g., Casas-Arce et al., 2022; Chui et al., 2018; Eichholz et al., 2023; Monteiro et al., 2022; Rohrbeck & Kum, 2018). Second, forecasting is a promising case of augmenting human decision-making with algorithmic input (Blattberg & Hoch, 1990; Myers & Ramsey, 2023). The progress in forecasting variables with high relevance for business is substantial (Panda & Mohanty, 2023; Shrivastava et al., 2023; Varghese Joseph et al., 2022). Whereas traditional human forecasting uses a limited set of methods and data sources (Casas-Arce et al., 2022; De Baets, 2021; Verstraete et al., 2020), AI-based algorithms allow integrating a much broader set of heterogeneous and structured as well as unstructured data sources. As a result, they can make more accurate and more timely forecasts that are unbiased by preferences, motivated reasoning or strategic gaming from human forecasters (Appelbaum et al., 2017; Blackburn et al., 2015; Zhang et al., 2022), resulting in so-called "predictive analytics" (Fehrenbacher et al., 2023; Göbel, 2023).

Given these advantages, introducing algorithms in management and accounting is a "trendy" and widely discussed topic (Bakarich & O'Brien, 2021; Enholm et al., 2022; Möller et al., 2020; Mosier & Skitka, 2018) and a persistent theme in research on management (Myers & Ramsey, 2023), management information systems (Blackburn et al., 2015) and related domains (Afsay et al., 2023; Krieger et al., 2021). Still, studies on factual usage of algorithms indicate that, notably in managerial settings, algorithmic support in case-by-case decisions, e.g., providing a forecast for a single case, is avoided (see e.g., Brink et al., 2024; Kaufmann et al., 2023; Monod et al., 2023).

Based on this traditional conceptualization of algorithm aversion, its measurement follows a more or less standardized procedure going back to Dietvorst et al. (2015). Applied to a forecasting task, it works as follows: First, users receive raw data, form a mental model of the drivers and causal relationships involved and then make their own forecast based on this model. Next, users are presented with a competing algorithmic forecast and have to decide, whether to stick with their forecast or to replace it with the algorithmic forecast. While details of the procedure vary regarding the aspects motivating a particular study (see, e.g., Daschner & Obermaier, 2022; Jung & Seiter, 2021; Önkal et al., 2009; Prahl & van Swol, 2017), the underlying either-or procedure applies.

In the light of more general research on how decision-makers deal with advice and different sources of input to formulate a forecast (see the review by Leitner & Leopold-Wildburger, 2011), such a measurement is problematic as it may lead to over-estimation the phenomenon of algorithm aversion. There are several reasons for this.

First, the issue of psychological ownership (Haesebrouck, 2021; Liu et al., 2012; Pierce et al., 2001) could play a role which applies to forecasts that users have produced with more or less substantial effort. As a consequence, users have a stake in this forecast because abandoning a personal output comes at the price of cognitive dissonance (Arkes & Blumer, 1985; Schultze et al., 2012). Thus, pitching a forecast produced by a user against a forecast coming from an algorithm may find aversion against giving up one's own forecast: Not because of algorithm aversion but because of a feeling of psychological ownership regarding one's forecast. Furthermore, having decided – here about a forecast – also affects how incoming information is used (Bronfman et al., 2015), how information from various sources is combined (Bonaccio & Dalal, 2006; Leitner & Leopold-Wildburger, 2011) and how a decision aid is evaluated (Bronner & de Hoog, 1984; Chaxel et al., 2013). Employing the standard design with a forecast produced by the user will per se lower the chances that an alternative forecast, regardless of its source, is used. Even though this is not algorithm aversion, it would be interpreted as such in the standardized experimental setting described above.

Second, a major objection to algorithmic advice is its lack of transparency (Park & Yoon, 2024; Schmidt et al., 2020; Shin, 2021). Regarding this issue, letting users produce their own forecast also

stacks the odds against using an alternative: Users are aware of the reasons and justifications for the forecast they produce, but are lacking this information when receiving an algorithmic forecast. This can also induce them to rely more on their own judgment (Yaniv, 2004; Yaniv & Kleinberger, 2000). Again, this is no expression of algorithm aversion but applies equally to all forecasts without supporting elaborations, independent of their origin. However, most existing research lacks a reference, i.e., opaque human advice, to capture this effect. It is therefore unclear to what degree the standard measurement procedure captures algorithm aversion rather than potentially confounding factors.

Third, it is questionable to what extent the standard measurement approach corresponds to the practice in firms, for example in the finance and accounting function (Fildes et al., 2019). While management accountants at lower levels personally prepare forecasts and receive algorithmic advice, making them subject to the problems mentioned, this is not the case for the higher-level management. Here, the choice is not between a personally produced forecast and an alternative, but between alternatives with different performance records and different origins, like humans or algorithms. Given these issues, it is necessary to test for the robustness of algorithm aversion as a general phenomenon.

2.2. Hypotheses development

Existing research discusses several competing mechanisms underlying the phenomenon of algorithm aversion. Identifying the relevance of these mechanisms is particularly important for management practice, as any interventions to address algorithm aversion will need to address these mechanisms. We therefore focus on several prominent hypothesized mechanisms that have been discussed in research so far (Baines et al., 2024; Glikson & Williams Woolley, 2020; Jussupow et al., 2020; Mahmud et al., 2022). To structure this section, we suggest three stages which have to be passed for the usage of algorithmic decision support. First, the algorithm's performance must be perceived correctly, as using inferior input would be inappropriate. Second, there must be trust in any unobservable properties of the algorithm, the notion of trust will be differentiated in the following elaborations. Third, there may be barriers to usage arising from the organizational context in which the user is embedded that prevent the user from using an algorithm despite a personal willingness to do so. Potential mechanisms can be arranged along this timeline.

2.2.1. Performance and usage

The very definition of algorithm aversion depends on the superior performance of algorithms. Beginning with Dietvorst et al. (2015), or even with Meehl (1954), the assumption is that if algorithms outperform human decision-making, users should rely on algorithmic decisions rather than on their own judgement or intuition. It follows that a typical feature of experimental studies on algorithm usage is that participants receive advice from an objectively superior algorithm. The superior performance can be presented as information or demonstrated by high forecasting accuracy during the experiment. The assumption that an algorithm's superior performance automatically leads to increased usage is not corroborated, as in many experiments participants refrain from using superior algorithms for their decision-making (Brink et al., 2024; Commerford et al., 2022; Fildes et al., 2019; Johnson et al., 2021; Longoni et al., 2019; Raisch & Krakowski, 2021).

One interpretation of the practical observation of non-usage of algorithms may be that even though users see the higher performance, they choose not to use it due to aversion to algorithms. However, whether participants in experiments actually recognize the algorithm's superiority is not always ascertained. We therefore suggest taking a differentiated view on performance as a factor in algorithm usage. It can be convincingly argued that performance matters for usage (Daschner & Obermaier, 2022; Lim & O'Connor, 1995; Schaap et al., 2024) albeit algorithm aversion obstructs usage. The ambiguous experimental results on usage of superior algorithms might be caused by selective attention: Participants – but also users in general – may ignore the algorithm's higher performance and thus its appropriateness as a source of advice. Selective attention biases have been found in many settings (e.g., Hales, 2007; Woolley & Risen, 2018), and may be at work here, too. We test for this issue and suggest that algorithm aversion expresses itself in the fact that users do not pay attention to the algorithm's objectively superior performance for the very reason that it is an algorithm. As a result, if this bias was

at work, the origin of a forecast source (human vs. algorithmic) would affect how its performance is perceived and thereby – indirectly – used.

Hypothesis 1a: *Higher perceived past performance of a forecasting source positively affects the use of a forecast provided by this source.*

Hypothesis 1b: *Superior forecasting performance by an algorithm is perceived to a lesser degree than superior forecasting performance by a human expert.*

2.2.2. Trust and usage

Trust is considered relevant for usage of sources on which forecasts are based (e.g., Gönül et al., 2012; Önköl et al., 2019). Consequently, lack of trust can be considered as a barrier to usage of algorithmic decision support (Bareis, 2024; Nakashima et al., 2024). Indeed, extant literature found evidence of trust as an antecedent to the use of advice from (non-)human sources (Daschner & Obermaier, 2022; Davenport, 2018; Glikson & Williams Woolley, 2020; Li et al., 2008). However, the direct role of trust has been challenged by more recent research (Bayer et al., 2022; Schaap et al., 2024). Applied to the issue of forecasting, the effect of trust should express itself in that trust matters for usage of a given source of decision support. Existing research has shown that algorithms typically receive less trust than a human (see for many Glikson & Williams Woolley, 2020). We therefore state the following general hypotheses.

Hypothesis 2a: *Trust in a forecasting source positively affects its usage.*

Hypothesis 2b: *Participants trust algorithmic forecasts to a lesser degree than forecasts based on human expertise.*

2.2.3. Barriers in the decision-maker's environment

Even if a decision-maker would be inclined to rely on an algorithm, he or she may refrain from doing so because of barriers in the organizational environment. There are several studies, notably from the medical field, which indicate that usage of algorithms – regardless of the appropriateness of doing so – is viewed critical by many actors (Arkes et al., 2007; Efendić et al., 2024; Shaffer et al., 2013). As a result, potential users may assume that other agents would disapprove of using algorithms as a source of advice, and thus, regardless of their own opinion, abstain from using them as well. As business forecasting is embedded in a larger organizational context also involving relationships of accountability, barriers to the usage of algorithms can therefore be derived in an analogous fashion, e.g., from upper echelon theory (Hambrick & Mason, 1984). Hambrick and Mason posit that organizational outcomes to a large degree have to be interpreted in the light of values and cognitive bases of a firm's top management. Important is not the actual view of the top-level management but the perceptions held among the lower-level managers. As a result, if users presume that their superiors are algorithm averse (Alexander et al., 2018) or if users believe that they are judged negatively if they rely on algorithms (Arkes et al., 2007; Shaffer et al., 2013), this precludes usage of algorithmic advice regardless of the user's personal attitudes on usage. We therefore formulate the following two hypotheses:

Hypothesis 3a: *Presumed trust of a superior in a forecasting source is positively related to its usage.*

Hypothesis 3b: *The origin of a forecasting source affects presumed trust, with an algorithmic source receiving less presumed trust than a human expert.*

A second type of barrier in this context is the lack of explanation on how a specific forecast came about. Advanced algorithmic forecasting technologies, like, e.g., deep learning with neural networks, typically cannot produce an explanation on how they produced an output. Instead, they are black boxes (Giboney et al., 2015; London, 2019; Merendino et al., 2018; Vinson et al., 2018; Zhang et al., 2022). In contrast, human experts providing advice can usually provide such an explanation for superiors or other stakeholders, either ex ante or ex post if needed. However, whether reactions to intransparent human vs. algorithmic advice differ, has, to the best of our knowledge, not yet ascertained, albeit there

is a broad literature on using so-called interpretable (or explainable) artificial intelligence (Grimmelikhuijsen, 2023; Lösse & Weißenberger, 2023). Based on the arguments made and existing evidence on transparency of algorithms, we hypothesize that explanations, in particular elaborations on how a forecast came about, are relevant for usage (see e.g., Shin, 2021 for evidence, notably on human input; Tzioti et al., 2014). Explainability should increase usage per se, and it should matter more, when black box algorithms are the source of the forecast.

