Jung, Markus; Seiter, Mischa

Towards a better understanding on mitigating algorithm aversion in forecasting: an experimental study

Journal of Management Control

Provided in Cooperation with:
Springer Nature

Suggested Citation: Jung, Markus; Seiter, Mischa (2021) : Towards a better understanding on mitigating algorithm aversion in forecasting: an experimental study, Journal of Management Control, ISSN 2191-477X, Springer, Berlin, Heidelberg, Vol. 32, Iss. 4, pp. 495-516, https://doi.org/10.1007/s00187-021-00326-3

This Version is available at:
https://hdl.handle.net/10419/286788

Journal of Management Control (2021) 32:495–516
https://doi.org/10.1007/s00187-021-00326-3

ORIGINAL PAPER

Towards a better understanding on mitigating algorithm aversion in forecasting: an experimental study

Markus Jung[1] · **Mischa Seiter**[1]

Accepted: 7 September 2021 / Published online: 28 September 2021

Abstract

Forecasts serve as the basis for a wide range of managerial decisions. With the potential of new data sources and new techniques for data analysis, human forecasters are increasingly interacting with algorithms. Although algorithms can show better forecasting performance than humans, forecasters do not always accept these algorithms and instead show aversion to them. Algorithm aversion has become a widely known phenomenon. Drawing on the seminal study of Dietvorst et al. (J Exp Psychol Gen 144(1):114–126, 2015), we extend the evidence on algorithm aversion by introducing three environmental variables from the management accounting literature. We argue that time pressure, "do your best" goals, and forecasters' data input decision rights on the algorithms input mitigate algorithm aversion. To test our hypotheses, we conducted an experimental study with 1,840 participants overall. We found support for our hypothesis that time pressure mitigates algorithm aversion. We found evidence that the mitigation effect is based on forecasters' loss of confidence in their own forecast when they are under time pressure. We found no support for our hypothesis on "do your best" goals or forecasters' data input decision rights.

Keywords Algorithm aversion · Forecasting · Time pressure · Experimental study

1 Introduction

Forecasts are the basis for a wide range of managerial decisions (Butler & Ghosh, 2015; Chen et al., 2015a). We have seen several approaches to using new forms of data as well as new techniques to analyse them (Teoh, 2018). A prominent example is the use of customer data to implement smart, connected products (Porter & Heppelmann, 2014). In the field of management accounting, new forecasting approaches

✉ Markus Jung
markus.jung@uni-ulm.de

Mischa Seiter
mischa.seiter@uni-ulm.de

[1] Ulm University, Helmholtzstraße 22, Ulm, Germany

are a manifestation of this trend, and human forecasters are increasingly forced to interact with complex algorithms (Dietvorst et al., 2018).

With regard to forecasting, algorithms have been shown to be more accurate than human forecasters in many fields (Grove et al., 2000). This holds true even for simple forms of algorithms like linear models (Dawes, 1979). Due to algorithms' superior performance in terms of forecasting accuracy, one would expect to see human forecasters integrate algorithms into their daily work (Dietvorst et al., 2015; Logg et al., 2019; Venkatesh & Davis, 2000). However, the literature shows a variety of cases in which forecasters reject superior algorithms at the expense of their work results (Castelo et al., 2019; Dietvorst et al., 2015; Grove et al., 2000; Prahl & van Swol, 2017). This phenomenon has a long history (Dawes, 1979; Meehl, 1954), and reasons for rejecting superior algorithms range from forecasters' overconfidence (Logg et al., 2019) to irrational overreaction to new information (Remus et al., 1995).

A seminal study by Dietvorst et al. (2015) defines algorithm aversion as follows: human forecasters refuse to use algorithms that perform far better than themselves after they see an algorithm err. This effect is of high relevance because perfect accuracy is impossible in forecasting (Dietvorst et al., 2015). However, previous research also shows that forecasters are willing to use algorithms with which they have not previously interacted to improve their forecasts. Logg et al. (2019) refer to this willingness as algorithm appreciation. In contrast, algorithm aversion seems to be affected by different aspects of forecasters' work environment or task specificities. For example, it is mitigated when forecasters can modify the results of an algorithm, and therefore worsening forecasting accuracy (Dietvorst et al., 2018), or if the task is perceived as objective rather than subjective (Castelo et al., 2019).

Although research on algorithms in forecasting and algorithm aversion has been ongoing for more than 60 years, there are still large gaps in the literature (Burton et al., 2019). The growing availability of algorithms increases the need for a better understanding of how algorithm aversion occurs and how working environments affect it (Castelo et al., 2019). Prior studies have largely neglected the working environment to which forecasters are subjected when they create forecasts and suffer from algorithm aversion. Following Shaw and Gupta (2015) and Merchant and van der Stede (2017), future work must address the question of under what conditions forecasters direct their actions toward the best performance, thus achieving the best forecasting accuracy. Logg et al. (2019) summed up this discussion by noting that "algorithm aversion is not as straightforward as prior literature suggests, nor as contemporary researchers predict." This study is therefore motivated by the question: does the working environment of forecasters mitigate algorithm aversion?

We study the environment of forecasters and their decision-making process when interacting with algorithms. More specifically, we draw on the work of Bonner and Sprinkle (2002) to define the most relevant environmental variables for the forecasting task of Dietvorst et al. (2015). This replicates and extends the work of Dietvorst et al. (2015) with three emerging environmental variables from forecasters' daily work routines. First, we induce time pressure to investigate when forecasters only have a limited amount of time for their forecasts. Second, we assign a "do your best" goal regarding the best forecasting accuracy. Third, we give our participants

data input decision rights and therefore decision rights to choose the information input for the algorithm. To integrate these management accounting variables into the work of Dietvorst et al. (2015), we conducted an experimental study on Amazon's Mechanical Turk with 1840 participants.

Our study contributes to the management accounting and forecasting literature in four ways: first, we contribute to the literature on algorithm aversion by demonstrating that time pressure mitigates algorithm aversion—without the need of forecasters to affect the algorithms' outcome and therefore worsening the algorithms forecasting accuracy. In doing so, we also add important knowledge to the literature on time pressure by showing, second, that under time pressure forecasters willingly use algorithms as a meaningful way to escape the uncertainty and tense situation they experience under time pressure. Third, conversely, a "do your best" goal is ineffective in mitigating algorithm aversion. Fourth, giving forecasters more decision rights by letting them choose the data that the algorithm processes does not impact their algorithm aversion.