Hypothesis 4a: *Explainability on how a forecast came about increases its usage.*

Hypothesis 4b: *Explainability increases usage of forecasts by algorithmic sources more than they increase forecasts provided by human experts.*

3. Method

To test our hypotheses, we conducted two factorial online survey experiments (Oll et al., 2018). The first experiment aims at replicating the standard observation of algorithm aversion, using the measurement procedure outlined above, and also at investigating potential mechanisms underlying algorithm aversion. The second experiment specifically investigates the role of explainability for usage of algorithmic output.

Both experiments share the basic research design. Participants were asked to put themselves into the role of a division's head of finance (i.e., chief management accountant). They were tasked with providing as accurately as possible a sales growth forecast for one of the firm's divisions for the upcoming quarter to the top management, with no other information than two forecasts coming from two sources. Participants received a fixed participation fee and were incentivized by a variable payoff, linked to the forecast/actual-deviation, which was determined after the experiment.

In both experiments, one forecast the participants received was always described as coming from a human staff member ("management accountant A") regularly in charge of this task. No further information on how this forecast came about was given besides from the fact that "traditional forecasting instruments" were used. Simultaneously, participants were provided with another forecast from a second source.

The first manipulation in the 2×2 between-subjects design used in both experiments concerns the origin of this second forecast. It is either produced by an algorithm (referred to as an artificial intelligence, ALGO) or by another human expert (referred to as another management accountant "B", HUMAN). The second manipulation concerns the sources' past performance, which was presented to the participants using comparisons of past forecasts and with the realized values (actuals). In the EQUAL condition, both forecasts provided were equally accurate in the past. The sum of forecast/actual-deviations is identical, no forecast features a bias in one direction and both were right on target for the same number of quarters. In the BETTER condition, the second forecaster is substantially more accurate (the sum of absolute forecast/actual-deviations is smaller), see Appendix A. Given the strong evidence on the relevance of presentation formats, in particular tables vs. graphs (Paivio, 1979; Perkhofer et al., 2020), the accuracy-related information was presented in both formats, to maximize its effect on the participants. In both experiments, we used a between-subjects design.

To provide their own forecast of the division's expected sales growth rate, participants received simultaneously two forecasts for the upcoming quarter. Both input forecasts diverged, but both were equally realistic given past actuals. Participants were instructed to derive their own forecast for the upcoming quarter to report to the top management, using both input forecasts as they saw fit. The forecast reported by the participants was the dependent variable of the experiments.

A post-experimental questionnaire covered basic demographics, but also questions related to the mechanisms we want to test. Specifically, we measured the participant's perception of performance, procedural trust in how a forecast came about, and the participant's perceptions about the trust of the management in algorithms.

In experiment 2, we replicated the design of the first experiment with two modifications: First, we made the difference between the EQUAL and BETTER condition more pronounced by increasing

relative forecast accuracy of the second source compared to experiment 1. Second, we added an explanation for the second forecast which was identically in both cases, i.e., for the second human forecaster as well as for the algorithm. The explanation illustrates the forecasting process, but beyond this does not deliver any other information relating to the content of the forecast (see [Appendix C](#)).

4. Experiment 1

4.1. Experimental design

In the first experiment, the origin of the second forecast source (HUMAN vs. ALGO) as well as its past performance (EQUAL vs. BETTER) were manipulated (2×2 between subjects design, for performance presentation see [Appendix A](#)). In the EQUAL condition, both forecasts had in total an absolute deviation of 13 percentage points for the period under consideration. In the BETTER condition, the first (human) forecasting source had in total an absolute deviation of 16.7 percentage points, the second forecasting source (either human or algorithmic origin) had in total an absolute deviation of 7.3 percentage points. No additional information on how the forecasts came about was provided. For the quarter in question, the first forecast from the human was for 4.2% sales growth, the second forecast (from a second human or an algorithm) for 5.9%. The experiment was pretested with senior managers from corporate finance and accounting departments to assure the plausibility of this setting.

4.2. Sample characteristics

Participants were practitioners, management accountants and other people, who can be assumed to be familiar with the forecasting scenario used in the experiment. They were recruited via a provider of online surveys. Overall, 138 persons participated, 58% of which are male and 30% hold a university degree of which roughly a third (i.e., 9% of all participants) hold a degree in economics or business administration. About 35% of the participants work in sales and distribution functions and 19% in management accounting functions. The average age is 51 years, ranging from 35 to 79 years with an average work experience of 27 years. A randomization check showed the demographic features to be uncorrelated with the manipulations.

4.3. Algorithm aversion

In our setting, algorithm aversion would express itself in how participants weight the two input forecasts when producing their own forecast to be reported to top management (see the notions of aversion in [Önköl et al., 2009](#), [Promberger & Baron, 2006](#), or [Dietvorst et al., 2015](#), all of which are based on the weighting of advice given by humans as opposed to algorithms).

Aversion to algorithms as a source's origin is expected to result in discounting the algorithmic forecast, which in turn leads to a reported forecast shifted away from the algorithmic forecast towards the human forecast. In our experiment, algorithm aversion should manifest itself as follows (see [Hogarth & Einhorn, 1992](#)): In the EQUAL conditions, a rational participant, weighting both forecasts equally, would report 5.1% which is the mid-point between the forecast of the management accountant as default source (4.2% sales growth rate) and the alternative (5.9% sales growth rate). In the BETTER conditions, a rational participant would report 5.9%, which is the information provided by the more accurate forecaster 2, regardless of its origin (human or algorithm). Under algorithm aversion, we would expect these results to hold for both HUMAN conditions (HUMAN-EQUAL and HUMAN-BETTER). But a participant averse to using the algorithmic forecast would in both ALGO conditions assign greater weight on the human forecaster 1 and, as a result, report a forecast closer to 4.2% compared to 5.1% (ALGO-EQUAL) or to 5.9% (ALGO-BETTER).

Our results are visualized in [Figure 1](#) that shows the means (M) and 95% confidence intervals (CI) for forecasts obtained in the four experimental conditions. The descriptive statistics for the dependent variable are given in [Appendix B](#).

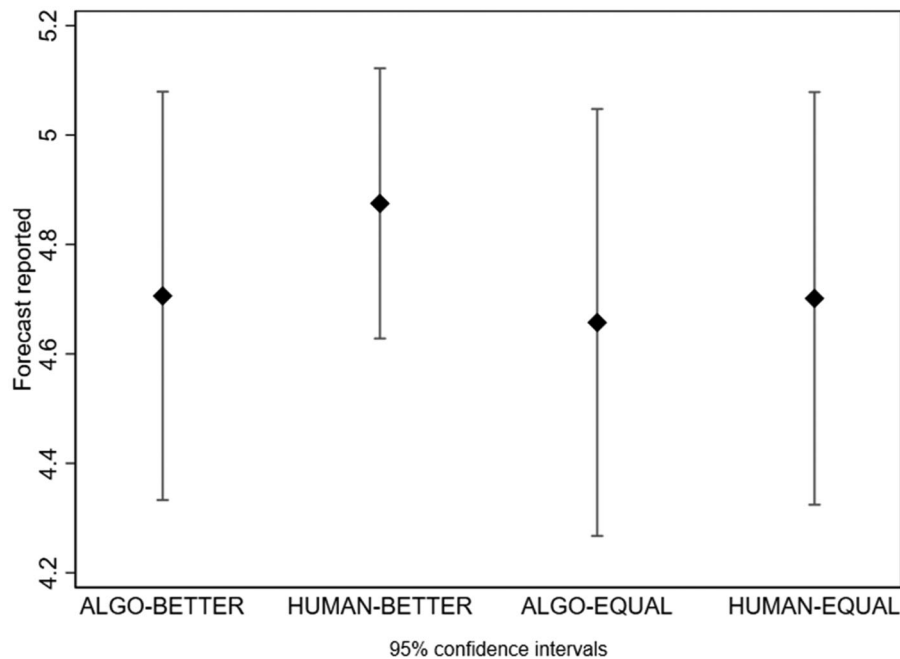


Figure 1. Reported forecasts for sales growth rates by experimental condition.

Notes: The forecast from the human expert is for 4.2% sales growth, the alternative (algorithm or another human forecaster) for 5.9%. ALGO-BETTER: mean (M): 4.71%; 95% confidence interval (CI): 4.34%–5.07%; HUMAN-BETTER: M: 4.88%; CI: 4.64%–5.12%; ALGO-EQUAL: M: 4.66%; CI 4.28%–5.04%; HUMAN-EQUAL: M: 4.70; CI: 4.34%–5.07%.

If one source has been more accurate in the past, a rational participant caring only about the accuracy of the source, is expected to report the forecast coming from the more accurate source. If both forecasts have been equally accurate in the past, a rational participant is expected to put equal weight on both sources and report the average of both input forecasts (see Hogarth & Einhorn, 1992). Aversion to algorithms as a source's origin is expected to result in discounting the forecast of this source, which results in a reported forecast shifted away from the source subject to aversion.

As is evident, the weightings of the inputs for the reported forecasts vary among experimental conditions, though not significantly as the confidence intervals overlap. The positioning of all reported forecasts within the range defined by the two input forecasts (4.2% to 5.9%) indicates a slight tendency to prioritize the 4.2%-forecast from the first source, as the means in all experimental conditions remain below the midpoint (~5.0). However, the alternative source is never discounted completely, as the average forecasts in all conditions differ significantly from the first source (4.2%, which is not in the 95% confidence interval, CI). We controlled for risk propensity, but we found no significant effect on the reported forecast.

Our finding contrasts with the results in experiments using the standard measurement approach as proposed by Dietvorst et al. (2015). According to their results, algorithm aversion should be reflected in a constant negative effect in our measurement, but we find that there are no differences in the response to human vs. algorithmic information. Thus, there is no evidence of algorithm aversion per se.

It follows, as we have argued in the theoretical section, that an experimental design that replaces a personally made forecast with a forecast coming from an algorithm, reflects two other mechanisms that are distinct from algorithm aversion in the narrow sense of the meaning. First, these results are subject to psychological ownership effects, which are excluded in our design. Second, a forecast provided by an external source is always less transparent compared to a forecast made by the participants themselves based on raw data and a mental model. Both mechanisms induce aversion, but it is rather an aversion either to giving up a forecast made with personal effort or to using an intransparent forecast provided by an unquestionable black-box mechanism.

Table 1. Perceived performance of forecasting sources.

	All	Algorithm	Human
BETTER	0.336	0.206	0.462*
ALGO	-0.365*		
constant	4.242***	3.941***	4.176***
R ²	0.039	0.006	0.040
R ² _{adj}	0.025	-0.009	0.026
N	138	68	70

Legend: * $p < 0.1$; ** $p < 0.01$; *** $p < 0.001$. R^2 = coefficient of determination, R^2_{adj} = adjusted R^2 ; N = number of participants.

Notes: Results obtained by an OLS regression. Dependent variable: Perceived Performance, i.e., participant's belief regarding the past performance of the two sources, 7-point Likert scale with high values indicating the belief that the second forecast performed better in the past. BETTER: Dummy variable indicating that the second forecast was objectively better in the past (reference: equal objective performance). ALGO: Dummy variable, indicating that the second forecast was produced by an algorithm (reference: forecast was produced by a human).

4.4. Mechanism 1: Biased performance perception

We hypothesized that even though basically the perception of a source's performance drives usage (H1a), users are biased against algorithms and pay attention and recognize superior performance if the source's origin is human and ignore it if the origin is algorithmic (H1b).

Starting with the latter argument, Table 1 provides results on determinants of perceived performance of the second forecasting source relative to the first source.

Perceived performance was measured as the participants' beliefs that one of the sources was superior to the other in the past, on a 7-point Likert scale ranging from 1 (the first source, i.e., the management accountant A, was clearly perceived to have performed better in the past) to 7 (the second source, i.e., either another human expert or an algorithm, was perceived to have performed better) with the mid-point of 4 indicating belief that both sources performed equally in the past. Perceived performance was regressed on the origin of the second forecast source (ALGO) and whether the second forecast was actually better in the past (BETTER).

Prima facie, there is a significant effect of a forecast originating from an algorithm being perceived as performing worse (see panel "All"). But this effect is due to differences in the participants' perception of performance: Participants perceive the objectively higher performance if the forecaster is human (panel "Human", i.e., the second source was from human origin) but ignore it if the forecast originates from an algorithm (panel "Algorithm", i.e., the second source was from algorithmic origin), thus supporting H1b.

We then test H1a by inquiring into the relevance of perceived performance for usage. To do so, we regress the participants' reported forecast on perceived performance (PerceivedPerformance), the objective past performance (BETTER) and the nature of the source (ALGO). As the first forecast was a sales growth rate of 4.2% and the second forecast of 5.9%, positive coefficients indicate a higher weight for the forecast from the second source, characterized by its type and perceived performance.

The regression coefficients in the panel ("All") indicate that neither the origin of the forecast per se nor the objective performance of the source matters for usage. Any algorithm aversion would show up as a negative coefficient for the dummy variable ALGO, as participants would then place more weight on the first forecast. But we see no such effect. Instead, only perceptions of performance matter and result in a bias against algorithmic advice.

Overall, the results in panel "All" indicate that perceived performance increases usage, supporting H1a. Nevertheless, as hypothesized in H1b, further inquiry shows that this is not always the case: As columns for both panels "Algorithm" and "Human" in Table 2 indicate, perceived performance matters for usage only if the human expert provides the forecast. For a forecast produced by an algorithm, objective superior performance translates neither into perceived performance nor into increased usage.

Taken together, these findings support H1a and H1b and identify biased performance perceptions as a first relevant mechanism explaining inappropriate usage of algorithmic advice.

4.5. Mechanism 2: Trust

A second mechanism driving algorithm aversion suggested in existing research is lack of trust in algorithms. Specifically, the origin of the source of advice – either human or algorithm – is supposed to

Table 2. Use of alternative forecasts.

	All	Algorithm	Human
PerceivedPerformance	0.234***	0.137	0.365***
BETTER	0.033	0.020	0.005
ALGO	-0.022		
constant	3.740***	4.116***	3.177***
R ²	0.092	0.031	0.215
R ² _{adj}	0.072	0.000	0.192
N	137	67	70

Legend: * $p < 0.1$; ** $p < 0.01$; *** $p < 0.001$. R^2 = coefficient of determination, R^2_{adj} = adjusted R^2 ; N = number of participants.

Notes: Results obtained by an OLS regression. Dependent variable: Usage measured by the forecast reported. PerceivedPerformance: Participants' belief regarding the past performance of the two sources, 7-point Likert scale with high values indicating the belief that the second forecast performed better in the past. BETTER: Dummy variable indicating that the second forecast was objectively better in the past (reference: equal objective performance). ALGO: Dummy variable, indicating that the second forecast was produced by an algorithm (reference: forecast was produced by a human).

Table 3. Determinants of general and procedural trust.

	Personal	Correct	Complete	Interpretation
PerceivedPerformance	0.710***	0.625***	0.390***	0.390***
BETTER	-0.096	0.022	0.203	0.152
ALGO	-0.001	-0.004	-0.015	-0.405*
constant	1.003**	1.215***	2.175***	2.444***
R ²	0.454	0.355	0.151	0.211
R ² _{adj}	0.442	0.341	0.132	0.193
N	138	138	138	138

Legend: * $p < 0.1$; ** $p < 0.01$; *** $p < 0.001$. R^2 = coefficient of determination, R^2_{adj} = adjusted R^2 ; N = number of participants.

Notes: Results obtained by OLS regression. Dependent variables: trust variables, see text. All were measured by 7-point Likert scales, with high values indicating that the participants believed the second forecast to be more trustworthy in this regard. PerceivedPerformance: Participants' belief regarding the past performance of the two sources, 7-point Likert scale with high values indicating the belief that the second forecast was perceived as having performed better in the past. BETTER: Dummy variable indicating that the second forecast was objectively better in the past (reference: equal objective performance). ALGO: Dummy variable, indicating that the second forecast was produced by an algorithm (reference: forecast was produced by a human).

affect users' trust in the forecast produced (H2b) and this trust is then supposed to affect usage (H2a). To test this mechanism in our experiment, we consider on the one hand participants' overall trust in the source by asking whether the participant personally had higher trust in either the first or the second forecast ("Which of the two forecasts do you personally think is more trustworthy?"/"Personal").

But as such a notion of trust is rather broad and thus has been criticized in existing research (e.g., Choung et al., 2023), we also capture procedural aspects of trust using the approach suggested by Bhattacharjee (2002), and survey participants' opinions on different dimensions of the forecasting process. Did they expect the data to be gathered correctly ("For which of the two forecasts do you have more confidence that the information and data that goes into the forecast have been determined correctly?"/"Correct")? Did they believe that all relevant data was collected ("Which of the two forecasts do you have more confidence in in terms of the completeness of the information base? In other words, which forecast do you think has taken all relevant data from all relevant areas into account when determining its values?"/"Complete")? Was in their view the data underlying the forecast correctly interpreted ("Which of the two forecasts do you have more confidence in, knowing that the data and information collected have been correctly interpreted for forecasting purposes?"/"Interpretation")?

All four trust variables were measured by 7-point Likert scales, with high values indicating that the participants believed the second forecast to be more trustworthy in this regard. In line with the established methodological literature on measurement, the aspects were captured using single items (Allen et al., 2022; Wanous & Reichers, 1996). Further, a 7-point scale is sufficiently sensitive for measurements (Norman, 2010).

Table 3 gives the results of regressing general and procedural trust on the origin (ALGO), objective performance (BETTER) and perceived performance (PerceivedPerformance) of the second forecast. The constellation, in which the second forecast was also produced by a human, serves as a reference category.

Table 4. Trust as a determinant of usage.

	All		Algorithm		Human	
	1	2	1	2	1	2
PerceivedPerformance	0.132	0.240**	0.015	0.053	0.350**	0.389***
BETTER	0.047	0.010	0.021	0.015	0.009	-0.075
ALGO	-0.021	-0.004				
Personal	0.143		0.216		0.017	
Correct		-0.092		0.054		-0.179
Complete		0.091		0.147		-0.053
Interpretation		0.042		-0.036		0.214
constant	3.596***	3.552***	3.784***	3.823***	3.174***	3.132***
R ²	0.113	0.105	0.070	0.071	0.215	0.259
R ² _{adj}	0.086	0.064	0.025	-0.006	0.180	0.201
N	137	137	67	67	70	70

Legend: * $p < 0.1$; ** $p < 0.01$; *** $p < 0.001$. R^2 = coefficient of determination, R^2_{adj} = adjusted R^2 ; N = number of participants.

Notes: Results obtained by OLS regression. Dependent variable: Forecast reported. PerceivedPerformance: Participants' belief regarding the past performance of the two sources, 7-point Likert scale with high values indicating the belief that the second forecast performed better in the past. BETTER: Dummy variable indicating that the second forecast was objectively better in the past (reference: equal objective performance). ALGO: Dummy variable, indicating that the second forecast was produced by an algorithm (reference: forecast was produced by a human). Personal: Participants' general trust in the source. Correct: Participant's trust that correct data were used to produce the forecast. Complete: Participants' trust that data gathered for the forecast were complete. Interpretation: Participants' trust that the data were interpreted correctly for formulating the forecast. All variables are measured on 7-point Likert scales with high values indicating that the alternative source received more trust.

First, the results indicate that the origin of a forecast (ALGO) affects trust only regarding correct interpretation of the data. Participants believe this to be less the case for an algorithm (see column "Interpretation"). Overall, algorithms do not significantly induce less trust than humans. Thus, there is only limited support for hypothesis 2b.