2 Theoretical background on algorithm aversion in forecasting

An algorithm is essentially a sequence of mathematical calculations with a specific goal, defined as "a procedure for computing a function" (Rogers, 1987). Algorithms can capture and evaluate large, unstructured sets of data, often referred to as "big data" (Logg et al., 2019; Vitale et al., 2020). Research on algorithms in decision-making dates back to the work of Meehl (1954) and Dawes (1979). Meehl (1954) was the first to describe the psychological aversion to algorithms, emphasizing the fundamental superiority of statistical models over expert opinions and how experts still assess these superior performing models with irrational scepticism. In a similar fashion, Dawes (1979) shows how simple linear models can predict students' success better than experts can. Highhouse (2008) argues that the phenomenon of a forecaster rejecting an algorithm may be rooted in the belief that human forecasters tend to be able to act perfectly, whereas algorithms cannot do so. Grove et al. (2000) verified this phenomenon in a meta-analysis of 136 studies; however, the authors also found algorithms to display an average superiority of 10% over expert opinions.

Dietvorst et al. (2015) proposed the term *algorithm aversion* to describe this phenomenon in the presence of performance feedback in forecasting. Performance refers to the accuracy of a forecast, while feedback refers to receiving insights on previous performance—in this case, the performance of forecasts created either by humans or algorithms. The findings by Dietvorst et al. (2015) serve as a framework for research on algorithm aversion in our study. In several incentivized forecasting experiments, they gave participants the opportunity to choose either their own assessment or that of an algorithm. Most participants initially opted for the algorithm and subsequently were shown the results of the algorithm. The transparency of the algorithm's forecast made the participants aware that the algorithm did not produce perfect results. As soon as the participants became aware of the algorithm's potential for erroneous results, the rate of rejection increased. This rejection persisted after participants were told that the algorithm,

despite not being perfect, performed better than humans on average. This finding is not in line with agency theory, because agency theory assumes that individuals maximize their benefits and thus choose the forecast that generates the highest revenues in the long run (Baiman, 1990; Eisenhardt, 1989). However, this rationale seems to be no longer given in the context of algorithm aversion. It is important to note that Dietvorst et al. (2015) do not postulate a general algorithm aversion in the sense that people will consistently reject any algorithm. Rather, algorithm aversion refers to the increasing rejection of an algorithm once it has been perceived as capable of producing erroneous results.

Performance feedback has often been studied in the literature on forecasting and decision-making, typically focusing on the search for causes and possibilities of the effect of feedback on performance (Ashton, 1990; Chen et al., 2015). Research has shown that performance feedback has different effects on forecasters. Depending on the characteristics of forecasters, it can improve the performance of forecasters (Ashton, 1990; Lourenco et al., 2018), or worsen it (Akın & Karagözoğlu, 2017), or even not affect it at all when forecasters are self-confident. In the context of algorithm aversion, there is a clear picture of how performance feedback affects decision-making: the effect of feedback is not examined for the accuracy of the forecaster, but rather on the likelihood that the forecaster will trust an algorithm or themselves. It is postulated that forecasters prefer their own forecasts if they receive performance feedback on both their own and the superior performance of an algorithm. Therefore, performance feedback on the algorithm's forecasting accuracy is the cause of algorithm aversion (Dietvorst et al., 2015). If there is no performance indication of the algorithm's forecasting accuracy, algorithm aversion does not occur. This creates a new dimension of the impact of feedback, making existing approaches less explanatory. However, there is no evidence regarding whether this effect of performance feedback in algorithm aversion might be affected by other variables.

A further explanation of algorithm aversion considers confidence in the algorithm as a mediating effect. Algorithm aversion can be seen as a loss of confidence in an algorithm when its poor performance is perceived by a forecaster. Confidence describes expectations of the forecaster regarding the accuracy of a forecast. The initial expectations a forecaster regarding an algorithm might be driven by past experiences and are therefore unstable over time. Interaction with algorithms over many years in everyday and professional life and the continuous improvement of algorithm performance might lead to increased positive experiences with the outcome of algorithms and therefore higher confidence in unknown algorithms (Al-Htaybat & Alberti-Alhtaybat, 2017; Appelbaum et al., 2017; Quattrone, 2016). This could result in two developments concerning algorithm aversion. First, without feedback regarding the specific algorithm's performance, the use of algorithms in forecasts will increase because there is less evidence leading forecasters to set low expectations and reject unknown algorithms. Second, if the forecaster receives feedback about the imperfect performance of the algorithm, they gain uncertainty about the performance of that algorithm. This means that with a higher starting level of confidence, a feasible loss of confidence could arise. This could increase algorithm aversion over time. In addition, it must be noted that algorithm aversion is only harmful for organizations if it leads to worse forecasting accuracy. Thus, theory states

that even algorithms with a better accuracy than the forecasters are rejected by the forecaster.

Organizations cannot prevent repeated collaboration of forecasters with algorithms and thus forecasters receiving performance feedback in the future. It is therefore important to investigate how algorithm aversion can be lowered. The following hypothesis development aims to discuss influencing variables that we expect to mitigate algorithm aversion when performance feedback is present.

3 Hypothesis development on mitigating algorithm aversion

Dietvorst et al. (2018) were able to show that algorithm aversion is mitigated by the possibility of forecaster intervention in the results of the algorithm, although such interventions generally worsened the forecasts (see also Carbone et al., 1983 or Goodwin & Fildes, 1999). Remarkably, it does not matter whether the forecaster can change the result of the algorithm strongly or only marginally; even small adjustments significantly mitigate algorithm aversion. Due to the effect of worsening forecast accuracy, this might not always be a practical solution.

Logg et al. (2019) follow Dietvorst et al., (2015, 2018) in conducting a series of experiments to answer the following research question: In what cases do forecasters trust the advice of algorithms more than that of experts? Logg et al. (2019) coined the term *algorithm appreciation* as a parallel to the term *algorithm aversion* to describe people's acceptance of algorithms before receiving performance feedback. They show that algorithm appreciation is boosted when the role of the self is involved. In other words, people disregard advice in general and show a preference for their own judgment. Due to potential large errors or unintended functions of algorithms, humans need a critical scepticism toward algorithms with which they do not have experience (McKinney et al., 2017). Castelo et al. (2019) show that algorithm aversion is especially likely to occur when tasks have a subjective character. Consequently, algorithm aversion can be mitigated by increasing the perceived objectivity of a task. The research thus shows that other factors influence algorithm aversion.

Forecasters are "social beings with complex and somewhat changeable motivations, not as isolated operators of stable (probably profit-maximizing) decision models" (Luft, 2016, p. 9). As a result, their decision-making processes are subject to several environmental influencing factors. Besides the opportunity to change the algorithm results or the differentiation of more subjective or more objective tasks, there is a large gap regarding forecasters' daily work environment.

First, during the preparation of forecasts, deadlines play a major role. Data can arrive on short notice and must be processed quickly, which usually leads to time pressure in the preparation of forecasts. Second, realistic payment is usually a fixed remuneration; forecasters' salaries are rarely tied to the accuracy of a forecast, which is why they tend to simply "do their best" for a forecast without receiving a reduced remuneration in case of poor performance. Third, forecasts are a recurring task, and forecasters know the basis of their data. They can therefore judge to what extent they are suitable for the preparation of a forecast and decide for themselves which data

their forecasting systems should work with. Linking these circumstances to accounting relevant variables, we draw on the framework of Bonner and Sprinkle (2002), who point out environmental variables that have been widely investigated in the management accounting literature: time pressure, assigning a "do your best" goal, and assigning decision rights for the data input (Bonner & Sprinkle, 2002).

3.1 The impact of time pressure on algorithm aversion

The availability of data is continuously increasing in terms of both scope and speed. In a fast-moving and volatile environment, there is less time for organizations and their respective employees to turn vast amounts of data into accurate forecasts (Camerer et al., 2004; Spiliopoulos & Ortmann, 2018). In cases where data arrive very soon before the forecast is due, the processing of the information for a forecast can become time critical. While time pressure can result from the forecasting task as such, a reduction in the time available for a task can also be induced by supervisors. For example, deadlines can be shortened, or data can be passed on later. The importance and effects of time pressure have been examined widely in the literature (Kelly et al., 2011; Lambert et al., 2017; Pietsch & Messier, 2017; Spilker, 1995; Wegier & Spaniol, 2015).

In the existing literature on algorithm aversion, forecasters do not have any time constraints for their forecasts. When time pressure occurs or is induced, there is less time for the forecaster to think of different options and possibilities that might lead to a change in behaviour (Wegier & Spaniol, 2015). Behavioural research therefore has shown that time pressure causes people to feel anxious, lose confidence in their own judgment, and decrease their effort towards a task (Pietsch & Messier, 2017). These findings appear in the literature on choking under pressure, which finds that people's performance worsens when they lose confidence in their own skills (Beilock & Carr, 2001) and that people even may quit when performance expectations are very high and they no longer believe in themselves (Dai et al., 2018). These negative effects on forecasters mainly affect their self-confidence. Time pressure can also improve the change towards new ways and strategies of problem-solving (Mather & Lighthall, 2012).

In summary, the literature suggests that introducing time pressure will lower forecasters' confidence in their own forecasts. Time pressure should not affect the way forecasters evaluate the performance of the algorithm. By decreasing confidence only in their own forecasts, we expect time pressure to mitigate algorithm aversion. We therefore hypothesize:

H1 Time pressure on the human forecaster mitigates algorithm aversion.

3.2 The impact of a "do your best" goal on algorithm aversion

In the literature on algorithm aversion, a payment scheme that encourages forecasters to achieve the best possible, if not perfect, forecast is typically applied. This means that a specific and difficult goal is assigned by an incentive structure

that allows for few errors according to forecasting accuracy. In the real working environment of forecasters, there are usually no employment contracts that link payment to the accuracy of a forecast. Forecasters receive a fixed salary independent of their individual performance. The targets for a forecast are therefore formulated independently of monetary incentive and can therefore be assigned in the sense of a "do your best" goal. We therefore discuss in the following how different goal settings can affect algorithm aversion and why we expect a more realistic setting with a “do your best” goal to mitigate algorithm aversion.

One of the most widely proven effects in psychology and management science is that specific and difficult goals lead to better performance than vague “do your best” goals (Locke & Latham, 1990, 2013). This effect has been shown convincingly in over 90% of laboratory and field studies (Locke et al., 1981). In some cases, research has shown that specific goals do not lead to better performance (Chen et al., 2015a) or do not affect performance (Akın & Karagözoğlu, 2017). Ordóñez et al. (2009) note that while goals focus attention, “[u]nfortunately, goals can focus attention so narrowly that people overlook other important features of a task”. They present several potential reasons why specific goals have so far led to algorithm aversion and why a “do your best” goal could reduce it. These include the fact that people limit their focus by setting specific goals, thus preventing adaptation, working methods, or learning effects (see also Earley et al., 1989; Wood et al., 1990). This might focus forecasters narrowly on their own performance instead of carefully comparing it with the performance of the algorithm. In addition, Webb et al. (2013) add that when goals are more challenging, they hinder people from using new strategies to be more efficient. Instead, people start working harder using their existing methods. This means that difficult goals can increase performance but can also hinder the discovery of new, more efficient working routines. Polzer and Neale (1995) show that goals hinder people from thinking about task dimensions that are not directly affected by the goal.

Following van Dyck et al. (2005), difficult goals hinder people from truly processing new information within the task. Latham and Locke (2013) also note that the fear of making mistakes increases because of difficult goals. They suggest that mistakes must be accepted and embraced so that people can properly judge the mistake and learn from it. Such a work environment where errors are accepted can be established with a “do you best” goal. Setting a “do your best” goal by telling people they receive their incentive no matter their performance is supposed to be positively related to performance and goal achievement (Gold et al., 2014; Seckler et al., 2017; van Dyck et al., 2005). Arguing with the impact of a “do your best” goal, we conclude that difficult goals hinder people from truly processing the superior yet not perfect performance of the algorithm because of forecasters’ narrow focus on the algorithm’s mistakes. Therefore, to mitigate algorithmic aversion, a “do your best” goal is applied.

H2 A “do your best” goal regarding forecasting accuracy mitigates algorithm aversion.

3.3 The impact of data input decision rights on algorithm aversion

Fildes et al. (2009) show that 90% of the results of algorithms in forecasts are not directly accepted but rather are changed by the forecaster. This shows a desire of forecasters for decision rights when interacting with algorithms. Dietvorst et al. (2018) show that algorithm aversion can be mitigated consistently by letting people modify imperfect forecasts produced by an algorithm. However, modifying the forecast of an algorithm has an unintended effect: adjusting the result of the algorithm causes the accuracy of the forecast to decrease (Carbone et al., 1983; Dietvorst et al., 2018; Goodwin & Fildes, 1999; Remus et al., 1995). Fildes et al. (2009) note that small changes in forecasts impair their accuracy, while bigger changes often lead to improvement. Particularly in cases where forecasters think they have contemporary negative evidence from certain information in certain areas that the algorithm does not have in its historical data, it might be useful to forecasters to delete this information from the algorithm's information processing (Remus et al., 1995). This implies there is some need for forecasters in selecting the data when forecasters interact with algorithms. It is important to understand how forecasters handle and value information for their own forecast and that of the algorithm.

The literature on algorithm aversion often describes a predefined algorithm that is built up on a particular database. Forecasters also must work with that database. By interacting with the database and the algorithm, forecasters get a feel for the difficulty of the task and how to derive a forecast from the data. It is important to understand that not all data has the same relevance to the quality of a forecast. The forecaster thus learns over time which data are particularly important for a good forecast.