Second, past performance matters for trust, but not straightforward, as there is no effect of an alternative being objectively better (BETTER) on any trust variable, but only of perceived performance. Instead, what matters is the participants' belief that the second forecast is superior. The more participants perceive the second forecast as outperforming the first, the more trust they have in this source, in both, general and procedural aspects. However, whether participants perceive this to be the case depends on some degree, as we saw earlier on, on the source, as superior performance is perceived in the case of humans, but not in the case of algorithms.

Regarding the mechanism's second step, the link between trust and usage, Table 4 reports the effects of different aspects of trust on usage, measured as the forecast reported. We ran two separate regressions, one for overall trust ("Personal"), one for the three aspects of procedural trust ("Correct", "Complete", "Interpretation"), to obtain a differentiated picture. In terms of empirical expectations, more trust in the alternative source should shift the reported combination of the two input forecasts towards the forecast from the second source, i.e., from 5% towards 5.9%. We therefore expect positive coefficients for the trust variables.

Table 4 shows that contrary to much of the existing research and our hypothesis 2a, trust is largely irrelevant for factual usage (see panel "All"). Neither trust in general (columns 1), nor procedural aspects of trust (columns 2) significantly affect usage. Even more interesting, this pattern applies equally to alternative forecasts originating from humans and algorithms (see panels "Algorithm" and "Human"). Thus, there is no evidence supporting the origin-trust-usage mechanism, as there is only a weak and highly specific link from a forecast's origin to trust and basically no link from trust to usage.

4.6. Mechanism 3: Presumed trust of others

External barriers, notably in the organizational context, may prevent users from using algorithms. A first barrier may arise from the organizational hierarchy, as participants are asked to put themselves in the situation of the division's head of finance and provide a forecast for the division's sales growth rate to the company's top management. As pointed out, we argue that if users in a firm assume their superiors to be averse to algorithmic advice, they will not use algorithms, regardless of their personal opinion on the issue. To evaluate this mechanism, we capture the impact of the source of a forecast on the

Table 5. Relationship between origin, presumed trust and usage.

	Panel A: Origin and presumed trust			Panel B: Presumed trust and usage		
	All	Algorithm	Human	All	Algorithm	Human
BETTER	-0.046	0.117	-0.269	0.033	0.024	0.004
ALGO	-0.235			-0.023		
PerceivedPerformance	0.433***	0.289*	0.636***	0.236**	0.148	0.368***
PresumedTrust				-0.004	-0.037	-0.005
constant	2.184***	2.448***	1.404*	3.750***	4.205***	3.184***
R ²	0.170	0.081	0.280	0.092	0.033	0.215
R _{adj} ²	0.152	0.053	0.259	0.065	-0.013	0.179
N	138	68	70	137	67	70

Legend: * $p < 0.1$; ** $p < 0.01$; *** $p < 0.001$. R^2 = coefficient of determination, R_{adj}^2 = adjusted R^2 ; N = number of participants.

Notes: Results obtained by OLS regression. PerceivedPerformance: Participants' belief regarding the past performance of the two sources, 7-point Likert scale with high values indicating the belief that the second forecast was perceived as having performed better in the past. BETTER: Dummy variable indicating that the second forecast was objectively better in the past (reference: equal objective performance). ALGO: Dummy variable, indicating that the second forecast was produced by an algorithm (reference: forecast was produced by a human). Panel A: Dependent variable: PresumedTrust, defined as the level of trust in a forecast the participant presumes the top management as ultimate recipient of the forecast to have. Panel B: Dependent variable: Usage of the forecasts, measured as reported forecast, for explanatory variables see above. All dependent variables are measured on 7-point Likert scales, high values indicating that the second source received more trust.

trust the participant presumes the recipient to have, and of this “presumed trust” on the participant’s usage of forecasts.

If algorithm aversion indeed operates through this mechanism, we expect to find an effect of the forecast’s origin on the trust participants presume the top management as recipient to have (H3b) and to find that this presumed trust affects factual usage (H3a). Table 5 indicates, at best, mixed support for this mechanism.

Panel A gives the results of regressing presumed trust on the origin and perceived performance of the second forecast. While in the case of an algorithm, participants presume top management to be somewhat less trusting, the effect does not reach significance (see column “All”). Thus, participants do not systematically presume the top management to be more skeptical regarding an algorithm-based forecast than a human-made forecast. Again, algorithm aversion expresses itself more indirectly, via perceived performance: If participants personally believe a source to be more accurate (PerceivedPerformance), they tend to presume that the top management as recipient would also have more trust in this source. Perceived performance is relevant for presumed trust (see panel A, column “All”), but much more so in the case of a human than in the case of an algorithm (see panel A, columns “Algorithm”, and “Human”, respectively). As for the second step of the mechanism, the relevance of presumed trust for usage, panel B indicates that presumed trust is irrelevant for usage. In no constellation, regardless of whether all cases or the subsamples are considered, the variable PresumedTrust exerts an effect on usage. Even if participants presume that the ultimate recipients of a forecast are skeptical of algorithms, this presumption does not affect how the output of an algorithm is used. Thus, participants’ presumption that the ultimate recipients would not accept usage of an algorithm is no obstacle. As a result, neither H3a nor H3b are supported.

5. Experiment 2

5.1. Design modifications

The second experiment evaluates whether usage of algorithms is obstructed by the black box nature of algorithms. Our research design and experimental setting were adapted from the first experiment with two modifications.

First, the difference in performance was increased. In the EQUAL condition, both forecasts had in total an absolute deviation of 16.1 percentage points for the periods under consideration, in the BETTER condition, the first forecaster, the human, had in total an absolute deviation of 19.8 percentage points, the better forecasting source (of human vs. algorithmic origin) had in total an absolute deviation of 6.3 percentage points. The values were chosen in contact with practitioners to be both, perceptibly different but still realistic.

Second, we added an explanation elaborating on how the second forecast came about. The explanation was given for the second source identically in both cases, i.e., for the second human forecaster (“management accountant B”) as well as for the algorithm. The explanation elaborates on the forecasting process, but provide information relating to the content of the forecast (see Appendix C).

The other stimuli – information on the nature and past accuracy of the forecasting sources – were provided in the same format as in experiment 1 (see Appendices A and C). Experiment 2 was pre-tested using students as participants, in particular regarding the plausibility of the explanation. For the quarter in question, the first (human) forecast from a management accountant was for a sales growth rate of 3.4%, the second forecast (either of human or of algorithmic origin) was for a sales growth rate of 4.9%. We would expect that a rational participant in the EQUAL condition would forecast the average of both values (4.2%) and in the BETTER condition an increased value shifting towards 4.9%. Aside from these issues, the dependent variable as well as the post-experimental questionnaire were identical to experiment 1.

5.2. Sample characteristics

Again, practitioners with a background in management accounting or other functions in business who can relate to the situation described in the experiment, were recruited via a provider of online surveys, taking care to recruit only subjects that did not participate in experiment 1. Overall, 204 persons participated with an average age of 45 years, ranging from 19 to 68 years. 41% of the participants were male, 38% held a university degree, of which about a fifth (8% of all participants) held a degree in economics or business administration. About 19% of our participants worked in sales and distribution functions and 4% in management accounting. The average work experience of all participants was 20 years.

The descriptive statistics of the dependent variable (forecast of sales growth rate for the upcoming quarter) for the eight experimental conditions are given in Appendix D. The findings in experiment 2 (not reported) on the relationships between origin, perceived performance, personal trust and presumed trust and usage replicate the findings from experiment 1, with the increased difference in performance resulting in an increased perception of performance. As in experiment 1, this matters only in the case of a human forecaster.

5.3. Mechanism 4: Explanation and usage

In terms of what a barrier to usage might be, we argue that an explanation of how a forecast was made increases usage (H4a) but that algorithms are more in need of an explicit explanation (H4b). Humans implicitly offer the option to obtain an explanation anytime, whereas algorithms do not. Results on the relevance of providing an explanation for factual usage are given in Table 6.

Contrary to H4a, providing an explanation on how the forecast came about does not increase usage of the source for which the explanation is provided (see panel “All”). Even though the coefficient obtained for EXPLANATION in the subsample, where the second forecast was from an algorithm, is

Table 6. Effect of explanation on usage.

	All	Algorithm	Human
ALGO	0.003		
PerceivedPerformance	0.152**	0.155	0.149*
EXPLANATION	0.211	0.228	0.194
constant	3.218***	3.197***	3.239***
R ²	0.047	0.041	0.053
R _{adj} ²	0.032	0.021	0.033
N	200	100	100

Legend: * $p < 0.1$; ** $p < 0.01$; *** $p < 0.001$. R^2 = coefficient of determination, R_{adj}^2 = adjusted R^2 ; N = number of participants.

Notes: Results obtained by OLS regression. Dependent variable: Forecast reported. Management accountant A predicted 3.4%, the second forecast (management accountant B vs. algorithm) was 4.9%. PerceivedPerformance: Participants' belief regarding the past performance of the two sources, 7-point Likert scale with high values indicating the belief that the second forecast was better in the past. ALGO: Dummy variable, indicating that the second forecast was produced by an algorithm (reference: forecast was produced by a human). EXPLANATION: Dummy variable, indicating that an explanation of how the forecast came about was presented (reference: no explanation provided).

slightly higher (see panel “Algorithm”), it is insignificant. This is also the case for the situation where the second forecast originates from a human expert (see panel “Human”). Thus, there is no support for H4b, which stated that an explanation matters more for algorithms than for humans. While not reported here, we also analyzed whether the provision of an explanation affects perceptions of performance, but found no such effect.

6. Discussion and conclusion

Improvements in the capabilities of algorithms have the potential to make the integration of algorithmic output into decision-making an option for increasing firms’ economic success (Brink et al., 2024; Enholm et al., 2022; KPMG, 2024; Mishra et al., 2022). However, research since the 1950s until today provides evidence of algorithm aversion (Baines et al., 2024; Dietvorst et al., 2015; Meehl, 1954), implying a general reluctance towards using advice produced by technologies like artificial intelligence. Our research objective is to investigate in how strong this aversion actually is in firms with a specific anchor in forecasting in accounting and control and on what behavioral mechanism it is based.