When gaining experience with the data that underlie an algorithm's forecast, a human forecaster might think the algorithm is weighting the information incorrectly. This implies that there should be a way to mitigate algorithm aversion by allowing the forecaster to select the information the algorithm processes and therefore giving them data input decision rights. We expect the ability to do so to raise confidence in the algorithm, especially when forecasters have experience with the data and might believe that they know where possible algorithm miscalculations are rooted. We therefore hypothesize:

H3 Data input decision rights for the human forecaster mitigate algorithm aversion.

4 Research design

We conducted an experimental study following the design and procedure of Dietvorst et al. (2015).[1] We collected a similar sample size and used the same data sources. However, we extended the study by introducing three additional variables, which we expect to find to mitigate algorithm aversion. The following chapters

[1] We replicate Study 3b, which is the largest study in Dietvorst et al. (2015) with regard to participants.

Table 1 Comparing sample to Dietvorst et al. (2015)

	Dietvorst et al. (2015)[a]	This study
Average age	33–34	39
Sex (% female)	46–53	46
Exclusion rate (%)	38	39
Final sample (average per condition)	259	230

[a]Dietvorst et al. (2015) reported these measures across several experimental groups and therefore presented several ranges

Table 2 Number of participants per condition

	Replication	Time pressure	"Do your best"	Data input decision rights
Algorithm appreciation	n = 219	n = 263	n = 208	n = 209
Algorithm aversion	n = 219	n = 223	n = 237	n = 262

illustrate how we conducted our study, the measures we employed, and the structure of each experimental condition.

4.1 Participants

Selection and payment of our participants followed Dietvorst et al. (2015). We conducted our study on Amazon's Mechanical Turk. Overall, we had 3032 participants and applied an exclusion rate of 38–39%, similar to Dietvorst (see Table 1). We excluded 865 participants who did not answer the main dependent variable of choosing the algorithm or their own forecast for their incentivized forecast. Another 47 participants were excluded for failing the attention check, and 280 participants failed to answer the belief questions.[2] This left us with 1840 participants. The sample averages $M = 39$ years of age and is 46% female. Table 1 offers an overview to compare the samples of our study and those in the study by Dietvorst et al. (2015).

4.2 Experimental design, manipulation, and measurement

In our setting, the participants were randomly assigned to one of eight conditions (n = 1840) (Table 2), which can be separated into two main clusters. One cluster did not gain any experience with the algorithm's forecasting performance and their own forecasting performance (algorithm appreciation conditions), and the other gained experience with the algorithm's forecasting performance and their own forecasting

[2] We looked the algorithm's worst forecast, which was 17 off, and added two standard deviations of the algorithm's performance so as not to exclude pessimistic forecasters. Participants who saw the algorithm as "on average more than 27 off" were not considered further in the study.

performance (algorithm aversion conditions) before making an incentivized forecast. In two experimental conditions no further adjustments were made. These conditions therefore replicate Dietvorst et al. (2015), and we refer to them as "replication". By comparing participants reliance on algorithms between these clusters, algorithm aversion can be shown.

We applied the following payment rule to determine the bonus payment for each participant: every participant received $1 for finishing the task. Participants received an additional $1 for a perfect forecast. For each unit of error, this amount was reduced by $0.15. The unit of error in each forecast is the measure for forecasting accuracy. These measures are important to compare our findings with prior findings on algorithm aversion, forecasting accuracy, and bonus payments.

For Hypotheses H1-H3, we investigate in those participants that receive performance feedback and therefore suffer from algorithm aversion. To do so, we set the algorithm aversion replication condition as our control condition and compare it to the other three algorithm aversion conditions.

Participants in the time pressure condition were given 12 s to make a decision. We pre-tested not only whether the given time would be short enough to trigger time pressure, but also whether the task could still be fulfilled with sufficient care under this pressure. The payment rule was the same as the replication conditions. We refer to this condition as the "time pressure" condition.

The participants in the condition with a "do your best" goal were told that they would receive a fixed incentive of $2, regardless of the accuracy of their incentivized forecasting performance. They were told to "do their best" to achieve the best accuracy in the upcoming forecasting task. We refer to this condition as the "do your best" condition.

The participants in the condition with data input decision rights for input information for the algorithm to process were told that the algorithm could only process 4 of the 5 variables shown, and that they could choose which ones to use. The payment rule was the same as for the replication conditions. We refer to this condition as "data input decision rights". Table 2 shows the final number of participants we gathered for each condition.

4.3 Experimental task and procedures

We started the study with a screening question to make sure participants were paying attention. After the screening question, participants were asked to solve a forecasting task. Specifically, they were asked to estimate "the rank of 1 U.S. state in terms of the number of airline passengers who departed from that state". They were given a list of five different pieces of information about the state. They were then told that they were going to receive a prediction from an algorithm that was "developed by transportation analysts".

The algorithm aversion conditions then started with 10 practice forecasts (stage 1 forecasts) to get accustomed to the task and the data. After each forecast, they received information about their forecasting accuracy, the algorithm's accuracy, and the state's true rank. The algorithm was the same as in Dietvorst et al. (2015) and

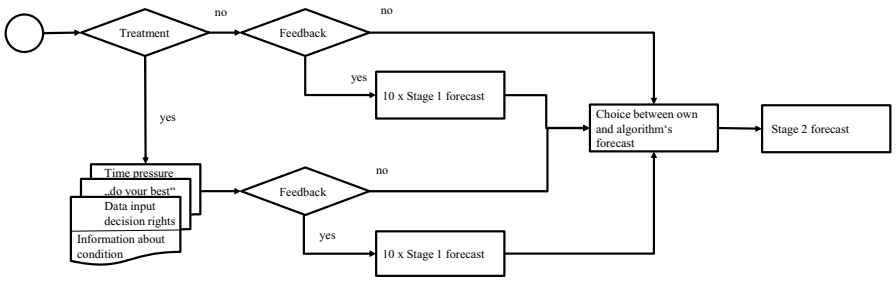

Fig. 1 Structure of experimental procedure

Table 3 Confidence and belief measures

Number	Question	Scale
1	On average, how many ranks do you think your estimates are away from states' actual ranks?	0–50
2	On average, how many ranks do you think the model's estimates are away from states' actual ranks?	0–50
3	How much confidence do you have in the statistical model's estimates?	1 = none; 5 = a lot
4	How much confidence do you have in your estimates?	1 = none; 5 = a lot

could perform better or worse than the participants. The algorithm appreciation conditions proceeded directly to their incentivized forecast (stage 2 forecast) without gaining experience with the data or the algorithm. Before the incentivized forecast, each participant had to decide whether to tie their incentive to their own forecast or to the algorithm's forecast.

Since we conducted the study on the same platform as the original, we added a question asking respondents to indicate whether they had previously participated in a similar task. Following Logg et al. (2019), we excluded these participants. We also excluded some explorative questions that were asked chronologically after the main dependent variable question was asked.

Each participant received information about their specific condition. All eight conditions were told that not all of the information shown to them would be equally relevant to the forecasting result. To avoid the data input decision rights condition, regretting their first selection and therefore rejecting the algorithm, they could adjust their decision (4 out of 5 input variables) before the incentivized stage 2 forecast. During stage 1 forecasts, the actual algorithm was not changed; thus, the participants always received the same feedback on the performance of the algorithm as the other conditions. Figure 1 provides a brief description of the experimental procedure.

Each participant also had to answer questions regarding confidence and belief measures from Dietvorst et al. (2015) (see Table 3). Decreasing confidence in the algorithm due to performance feedback operates as a mediator for algorithm aversion in the literature. Belief measures were conducted to closely replicate the study by Dietvorst et al. (2015), and we used them to test whether participants understood

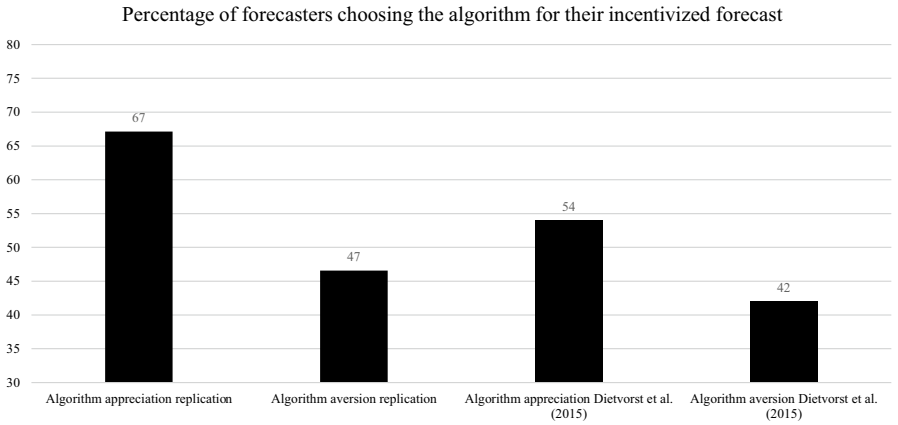

Fig. 2 Descriptive statistics on algorithm aversion compared to Dietvorst et al. (2015)

the task.[3] Belief questions 1 and 2, as well as confidence questions 3 and 4, were asked in a randomized order. We also asked participants for their age, gender, and highest level of education.

5 Results

In the following, we first present the analyses for algorithm aversion in our replication conditions. For a reliable evaluation of the results, we compare our results with those of Dietvorst et al. (2015). Second, we present the hypothesis tests on the mitigating effects of time pressure, "do your best" goals, and forecasters' data input decision rights on the algorithms input on algorithm aversion.

5.1 Comparing results to Dietvorst et al. (2015)

In the algorithm appreciation replication condition, 67% of participants chose the algorithm to determine their incentivized forecast. Feedback on their own performance and that of the algorithm reduces this figure to 47% in the algorithm aversion replication condition (see Fig. 2). This means that the algorithm aversion in our replications conditions is significant: χ^2 (1, N = 438) = 18,85, p = < 0.001.

Comparing our findings with those of Dietvorst et al. (2015), algorithm aversion seems to have increased since 2014. The percentage of participants who chose to tie their incentive to the algorithm without feedback increased from 54% in Dietvorst et al. (2015) to 67% in the present study. The isolated consideration of feedback

[3] We also collected the measures "How likely is the model to make a really bad estimate?" and "Do you think that your estimate or the model's estimate is closer to the last state's true rank?" We did not use these measures further.

Table 4 Forecasting performance: means (standard deviation)

	Algorithm	Forecaster	Difference	Paired t-test
Forecasting accuracy				
Replication conditions	4.38 (4.35)	10.11 (9.21)	5.73 (9.49)	$t(437) = 12.63$, $p < 0.001$
Dietvorst et al. (2015)	4.38 (4.34)	9.59 (9,37)	5.21 (9.66)	$t(520) = 12.31$, $p < 0.001$

Control- and Algorithm-Forecaster Condition from Dietvorst et al. (2015). Since in the published paper the performance was given across all 4 groups, but 3 out of 4 groups received feedback, these results have a bias towards good performance. We recalculated with the original dataset with the control and model and human condition

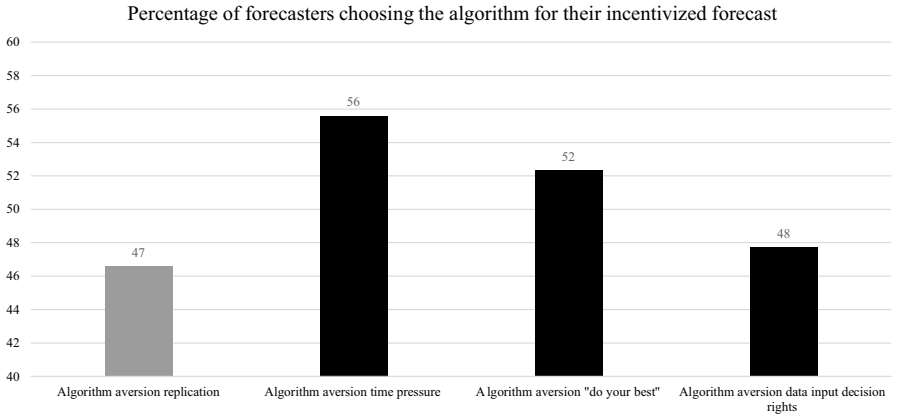

Fig. 3 Descriptive statistics on mitigating algorithm aversion

conditions also shows an increase in the general acceptance of algorithms from 42 to 47%. The delta representing algorithm aversion has increased from 12 to 19% from Dietvorst et al. (2015) in our study.

The rejection of the algorithm is only harmful to forecasters—and, in the long term, to organizations—when the algorithm performs better than the forecaster. We therefore provide further evidence of the superior performance of the utilized algorithm (see Table 4). This lack of accuracy shown in Table 4 leads to low bonuses for our participants. Forecasters who relied on their own forecast earned a \$0.27 bonus, while those who relied on the algorithm earned \$0.49. This difference of \$0.22 is significant ($t(437) = -10.53$, $p < 0.001$).