Before discussing the implications of our findings, we need to discuss the study’s limitations. First, as in all experimental studies, the setting used is to some degree stylized and abstract, and the situation’s specificity may affect the generalizability of the results. The design was pretested with experienced practitioners, who indicated that the task was straightforward and comprehensible. In terms of realism, input from practitioners indicated furthermore, that the situation chosen is common in practice, as users in functions as the one we describe in the experiment often receive forecasts instead of producing forecasts from raw data. The explanation used in experiment 2 was also pre-tested in terms of being comprehensible, and was found to be so. Its limited impact on usage is an empirical result, not an artifact arising from presenting an incomprehensible explanation. Still, another type of explanation might lead to different results (Miller, 2019). Second, in terms of sampling, we sampled practitioners with substantial working experience and in particular with experience with the managerial task used. Most practitioners we interviewed in the run-up to the study reported, regardless of their formal function, that having to make forecasts is part of their job, albeit not in the formal way of a management accountant being in charge of producing forecasts for the management. We also tested for effects of professional background on usage, but found no effects. Third, the stakes in the experiment – the consequence of reporting a sales forecast which turns out to be wrong – for the participants are presumably lower than they would be in real life. This, however, would imply in our view an even higher handicap for a new technology, and an even higher tendency to rely on the forecast produced by a human, in particular due to the possibilities of blame shifting in the case that problems occur (e.g., Gazit et al., 2023). As it is, we followed up to the arguments on the size of the stakes in testing for effects of risk attitudes on usage. We found no such effects. Fourth, the results are potentially affected by the sequence in which the manipulations and the post-experimental questions, e.g., on trust, were presented. However, if we had asked about trust in algorithms first and then presented an algorithmic forecast, the very fact that we asked about trust may triggered considerations which affect usage. We thus acknowledge that our results – as the results in every experimental research – may be affected by the research design used. Finally, a one-shot-experiment (even though we were able to replicate our results from experiment 1 in experiment 2 with different participants) may not be conclusive about how trust and usage develop over time, with personal experience gained in repeated dealings, and how the need for supplementary explanations develops in the long run. Given that trust in various aspects of the algorithmic forecast varied substantially among participants, but was still irrelevant for usage, we may assume that increasing trust will not turn out to be relevant. Other personal features, such as experience with algorithms, may also be of relevance for algorithm usage, but in line with the standard in experimental studies, see Piercey (2023) for the statistical arguments, we did not include them as covariates as we do not formulate hypotheses about them.

Taken these limitations into account, our findings offer new insights and call for a discussion of existing findings, also because of their implications for practice.

First, if participants are free to use and to combine human and algorithmic forecasts that are presented simultaneously, there is little evidence of algorithm aversion. This contrasts with research using the measurement of algorithm aversion in line with Dietvorst et al. (2015). In the light of the

substantial arguments on psychological ownership arising from the invested effort involved in this measurement approach, we believe it to be plausible that the aversion found using the traditional measurement is more an aversion to abandoning the output of one's own efforts than aversion to algorithms. We would also argue that our measurement approach is closer to real life. To validate our experiments and in particular the measurement approach used, we conducted interviews with practitioners on the usage of algorithms in their firms. Their consistent answer was that currently, and for the foreseeable future, human forecasters and algorithms will operate in parallel, just as in the situation we used in our experimental design and different to the design implemented in earlier studies. In real life, for users, in particular higher up in an organization's hierarchy, the choice will typically be between a forecast produced by humans, say the management accountant in charge, and of an additional decision support system providing algorithm-based forecasts. They do not have to choose between their own forecast and an alternative. Both, the theoretical arguments on psychological ownership effects and the view of practitioners imply that the situation we use is closer to practice. This also suggests that the magnitude of algorithm aversion is more limited in practice than hitherto presumed.

Second, as for existing aversion, there is substantial support for a more indirect behavioral mechanism. Algorithm aversion operates at least in part by users not paying attention to an algorithm's superior performance. Comparing experiments 1 and 2, we find that if differences in past performance are more pronounced, this tendency is reduced. Still, to be perceived as outperforming, an algorithm has to outperform a human forecaster to a higher degree than one human forecaster has to outperform another human forecaster. This selective attention to performance is located very early in the decision processes on usage of algorithms than, for instance, trust. Trust strongly features in existing research (e.g., Glikson & Williams Woolley, 2020), but we found it to be largely irrelevant. In this, we are in line with more recent research (e.g., Schaap et al., 2024).

Third, there is strong evidence that the perception of forecasting performance matters conditional on the nature of the forecasting source. Participants follow their belief about the performance in the case of a human and rely more on the source they perceive to be better. But they act against their belief in the case of an algorithm by ignoring an objectively superior performance which they also perceive as superior. This mechanism is ignored in extant research, which typically does neither measure nor and use perceptions of performance.

We suggested that there are barriers external to the user, which prevent usage of algorithmic forecasts. A first barrier, often discussed in the literature, is the black box nature of the algorithm (Nakashima et al., 2024; Schmidt et al., 2020). There are persistent calls for supplementing algorithmic tools with explanation facilities, such as LIME (Bitzer et al., 2023), and considerable amounts of time and skill are invested into the development of explanations (Ali et al., 2023). Empirically, we find that opening up the "black box" by providing an explanation does not increase reliance on advice. This finding is in line with some very recent research on transparency, which also found no effect of transparency on usage (e.g., Adadi & Berrada, 2018; Ning et al., 2024; Schmidt et al., 2020; Shin, 2021). It is also not the case that algorithms *ceteris paribus* profit more (or less) from an explanation than human forecasters. This finding is of interest, as the consequences of intransparency in humans were up to now not compared with intransparency of algorithms. There were reasons to expect differences, as we argued in hypothesis 4b, as, in principle, humans providing input may be contacted to provide an explanation at any time. This option does not exist for algorithms, which usually do not provide explanations, thus, having an explanation for an algorithmic forecast is more relevant. We see that for the situation we present, combining forecasts and report them, there is no effect of an explanation and no difference in the relevance of an explanation. This implies, that for the participants, the focus was on the forecast, not on the reasons for the forecast. This finding is also in line with the finding that participants orient themselves at the perceived performance of a source. Still, to conclude, the relevance of explanations for usage is a theme which is need of additional research, as one might expect that the need to be able to justify a forecast, even if it is based on input, rather than on a personal analysis of raw data, differs among users of forecasts.

A second barrier we suggested is presumed algorithm aversion of others (Arkes et al., 2007; Efendić et al., 2024; Longoni et al., 2019; Promberger & Baron, 2006; Shaffer et al., 2013), in our case of managers higher up in the hierarchy. While participants indeed presume their superiors to be more skeptical

of algorithms, this does not affect their usage of algorithmic output. We take this to indicate that the effect of the organizational environment on usage of algorithms is weak compared to early studies, e.g., Arkes et al. (2007). This may result from the trend that artificial intelligence rapidly becomes more common also in private lives and thus acceptable.

In terms of practical implications of our findings, firms should not worry that employees or managers discard algorithmic information or advice simply because it is not coming from a human expert. But another impediment remains because there might be a selective attention bias with respect to the algorithms's performance. This could be a problem, especially, if performance differences are small. Simply presuming, as most experimental research and most proponents of algorithms do, that users do pay attention to objective differences in performance and are aware of them, is insufficient.

Given thus the crucial role of performance perceptions identified by us, a first task for a firm interested in increasing the usage of algorithmic advice is to make users aware of algorithms' actual performance, e.g., by implementing showcases or by actively pointing out the performance of different sources in detail and also over a longer period. Several practical examples are documented in literature (e.g., Fehrenbacher et al., 2023; Göbel, 2023; Reitzenstein et al., 2023). Future research should address other potential barriers, which obstruct users from following their belief about what is the better forecaster. Here too, algorithms are disadvantaged.

Concluding with regard to the robustness of algorithm aversion, we find that when the ownership effect arising from personally producing a forecast is absent, algorithm aversion is much less pronounced than in existing research where ownership effects were present. While mechanisms relying on some form of trust – general trust in algorithms, trust in the processes by which an algorithm works – are central to existing research, we find basically no relevance for either of them. True, the origin of a forecast is of some relevance for isolated aspects of trust, but this trust is of no relevance for usage. Providing an explanation does increase neither usage of algorithmic nor of human output, we take this as a hint to other mechanisms, like the inability of algorithms to bear responsibility for their output. Central to our findings is the perception of performance, which is affected by the origin and of conditional relevance for usage.

Ethical standards

The survey was conducted in accordance with ethical standards. The participation was voluntary and the personal data collected in the questionnaire is used exclusively for scientific purposes and stored anonymously. For data collection, the General Data Protection Regulation was applied.

Author contributions

CRedit: **Peter Kotzian**: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft; **Barbara E. Weißenberger**: Conceptualization, Funding acquisition, Investigation, Methodology, Writing – original draft; **Sonja G. Prinz**: Conceptualization, Methodology.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Prof. Dr. Barbara E. Weißenberger and Dr. Peter Kotzian gratefully acknowledge financial support from the Jürgen Manchot Foundation, Düsseldorf, within the Manchot research group “Decision-making with the Help of Artificial Intelligence”.