5.2 Analysis for mitigated algorithm aversion

In order to investigate whether algorithm aversion is mitigated by the three new variables, we first prove algorithm aversion within each condition. We found that it is significant within all conditions (comparing each algorithm appreciation with each algorithm aversion condition within the three new variables. Time pressure: χ^2 (1,

Table 5 Hypothesis test for mitigated algorithm aversion

Hypothesis	χ^2	Df	p-value	N
H1	3.606	1	0.029	442
H2	1.503	1	0.110	456
H3	0.062	1	0.402	481

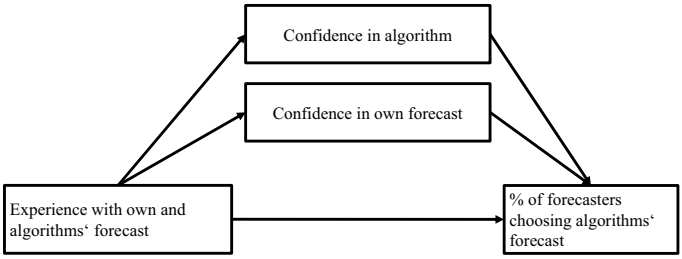

Fig. 4 Potential role of confidence on algorithm aversion

N = 486) = 9566, p = 0.002, "do your best" χ^2 (1, N = 445) = 24,422, p < 0.001, data input decision rights: χ^2 (1, N = 471) = 20,335, p < 0.001.).

Building on algorithm aversion within each condition, we further analysed algorithm aversion across conditions by comparing each treatment condition to the algorithm aversion replication condition (see Fig. 3). Participants' preference for the algorithm in the algorithm aversion conditions is as follows: in the replication condition, 47% of the participants chose the algorithm. The data input decision rights condition behaves very similarly to the control condition. While in the "do your best" condition more than 50% of the participants chose the algorithm, one can see the significant influence of time pressure in the almost 56% preference for the algorithm.

We tested Hypotheses H1–H3 using a one-sided chi-square test between the algorithm aversion conditions in each case, with the algorithm aversion replication condition as the reference (see Table 5).

Table 5 shows significant support for H1. That is, time pressure mitigates algorithm aversion. Meanwhile, we find no support for H2 or H3.

6 Additional analysis: the role of forecasters' confidence in their own forecast

The effect of performance feedback in the context of algorithm aversion is described by Dietvorst et al. (2015) as the influence on the confidence of the forecaster in the algorithm and self-confidence (see Fig. 4). The cause for algorithm aversion is described as the relationship between performance feedback as the independent variable, confidence in the algorithm as a mediator, and the choice between one's own and the algorithm's forecast as the dependent variable. Confidence in one's own

Table 6 Comparing confidence measures to the algorithm aversion replication condition

	Algorithm aversion conditions	Unpaired t-test
Confidence in algorithm's forecast	Difference (standard deviation)	
Time pressure	−0.080 (0.073)	t(440) = −1.094, p = 0.275
"Do your best" goal	0.044 (0.081)	t(454) = 0.629, p = 0.529
Data input decision rights	−0.013 (0.068)	t(479) = −0.191, p = 0.849
Confidence in own forecast	Difference (standard deviation)	
Time pressure	−0.184 (0.085)	t(440) = −2.151, p = 0.032
"Do your best" goal	0.102 (0.081)	t(454) = 1.257, p = 0.209
Data input decision rights	0.067 (0.082)	t(479) = 0.815, p = 0.416

forecast has also been tested as a mediator, but did not have any significant effects (Dietvorst et al., 2015).

A forecaster who receives performance feedback on their accuracy and on that of an algorithm will lose confidence in the algorithm but not in their own performance, even if the forecaster's performance is worse. Confidence in the algorithm acts as a mediator between the experience with the algorithm and the forecaster's choice. Confidence in one's own forecast, however, has no influence on the forecaster's decision-making process.

Due to this effect of the confidence measure, we now broaden the view to include the analysis of the influence of confidence in the algorithm's forecast and confidence in one's own forecast. It is of particular interest why we could only cause a mitigation in algorithm aversion through time pressure. A comparison of the confidence measures reveals a different picture. No significant changes in the confidence in the algorithm can be found. A significant change in the confidence in one's own forecast under time pressure can be observed. Thus, time pressure makes forecasters feel increasingly insecure about their own forecasts.

Table 6 shows that due to time pressure, confidence in the forecaster's own forecast is significantly lower than in the control condition. No other significant differences in confidence between our conditions can be shown.

We therefore test the significant mitigation in algorithm aversion due to time pressure (H1) for its relationship with mitigated confidence in the forecaster's own forecast. We find that confidence in one's own forecast mediates the effect of the mitigated algorithm aversion. We calculate a binary mediation analysis with 95% confidence intervals around the indirect effect. The sample size is 442 participants, including the algorithm aversion replication condition and time pressure condition. The dependent variable is the forecaster's choice to tie the incentive to either the algorithm or the forecaster's own forecast. The independent variable is whether a participant experienced time pressure, and the mediator is the confidence in one's own forecast. We can thus provide further explanation and support for our hypothesis on the influence of time pressure. With 95% CI [0.0145, 0.3177], confidence in one's own forecast mediates the mitigation of algorithm aversion when forecasters perceive time pressure. In the overall picture, the knowledge algorithm aversion

can be expanded as follows. The basic aversion to algorithms whose performance is known to a forecaster results from mitigated confidence in the algorithm. Here, the trust in the forecaster's own forecast does not yet play a role. If, however, the trust in the algorithm is already gone, the resulting aversion can be mitigated by the reduction of the confidence in one's own forecast.

7 Discussion

Algorithm aversion describes the effect whereby forecasters reject a superior-performing algorithm as soon as they recognize it does not have perfect forecasting performance. Algorithm aversion is counterintuitive, as one would expect forecasters to choose a superior algorithm and not choose their own forecasts to determine their forecasting incentive. This behaviour is harmful to any organization. Hence, as algorithms are increasingly common in forecasters' work, algorithm aversion needs to be better understood, as well as the circumstances that mitigate it.

In the first section of the results, we replicated Dietvorst et al. (2015) to show that people still reject algorithms after seeing their erroneous results. Since forecasters become increasingly used to working with algorithms, relying on them has become a daily routine for many people, leading them to build confidence in the algorithms, even when they do not have performance indications about that algorithm (Griffin & Wright, 2015; Quattrone, 2016). This means that in the absence of performance feedback regarding the algorithms' accuracy, forecasters have large confidence in such algorithms and show algorithm appreciation in their forecasting tasks. This is somewhat remarkable because participants have no knowledge about the structure and reliability of the algorithm—the algorithm could completely fail. Even though our participants earned more money when they were relying on the algorithm, because of this potential failure, there needs to be more scepticism towards unknown algorithms (McKinney et al., 2017). As any algorithm can fail, questioning its function and performance is crucial to prevent the negative effects of poorly performing algorithms (Fildes et al., 2009).

Once forecasters see an algorithm's performance, they dismiss it in favour of their objectively poorer individual human forecasting performance; even if the algorithm is not the most precise in terms of forecasting accuracy, it still outperforms the participants by far. We show that rejecting algorithms leads to lower accuracy of forecasts and thus to costly algorithm aversion. Thus, organizations need to ensure that their algorithms perform acceptably well as soon as they are released. In times when algorithms are used to make forecasts for events with significant social impact, it is all the more important with respect to the background of algorithm aversion to avoid issuing bad early forecast. Instead, an attempt should first be made to produce a reliable forecast, even if doing so takes longer, as it increases credibility and protects people in the long term. These findings make contributions to theory and have practical implications.

We extended these conclusions by integrating three variables from the management accounting literature that we expected to mitigate algorithm aversion. These variables represent specific circumstances of forecasters' daily work. We found that

time pressure can significantly mitigate algorithm aversion. The other two variables we tested—a "do your best" goal and data input decision rights—had no influence on algorithm aversion.

We hypothesized that time pressure mitigates algorithm aversion. Under conditions where forecasters receive performance feedback for both their own forecast and that of an algorithm, they usually choose their own forecast to rely on. When time pressure occurs, they start using the algorithm more and more. This results from the fact that confidence in one's own forecast mediates the forecaster's decision-making when that forecast is made under time pressure. When confidence in one's own forecast decreases due to time pressure, algorithm aversion decreases as well.

Based on the existing evidence on algorithm aversion, our results can offer additional suggestions on how to mitigate algorithm aversion without interfering with the result of an algorithm and therefore not worsening the algorithm's result (Dietvorst et al., 2018; Remus et al., 1995). This finding has several important implications and makes time pressure a meaningful instrument to mitigate algorithm aversion in practice. Time pressure can be used both as a deliberate control instrument and as a helpful result from some forecasting tasks as such. In the first case, data can be withheld by management, or deadlines for preparing a forecast can be shortened. In case of natural time pressure in certain forecasts, the resulting pressure on forecasters can be faced calmly. The stress of time pressure will make the forecaster rely on the algorithm because of less reliance on their own skills. These findings are in line with the literature on time pressure, which shows that time pressure causes people to feel anxious or lose confidence in their own judgment (Pietsch & Messier, 2017).

Our results add an important aspect to the literature on time pressure. Usually time pressure is perceived as something negative that leads to uncertainty (Pietsch & Messier, 2017; Wegier & Spaniol, 2015). This in turn leads to poor performance or limited creativity in finding new solutions. By combining two strands of the literature, an interesting extension emerges. In our study, the predictable effect of time pressure occurs, and participants lose confidence in their performance. At the same time, an algorithm is available to them as an alternative way of reaching the given goal that enables them to escape from the uncertainty. This shift is of particular importance because it allows participants to forecast more accurately and lower the pressure situation. They benefit in two ways—intentionally and unintentionally—lower pressure and better forecasts.

To prevent negative goal effects like a narrow focus on errors or slow learning effects, we established a "do your best" goal to mitigate algorithm aversion (Gold et al., 2014; Seckler et al., 2017). Contrary to a "do your best" goal, we conclude that difficult goals hinder people from truly processing the superior yet not perfect performance of the algorithm because of forecasters' narrow focus on the algorithm's mistakes. The "do your best" goal was established with a fixed incentive and participants, were told to "do your best" with respect to forecasting accuracy. Goal theory predicts that when people are told to do their best, they perceive a broader range of acceptable outcomes, which then shifts their focus from performing the task to finding new and alternative solutions (Locke & Latham, 2013; Webb et al., 2013). Surprisingly, assigning such a "do your best"

goal did not affect forecasters' algorithm aversion, and thus our corresponding hypothesis must be rejected. Along with prior findings (Akın & Karagözoğlu, 2017), participants in this condition behaved like those in the control conditions with one meaningful difference in terms of practical implications: due to the fixed incentive, they earned by far the highest bonus payments with the same algorithm version occurring. There seems to be no proper way to address algorithm aversion simply by incentivizing forecasters for any performance according to forecasting accuracy. Even though they perceive a broader range of acceptable outcomes, they do not change their behaviour toward the use of algorithms (Akın & Karagözoğlu, 2017). It seems like there is no such thing as testing the algorithm or using it out of curiosity. Due to increased payment without positive effects, this case is even to be avoided.

With regard to enabling forecasters with decision rights for the data the algorithm is to process, we hypothesized that by influencing the information input, participants would become more confident in the algorithm and mitigate algorithm aversion (Dietvorst et al., 2018; Kren & Liao, 1988). The literature suggested that algorithm aversion can be mitigated by letting people modify imperfect forecasts produced by an algorithm. However, adjusting the result of the algorithm causes the accuracy of the forecast to decrease (Carbone et al., 1983; Dietvorst et al., 2018; Goodwin & Fildes, 1999; Remus et al., 1995). With regard to the latter, we find almost no divergence in participants' behaviour when they have data input decision rights compared to those who do not. Regarding these decision rights when working with algorithms, it must therefore be stated that algorithm aversion can only be reduced when forecasters can manipulate the output of an algorithm.

Our study is subject to several limitations. We could not ensure that participants had not taken on a similar task before and therefore did have experience with the task. We used a screening question, but participants may not have answered truthfully. For the data input decision rights condition, there is a possibility that confidence in the algorithm is unintentionally influenced by the mitigated amount of information input compared to the other conditions. Participants thought their algorithm processed only four input sources. Third, in line with the study we replicate, our sample does not consist of forecasting experts. Data input decision rights might be a more effective way of reducing algorithm aversion when experts instead of laypeople are involved. It might raise their overconfidence and start working as a factor to decrease algorithm aversion (Logg et al., 2019). Based on the original incentive plan, forecasters who relied on their own forecast earned a $0.27 bonus, while those who relied on the algorithm earned $0.49. The expected value for the "do your best" conditions differed from this original payment scheme. A "do your best" goal implies a fixed incentive that is not linked to participant performance. That raised the expected value to a 1$ bonus, regardless of the forecaster's performance. This was necessary to ensure participants were not demotivated by smaller potential maximum incentives. Finally, it must be said that there is a bias in the experimental design with respect to forecasting accuracy because of training. Participants in the feedback conditions had an advantage of 10 trial runs over the participants who were not in the feedback conditions. We did not further investigate what impact besides algorithm aversion this might have had.

Our findings pose challenges and provide avenues for future research in several ways. First of all, it should be noted that the negative effect of seeing the algorithm perform (feedback) in this study is very stable and can only be affected by time pressure (Bandiera et al., 2013; Akın und Karagözoğlu, 2017; Lourenco et al., 2018). The question of how it can be further addressed therefore arises. The feedback was given to the participants in the form of a comparison of their absolute assessment and that of the algorithm. The feedback could be presented differently to improve the undesired feedback effect. Percentage deviations per forecast and the mean deviation of the forecasts can be shown to the forecaster. A performance summary can be given to the forecaster to show feedback in a neutral and objective way and to avoid creating effects that could lead to an overreaction due to large discrepancies in the algorithm's performance (Petropoulos et al., 2016; Remus et al., 1995), or a few very good human forecasts promoting overconfidence in the forecaster's performance (Choi & Hui, 2014; Grieco & Hogarth, 2009; Logg et al., 2019). Furthermore, the presentation of feedback can be optimized by new technologies such as voice output (Bentley et al., 2018) or KPI-dashboards, thus leading to greater acceptance.