ORCID

Peter Kotzian  <http://orcid.org/0000-0002-2141-4961>

Barbara E. Weißenberger  <http://orcid.org/0000-0002-4789-0595>

References

- Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
- Afsay, A., Tahriri, A., & Rezaee, Z. (2023). A meta-analysis of factors affecting acceptance of information technology in auditing. *International Journal of Accounting Information Systems*, 49, 100608. <https://doi.org/10.1016/j.accinf.2022.100608>
- 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>
- Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion*, 99, 101805. <https://doi.org/10.1016/j.inffus.2023.101805>
- Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science. A call to action. *European Journal of Psychological Assessment*, 38(1), 1–5. <https://doi.org/10.1027/1015-5759/a000699>
- Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of business analytics and enterprise systems on managerial accounting. *International Journal of Accounting Information Systems*, 25, 29–44. <https://doi.org/10.1016/j.accinf.2017.03.003>
- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk costs. *Organizational Behavior and Human Decision Processes*, 35(1), 124–140. [https://doi.org/10.1016/0749-5978\(85\)90049-4](https://doi.org/10.1016/0749-5978(85)90049-4)
- Arkes, H. R., Shaffer, V. A., & Medow, M. A. (2007). Patients derogate physicians who use a computer-assisted diagnostic aid. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 27(2), 189–202. <https://doi.org/10.1177/0272989X06297391>
- Aschauer, F., Sohn, M., & Hirsch, B. (2024). Managerial advice-taking - Sharing responsibility with (non)human advisors trumps decision accuracy. *European Management Review*, 21(1), 186–203. <https://doi.org/10.1111/emre.12575>
- Baines, J. I., Dalal, R. S., Ponce, L. P., & Tsai, H.-C. (2024). Advice from artificial intelligence: A review and practical implications. *Frontiers in Psychology*, 15, 1390182. <https://doi.org/10.3389/fpsyg.2024.1390182>
- Bakarich, K. M., & O'Brien, P. (2021). The robots are coming ... but aren't here yet: The use of artificial intelligence technologies in the public accounting profession. *Journal of Emerging Technologies in Accounting*, 18(1), 27–43. <https://doi.org/10.2308/JETA-19-11-20-47>
- Bareis, J. (2024). The trustification of AI. Disclosing the bridging pillars that tie trust and AI together. *Big Data & Society*, 11(2), 1–14. <https://doi.org/10.1177/20539517241249430>
- Bayer, S., Gimpel, H., & Markgraf, M. (2022). The role of domain expertise in trusting and following explainable AI decision support systems. *Journal of Decision Systems*, 32(1), 110–138. <https://doi.org/10.1080/12460125.2021.1958505>
- Bhattacharjee, A. (2002). Individual trust in online firms: Scale development and initial test. *Journal of Management Information Systems*, 19(1), 211–241.
- Bitzer, T., Wiener, M., & Cram, W. A., Technische Universität Dresden. (2023). Algorithmic transparency: Concepts, antecedents, and consequences – A review and research framework. *Communications of the Association for Information Systems*, 52, 293–331. <https://doi.org/10.17705/1CAIS.05214>
- Blackburn, R., Lurz, K., Priese, B., Göb, R., & Darkow, I.-L. (2015). A predictive analytics approach for demand forecasting in the process industry. *International Transactions in Operational Research*, 22(3), 407–428. <https://doi.org/10.1111/itor.12122>
- Blattberg, R. C., & Hoch, S. J. (1990). Database models and managerial intuition: 50% model + 50% manager. *Management Science*, 36(8), 887–899. <https://doi.org/10.1287/mnsc.36.8.887>
- Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, 101(2), 127–151. <https://doi.org/10.1016/j.obhdp.2006.07.001>
- Brink, A., Benyayer, L.-D., & Kupp, M. (2024). Decision-making in organizations: Should managers use AI? *Journal of Business Strategy*, 45(4), 267–274. <https://doi.org/10.1108/JBS-04-2023-0068>
- Bronfman, Z. Z., Brezis, N., Moran, R., Tsetsos, K., Donner, T., & Usher, M. (2015). Decisions reduce sensitivity to subsequent information. *Proceedings of the Royal Society B*, 282(1810). <https://doi.org/10.1098/rspb.2015.0228>
- Bronner, A. E., & de Hoog, R. (1984). The intertwining of information search and decision aiding. *Acta Psychologica*, 56(1–3), 125–139. [https://doi.org/10.1016/0001-6918\(84\)90013-1](https://doi.org/10.1016/0001-6918(84)90013-1)
- Casas-Arce, P., Cheng, M. M., Grabner, I., & Modell, S. (2022). Managerial accounting for decision-making and planning. *Journal of Management Accounting Research*, 34(1), 1–7. <https://doi.org/10.2308/jmar-10784>
- 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>
- Chaturvedi, A., Yadav, N., & Dasgupta, M. (2025). Tech-driven transformation: Unravelling the role of artificial intelligence in shaping strategic decision-making. *International Journal of Human-Computer Interaction*, 1–20. <https://doi.org/10.1080/10447318.2025.2456534>

- Chaxel, A.-S., Russo, J. E., & Kerimi, N. (2013). Preference-driven biases in decision makers' information search and evaluation. *Judgment and Decision Making*, 8(5), 561–576. <http://journal.sjdm.org/12/12227/jdm12227.pdf>. <https://doi.org/10.1017/S1930297500003661>
- Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction*, 39(9), 1727–1739. <https://doi.org/10.1080/10447318.2022.2050543>
- Choung, H., Seberger, J. S., & David, P. (2024). When AI is perceived to be fairer than a human: Understanding perceptions of algorithmic decisions in a job application context. *International Journal of Human-Computer Interaction*, 40(22), 7451–7468. <https://doi.org/10.1080/10447318.2023.2266244>
- Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P., & Malhotra, S. (2018). *Notes from the AI Frontier. Insights from hundred use cases*. McKinsey.
- Commerford, B. P., Dennis, S., Joe, J. R., & Ulla, J. W. (2022). Man versus machine: Complex estimates and auditor reliance on artificial intelligence. *Journal of Accounting Research*, 60(1), 171–201. <https://doi.org/10.1111/1475-679X.12407>
- Daschner, S., & Obermaier, R. (2022). Algorithm aversion? On the influence of advice accuracy on trust in algorithmic advice. *Journal of Decision Systems*, 31(sup1), 77–97. <https://doi.org/10.1080/12460125.2022.2070951>
- Davenport, T. H. (2018). Can we solve AI's 'trust problem'? *MIT Sloan Management Review*, 1–3. Retrieved September 3, 2020, from <https://sloanreview.mit.edu/article/can-we-solve-ais-trust-problem/>
- De Baets, S. (2021). Surveying forecasting: A review and directions for future research. *International Journal of Information and Decision Sciences*, 13(2), 111–126. <https://doi.org/10.1504/IJIDS.2021.116493>
- 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>
- Efendić, E., Van de Calseyde, P. P. F. M., Bahník, Š., & Vranka, M. A. (2024). Taking algorithmic (vs. human) advice reveals different goals to others. *International Journal of Human-Computer Interaction*, 40(1), 45–54. <https://doi.org/10.1080/10447318.2023.2210886>
- Eichholz, J., Knauer, T., & Winkelmann, S. (2023). Digital maturity of forecasting and its impact in times of crisis. *Schmalenbach Journal of Business Research*, 75, 443–481. <https://doi.org/10.1007/s41471-023-00171-1>
- Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial intelligence and business value: A literature review. *Information Systems Frontiers*, 24(5), 1709–1734. <https://doi.org/10.1007/s10796-021-10186-w>
- 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>
- Fildes, R., Goodwin, P., & Önkal, D. (2019). Use and misuse of information in supply chain forecasting of promotion effects. *International Journal of Forecasting*, 35(1), 144–156. <https://doi.org/10.1016/j.ijforecast.2017.12.006>
- Filiz, I., Judek, J. R., Lorenz, M., & Spiwojs, M. (2023). The extent of algorithm aversion in decision-making situations with varying gravity. *PloS One*, 18(2), e0278751. <https://doi.org/10.1371/journal.pone.0278751>
- Gazit, L., Arazy, O., & Hertz, U. (2023). Choosing between human and algorithmic advisors: The role of responsibility sharing. *Computers in Human Behavior: Artificial Humans*, 1(2), 100009. <https://doi.org/10.1016/j.chbah.2023.100009>
- Giboney, J. S., Brown, S. A., Lowry, P. B., & Nunamaker, J. F. (2015). User acceptance of knowledge-based system recommendations: Explanations, arguments, and fit. *Decision Support Systems*, 72(1), 1–10. <https://doi.org/10.1016/j.dss.2015.02.005>
- Glikson, E., & Williams Woolley, A. (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>
- Göbel, S., Arbeitskreis Digital Reporting der Schmalenbach-Gesellschaft. (2023). Implementierung von Predictive Analytics im Controlling: Herausforderungen und Erfolgsfaktoren. In Arbeitskreis Digital Reporting der Schmalenbach-Gesellschaft (Ed.), *Digital reporting: Transformation des Controllerbereichs durch den digitalen Wandel* (pp. 36–40). Franz Vahlen.
- Gönül, M. S., Önkal, D., & Goodwin, P. (2012). Why should I trust your forecasts? *Foresight: The International Journal of Applied Forecasting*, 27, 5–9.
- Grimmelikhuisen, S. (2023). Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review*, 83(2), 241–262. <https://doi.org/10.1111/puar.13483>
- Haesebrouck, K. (2021). The effects of information acquisition effort, psychological ownership, and reporting context on opportunistic managerial reporting. *Contemporary Accounting Research*, 38(4), 3085–3112. <https://doi.org/10.1111/1911-3846.12712>
- Hales, J. (2007). Directional preferences, information processing, and investors' forecasts of earnings. *Journal of Accounting Research*, 45(3), 607–628. <https://doi.org/10.1111/j.1475-679X.2007.00247.x>

- Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *The Academy of Management Review*, 9(2), 193–206. <https://doi.org/10.5465/amr.1984.4277628>
- Harrell, E. (2016). Managers shouldn't fear algorithm-based decision making. Harvard Business Review. <https://hbr.org/2016/09/managers-shouldnt-fear-algorithm-based-decision-making>
- Himmelstein, M., & Budescu, D. V. (2023). Preference for human or algorithmic forecasting advice does not predict if and how it is used. *Journal of Behavioral Decision Making*, 36(1), e2285. <https://doi.org/10.1002/bdm.2285>
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief adjustment model. *Cognitive Psychology*, 24(1), 1–55. [https://doi.org/10.1016/0010-0285\(92\)90002-J](https://doi.org/10.1016/0010-0285(92)90002-J)
- Johnson, E., Petersen, M. W., Sloan, J., & Valencia, A. (2021). The interest, knowledge, and usage of artificial intelligence in accounting: Evidence from accounting professionals. *Accounting & Taxation*, 13(1), 45–58.
- Jung, M., & Seiter, M. (2021). Towards a better understanding on mitigating algorithm aversion in forecasting: An experimental study. *Journal of Management Control*, 32(4), 495–516. <https://doi.org/10.1007/s00187-021-00326-3>
- Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. *Proceedings of the 28th European Conference on Information Systems (ECIS)*, June 15–17, 2020.
- Kaufmann, E., Chacon, A., Kausel, E. E., Herrera, N., & Reyes, T. (2023). Task-specific algorithm advice acceptance: A review and directions for future research. *Data and Information Management*, 7(3), 100040. <https://doi.org/10.1016/j.dim.2023.100040>
- Keding, C., & Meissner, P. (2021). Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions. *Technological Forecasting and Social Change*, 171, 120970. <https://doi.org/10.1016/j.techfore.2021.120970>
- Köchling, A., Riazzy, S., Wehner, M. C., & Simbeck, K. (2021). Highly accurate, but still discriminatory. A fairness evaluation of algorithmic video analysis in the recruitment context. *Business & Information Systems Engineering*, 63(1), 39–54. <https://doi.org/10.1007/s12599-020-00673-w>
- KPMG. (2024). *6 Industry use cases for intelligent forecasting*. KPMG.
- Krieger, F., Drews, P., & Velte, P. (2021). Explaining the (non-) adoption of advanced data analytics in auditing: A process theory. *International Journal of Accounting Information Systems*, 41, 100511. <https://doi.org/10.1016/j.accinf.2021.100511>
- Leitner, J., & Leopold-Wildburger, U. (2011). Experiments on forecasting behavior with several sources of information – A review of the literature. *European Journal of Operational Research*, 213(3), 459–469. <https://doi.org/10.1016/j.ejor.2011.01.006>
- Li, X., Hess, T. J., & Valacich, J. S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, 17(1), 39–71. <https://doi.org/10.1016/j.jsis.2008.01.001>
- Lim, J. S., & O'Connor, M. (1995). Judgemental adjustment of initial forecasts: Its effectiveness and biases. *Journal of Behavioral Decision Making*, 8(3), 149–168. <https://doi.org/10.1002/bdm.3960080302>
- Liu, J., Wang, H., Hui, C., & Lee, C. (2012). Psychological ownership: How having control matters. *Journal of Management Studies*, 49(5), 869–895. <https://doi.org/10.1111/j.1467-6486.2011.01028.x>
- 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>
- London, A. J. (2019). Artificial intelligence and black-box medical decisions: Accuracy versus explainability. *The Hastings Center Report*, 49(1), 15–21. <https://doi.org/10.1002/hast.973>
- Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629–650. <https://doi.org/10.1093/jcr/ucz013>
- Lösse, L. J., & Weißenberger, B. E. (2023). Using interpretable machine learning for accounting fraud detection – A multi-user perspective. *Die Unternehmung*, 77(2), 113–133. <https://doi.org/10.5771/0042-059X-2023-2-113>
- Lui, A. K. H., Lee, M. C. M., & Ngai, E. W. T. (2022). Impact of artificial intelligence investment on firm value. *Annals of Operations Research*, 308(1-2), 373–388. <https://doi.org/10.1007/s10479-020-03862-8>
- Maasland, C., & Weißmüller, K. S. (2022). Blame the machine? Insights from an experiment on algorithm aversion and blame avoidance in computer-aided human resource management. *Frontiers in Psychology*, 13, 779028. <https://doi.org/10.3389/fpsyg.2022.779028>
- 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>
- Mayer, B., Fuchs, F., & Lingnau, V. (2023). Decision-making in the era of AI support - How decision environment and individual decision preferences affect advice-taking in forecasts. *Journal of Neuroscience, Psychology, and Economics*, 16(1), 1–11. <https://doi.org/10.1037/npe0000170>
- Meehl, P. E. (1954). *Clinical versus statistical prediction*. University of Minnesota.
- Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L., & Canhoto, A. (2018). Big data, big decisions: The impact of big data on board level decision-making. *Journal of Business Research*, 93, 67–78. <https://doi.org/10.1016/j.jbusres.2018.08.029>

- 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>
- Mishra, S., Ewing, M. T., & Cooper, H. B. (2022). Artificial intelligence focus and firm performance. *Journal of the Academy of Marketing Science*, 50(6), 1176–1197. <https://doi.org/10.1007/s11747-022-00876-5>
- Möller, K., Schäffer, U., & Verbeeten, F. (2020). Digitalization in management accounting and control: An editorial. *Journal of Management Control*, 31(1), 1–8. <https://doi.org/10.1007/s00187-020-00300-5>
- Monod, E., Lissillour, R., Köster, A., & Qi, J. (2023). Does AI control or support?: power shifts after AI system implementation in customer relationship management. *Journal of Decision Systems*, 32(3), 542–565. <https://doi.org/10.1080/12460125.2022.2066051>
- Monteiro, A. P., Vale, J., Leite, E., Lis, M., & Kurowska-Pysz, J. (2022). The impact of information systems and non-financial information on company success. *International Journal of Accounting Information Systems*, 45, 100557. <https://doi.org/10.1016/j.accinf.2022.100557>
- Mosier, K. L., & Skitka, L. J. (2018). Human decision makers and automated decision aids: Made for each other? In M. Mouloua (Ed.), *Automation and human performance: Theory and applications* (pp. 201–220). Taylor & Francis Group.
- Myers, J., & Ramsey, H. (2023). Judgment work: The human component augmenting AI in strategic management. In J. Martinez-Falco, et al. (Eds.), *New perspectives and possibilities in strategic management in the 21st century: Between tradition and modernity* (pp. 155–167). IGI Global, Business Science Reference.
- Nakashima, H. H., Mantovani, D., & Machado Junior, C. (2024). Users' trust in black-box machine learning algorithms. *Revista De Gestão*, 31(2), 237–250. <https://doi.org/10.1108/REGE-06-2022-0100>
- Ning, X., Lu, Y., Li, W., & Gupta, S. (2024). How transparency affects algorithmic advice utilization: The mediating roles of trusting beliefs. *Decision Support Systems*, 183, 114273. <https://doi.org/10.1016/j.dss.2024.114273>
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education: theory and Practice*, 15(5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>
- Oll, J., Hahn, R., Reimsbach, D., & Kotzian, P. (2018). Tackling complexity in business and society research: The methodological and thematic potential of factorial surveys. *Business & Society*, 57(1), 26–59. <https://doi.org/10.1177/0007650316645337>
- Önkal, D., Gönül, M. S., & De Baets, S. (2019). Trusting forecasts. *Futures & Foresight Science*, 1(3–4), e19. <https://doi.org/10.1002/ffo2.19>
- Önkal, D., Goodwin, P., Thomson, M., Gönül, M. S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, 22(4), 390–409. <https://doi.org/10.1002/bdm.637>
- Oomen, T., Goncalves, J., & Mol, A. (2024). Rage against the artificial intelligence? Understanding contextuality of algorithm aversion and appreciation. *International Journal of Communication*, 18, 609–633.
- Paivio, A. (1979). *Imagery and verbal processes*. Lawrence Erlbaum Associates.
- Panda, S. K., & Mohanty, S. N. (2023). Time series forecasting and modeling of food demand supply chain based on regressors analysis. *IEEE Access*, 11, 42679–42700. <https://doi.org/10.1109/ACCESS.2023.3266275>
- Park, K., & Yoon, H. Y. (2024). Beyond the code: The impact of AI algorithm transparency signaling on user trust and relational satisfaction. *Public Relations Review*, 50(5), 102507. <https://doi.org/10.1016/j.pubrev.2024.102507>
- Patton, C. E., Clegg, B. A., Davis, B. C., Caglar, T., Siebert, C., & Blanchard, N. (2025). The choice to use automation: Improvements from evidence accumulation. *International Journal of Human-Computer Interaction*, 1–18. <https://doi.org/10.1080/10447318.2025.2452189>
- 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>
- Pierce, J. L., Kostova, T., & Dirks, K. T. (2001). Toward a theory of psychological ownership in organizations. *The Academy of Management Review*, 26(2), 298–310. <https://doi.org/10.5465/amr.2001.4378028>
- Piercey, M. D. (2023). “Throw it in as a covariate?” Common problems using measured control variables in experimental research. *Auditing: A Journal of Practice & Theory*, 42(2), 183–205. <https://doi.org/10.2308/AJPT-2020-011>
- Prahl, A., & van Swol, L. M. (2017). Towards an understanding of algorithm aversion: Why do decision-makers discount advice from automation. *Journal of Forecasting*, 36(6), 691–702. <https://doi.org/10.1002/for.2464>
- Promberger, M., & Baron, J. (2006). Do patients trust computers? *Journal of Behavioral Decision Making*, 19(5), 455–468. <https://doi.org/10.1002/bdm.542>
- Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation-augmentation paradox. *Academy of Management Review*, 46(1), 192–210. <https://doi.org/10.5465/2018.0072>
- Reitzenstein, B., Kösling-Guse, S., & Weißenberger, B. E., Arbeitskreis Digital Reporting der der Schmalenbach-Gesellschaft. (2023). Predictive forecasting. In *Digital reporting: Transformation des Controllerebereichs durch den digitalen Wandel* (pp. 41–47). Franz Vahlen.
- Rohrbeck, R., & Kum, M. E. (2018). Corporate foresight and its impact on firm performance: A longitudinal analysis. *Technological Forecasting and Social Change*, 129, 105–116. <https://doi.org/10.1016/j.techfore.2017.12.013>

- Rossmann, L., & Wald, A. (2024). The automation of management decisions: A systematic review and research agenda of the factors influencing the decision to increase the level of automation. *International Journal of Information Technology & Decision Making*, 23(1), 107–140. <https://doi.org/10.1142/S021962202330001X>
- Saragih, M., & Morrison, B. W. (2022). The effect of past algorithmic performance and decision significance on algorithmic advice acceptance. *International Journal of Human-Computer Interaction*, 38(13), 1228–1237. <https://doi.org/10.1080/10447318.2021.1990518>
- Schaap, G., Bosse, T., & Vettehen, P. H. (2024). The ABC of algorithmic aversion: Not agent, but benefits and control determine the acceptance of automated decision-making. *AI & SOCIETY*, 39(4), 1947–1960. <https://doi.org/10.1007/s00146-023-01649-6>
- Schecter, A., Bogert, E., & Lauharatanahirun, N. (2023). Algorithmic appreciation or aversion? The moderating effects of uncertainty on algorithmic decision making. Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, 1–8, Article30.
- Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. *Journal of Decision Systems*, 29(4), 260–278. <https://doi.org/10.1080/12460125.2020.1819094>
- Schultze, T., Pfeiffer, F., & Schulz-Hardt, S. (2012). Biased information processing in the escalation paradigm: Information search and information evaluation as potential mediators of escalating commitment. *The Journal of Applied Psychology*, 97(1), 16–32. <https://doi.org/10.1037/a0024739>
- Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., & Medow, M. A. (2013). Why do patients derogate physicians who use a computer-based diagnostic support system? *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 33(1), 108–118. <https://doi.org/10.1177/0272989X12453501>
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- Shrivastava, V. K., Shrivastava, A., Sharma, N., Mohanty, S. N., & Pattanaik, C. R. (2023). Deep learning model for temperature prediction: A case study in New Delhi. *Journal of Forecasting*, 42(6), 1445–1460. <https://doi.org/10.1002/for.2966>
- Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47–53.
- Tzioti, S. C., Wierenga, B., & van Osselaer, S. M. J. (2014). The effect of intuitive advice justification on advice taking. *Journal of Behavioral Decision Making*, 27(1), 66–77. <https://doi.org/10.1002/bdm.1790>
- Varghese Joseph, R., Mohanty, A., Tyagi, S., Mishra, S., Satapathy, S. K., & Mohanty, S. N. (2022). Hybrid deep learning framework with CNN and bi-directional LSTM for store item demand forecasting. *Computers and Electrical Engineering*, 103(2), 108358. <https://doi.org/10.1016/j.compeleceng.2022.108358>
- Venkatesh, V., Thong, J. Y. L., & Xu, X., University of Arkansas. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376. <https://doi.org/10.17705/1jais.00428>
- Verstraete, G., Aghezaf, E.-H., & Desmet, B. (2020). A leading macroeconomic indicators' based framework to automatically generate tactical sales forecasts. *Computers & Industrial Engineering*, 139, 106169. <https://doi.org/10.1016/j.cie.2019.106169>
- Vinson, N. G., Molyneaux, H., & Lapointe, J.-F. (2018). *A review of the role of explanations for user acceptance in black box systems* [Paper presentation]. International Conferences Interfaces and Human Computer Interaction 2018.
- Wanous, J. P., & Reichers, A. E. (1996). Estimating the reliability of a single-item measure. *Psychological Reports*, 78(2), 631–634. <https://doi.org/10.2466/pr0.1996.78.2>
- 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>
- Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, 93(1), 1–13. <https://doi.org/10.1016/j.obhdp.2003.08.002>
- Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation? *Organizational Behavior and Human Decision Processes*, 83(2), 260–281. <https://doi.org/10.1006/obhd.2000.2909>
- Zhang, C., Cho, S., & Vasarhelyi, M. (2022). Explainable artificial intelligence (XAI) in auditing. *International Journal of Accounting Information Systems*, 46, 100572. <https://doi.org/10.1016/j.accinf.2022.100572>

About the authors

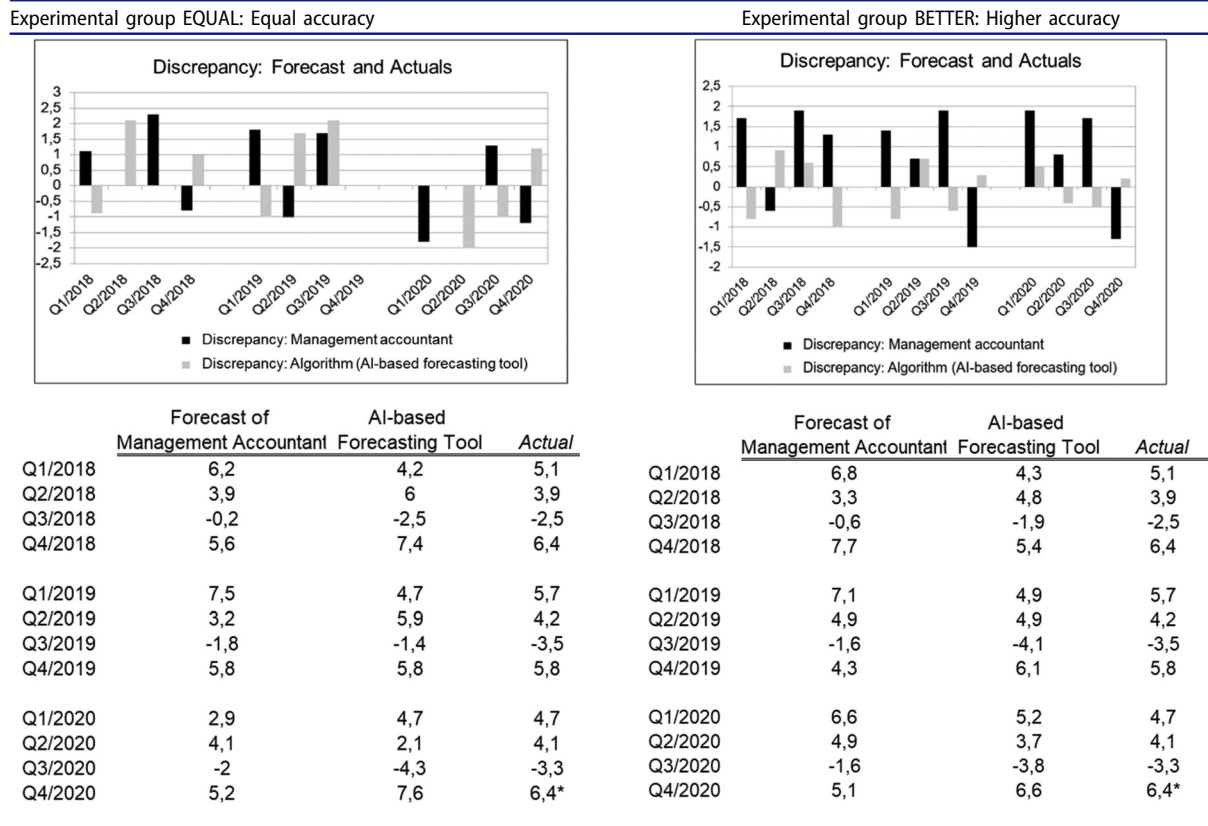
Peter Kotzian is research associate at the Chair of Management Control and Accounting at the Heinrich Heine University Düsseldorf. Social scientist by training, his research interests cover behavioral economics and the interplay between institutional settings and individual-level behavior in domains like decision support, investment behavior, and performance evaluation.

Barbara E. Weißenberger is Chair of Management Control and Accounting at Heinrich Heine University Düsseldorf. She has worked on numerous topic in both fields, including the integration of financial and managerial accounting and the use of management control systems with respect to digital transformation, sustainability and compliance.

Sonja G. Prinz had been working as research assistant at the Chair of Management Control and Accounting and is co-founder of the consulting agency Gradify. Her PhD thesis on “Management accounting in the age of digitalization” was published in 2024.

Appendix A. Experiment 1 – information on forecast accuracy

NOTE: Depicted here only for the ALGO condition, in the HUMAN condition the graphs were equivalent



Appendix B. Descriptives of the dependent variable in experiment 1

Dependent variable: Forecast of sales growth rate for the upcoming quarter

		Performance of second source is ...	
		... equal (EQUAL)	... better (BETTER)
Origin of second source is human (HUMAN)	N = 34 M = 4.70 SD = 1.08	N = 36 M = 4.88 SD = .73
	... algorithm (ALGO)	N = 34 M = 4.65 SD = 1.12	N = 33 M = 4.71 SD = 1.05

Note: N = number of participants, M = means, SD = standard deviation.

Appendix C. Experiment 2 – information on forecast accuracy and explanation

Forecasts of the sales growth rates of management accountants A and B (%) in the EQUAL condition (graphic representation were given analogous to experiment 1):

Quarter/Year	Forecast of		Actual
	Management Accountant A	Management Accountant B	
1/2018	4,3	4,3	5,4
2/2018	4,4	6,2	3,9
3/2018	-0,2	-3,3	-2,5
4/2018	5,6	6,9	6,4
1/2019	6,9	4,1	5,1
2/2019	3,9	6,7	4,9
3/2019	-1,8	-3,9	-3,5
4/2019	5,4	7,5	5,8
1/2020	7,1	4,1	5,3
2/2020	2,5	5,2	3,4
3/2020	-4,5	-4,2	-3,3
4/2020	7,6	7,6	6,4
1/2021	6,6	6,6	5,2

Forecasts of the sales growth rates of Management Accountants A and B (%) in the BETTER condition (graphic representation were given analogous to experiment 1):

Quarter/Year	Forecast of		Actual
	Management Accountant A	Management Accountant B	
1/2018	7,5	4,7	5,4
2/2018	2,3	4,8	3,9
3/2018	-1,3	-3,1	-2,5
4/2018	7,7	7,1	6,4
1/2019	6,5	4,3	5,1
2/2019	6,2	5,4	4,9
3/2019	-5,4	-3,9	-3,5
4/2019	4,3	6,1	5,8
1/2020	7,2	5,6	5,3
2/2020	2,2	3,0	3,4
3/2020	-1,6	-3,6	-3,3
4/2020	5,1	6,6	6,4
1/2021	6,6	5,4	5,2

Explanation in the ALGO condition

The artificial intelligence (AI)-based forecasting tool uses a large amount of current and past data related to the company, as well as data from social media. It determines patterns in this data and uses them to forecast sales development.

For the second quarter of 2021 forecast, the AI tool gives a seasonal trend pattern as the reason for its forecast. This pattern shows that sales development decreases slightly from the first to the second quarter, that the third quarter is typically very weak, and that the fourth quarter is again very strong. In addition to this seasonal pattern, the AI tool determined a base turnover and adjusted it according to the found seasonal pattern.

Explanation in the HUMAN condition

Management accountant B uses a large amount of current and past data related to the company, as well as data from social media. He determines patterns in this data and uses them to forecast sales development.

For the second quarter of 2021 forecast, the management accountant B gives a seasonal trend pattern as the reason for his forecast. This pattern shows that sales development decreases slightly from the first to the second quarter, that the third quarter is typically very weak, and that the fourth quarter is again very strong. In addition to this seasonal pattern, management accountant B determined a base turnover and adjusted it according to the found seasonal pattern.

Appendix D. Descriptives of the dependent variable in experiment 2

Dependent variable: Forecast of sales growth rate for the upcoming quarter

		Performance of second source is ...	
		... equal (EQUAL)	... better (BETTER)
No explanation	Origin of second source is human (HUMAN)	N = 25 M = 4.03 SD = .48	N = 36 M = 4.88 SD = .73
	Origin of second source is algorithm (ALGO)	N = 25 M = 3.99 SD = .67	N = 33 M = 4.71 SD = 1.05
With explanation	Origin of second source is human (HUMAN)	N = 25 M = 3.92 SD = 1.47	N = 25 M = 4.39 SD = .63
	Origin of second source is algorithm (ALGO)	N = 26 M = 3.79 SD = 1.55	N = 24 M = 4.41 SD = 1.01

Note: *N* = number of participants, *M* = means, *SD* = standard deviation.