Authors' contributions We would like to thank the Péter Horváth-Stiftung for funding our research and Berkeley J. Dietvorst for providing information to replicate his study.

Funding Open Access funding enabled and organized by Projekt DEAL. The study was funded by the Péter Horváth-Stiftung (€ 4,200).

Availability of data and material Experimental proceedings are available. All experimental data can be provided by the authors.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

References

Akın, Z., & Karagözoğlu, E. (2017). The role of goals and feedback in incentivizing performance. *Managerial aNd Decision Economics, 38*(2), 193–211. https://doi.org/10.1002/mde.2753

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

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

Ashton, R. H. (1990). Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *Journal of Accounting Research, 28*, 148. https://doi.org/10.2307/2491253

Baiman, S. (1990). Agency research in managerial accounting: A second look. *Accounting, Organizations and Society, 15*(4), 341–371. https://doi.org/10.1016/0361-3682(90)90023-N

Bandiera, O., Barankay, I., & Rasul, I. (2013). Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association, 11*(5), 1079–1114. https://doi.org/10.1111/jeea.12028

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*(4), 701–725. https://doi.org/10.1037/0096-3445.130.4.701

Bentley, F., Luvogt, C., Silverman, M., Wirasinghe, R., White, B., & Lottridge, D. (2018). Understanding the long-term use of smart speaker assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2*(3), 1–24. https://doi.org/10.1145/3264901

Bonner, S. E., & Sprinkle, G. B. (2002). The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society, 27*(4–5), 303–345. https://doi.org/10.1016/S0361-3682(01)00052-6

Burton, J. W., Stein, M.-K., & Jensen, T. B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making, 27*(11), 1309. https://doi.org/10.1002/bdm.2155

Butler, S. A., & Ghosh, D. (2015). Individual differences in managerial accounting judgments and decision making. *The British Accounting Review, 47*(1), 33–45. https://doi.org/10.1016/j.bar.2014.09.002

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics, 119*(3), 861–898. https://doi.org/10.1162/0033553041502225

Carbone, R., Andersen, A., Corriveau, Y., & Corson, P. P. (1983). Comparing for different time series methods the value of technical expertise individualized analysis, and judgmental adjustment. *Management Science, 29*(5), 559–566. https://doi.org/10.1287/mnsc.29.5.559

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

Chen, C. X., Rennekamp, K. M., & Zhou, F. H. (2015a). The effects of forecast type and performance-based incentives on the quality of management forecasts. *Accounting, Organizations and Society, 46*, 8–18. https://doi.org/10.1016/j.aos.2015.03.002

Chen, G., Crossland, C., & Luo, S. (2015b). Making the same mistake all over again: CEO overconfidence and corporate resistance to corrective feedback. *Strategic Management Journal, 36*(10), 1513–1535. https://doi.org/10.1002/smj.2291

Choi, D., & Hui, S. K. (2014). The role of surprise: Understanding overreaction and underreaction to unanticipated events using in-play soccer betting market. *Journal of Economic Behavior & Organization, 107*, 614–629. https://doi.org/10.1016/j.jebo.2014.02.009

Dai, H., Dietvorst, B. J., Tuckfield, B., Milkman, K. L., & Schweitzer, M. E. (2018). Quitting when the going gets tough: A downside of high performance expectations. *Academy of Management Journal, 61*(5), 1667–1691. https://doi.org/10.5465/amj.2014.1045

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34*(7), 571–582. https://doi.org/10.1037/0003-066X.34.7.571

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

Earley, P. C., Connolly, T., & Ekegren, G. (1989). Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology, 74*(1), 24–33.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *The Academy of Management Review, 14*(1), 57. https://doi.org/10.2307/258191

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: An empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting, 25*(1), 3–23. https://doi.org/10.1016/j.ijforecast.2008.11.010

Gold, A., Gronewold, U., & Salterio, S. E. (2014). Error management in audit firms: Error climate, type, and originator. *The Accounting Review, 89*(1), 303–330. https://doi.org/10.2308/accr-50592

Goodwin, P., & Fildes, R. (1999). Judgmental forecasts of time series affected by special events: Does providing a statistical forecast improve accuracy? *Journal of Behavioral Decision Making, 12*(1), 37–53. https://doi.org/10.1002/(SICI)1099-0771(199903)12:1%3c37::AID-BDM319%3e3.0.CO;2-8

Grieco, D., & Hogarth, R. M. (2009). Overconfidence in absolute and relative performance: The regression hypothesis and bayesian updating. *Journal of Economic Psychology, 30*(5), 756–771. https://doi.org/10.1016/j.joep.2009.06.007

Griffin, P. A., & Wright, A. M. (2015). Commentaries on big data's importance for accounting and auditing. *Accounting Horizons, 29*(2), 377–379. https://doi.org/10.2308/acch-51066

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19–30. https://doi.org/10.1037//1040-3590.12.1.19

Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology, 1*(03), 333–342. https://doi.org/10.1111/j.1754-9434.2008.00058.x

Kelly, T., Margheim, L., & Pattison, D. (2011). Survey on the differential effects of time deadline pressure versus time budget pressure on auditor behavior. *Journal of Applied Business Research (JABR), 15*(4), 117. https://doi.org/10.19030/jabr.v15i4.5666

Kren, L., & Liao, W. (1988). The role of accounting information in the control of organizations: A review of the evidence. *Journal of Accounting Literature, 7*, 280–309.

Lambert, T. A., Jones, K. L., Brazel, J. F., & Showalter, D. S. (2017). Audit time pressure and earnings quality: An examination of accelerated filings. *Accounting, Organizations and Society, 58*, 50–66. https://doi.org/10.1016/j.aos.2017.03.003

Latham, G. P., & Locke, E. A. (2013). Potential pitfalls in goal setting and how to avoid them. In *New developments in goal setting and task performance* (pp. 569–579). Routledge/Taylor & Francis Group.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice Hall.

Locke, E. A., & Latham, G. P. (2013). Goal setting theory, 1990. In E. A. Locke & G. P. Latham (Eds.), *New developments in goal setting and task performance* (pp. 3–16). Routledge.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin, 90*(1), 125–152. https://doi.org/10.1037/0033-2909.90.1.125

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

Lourenco, S., Greenberg, J. O., Littlefield, M., Bates, D. W., & Narayanan, V. G. (2018). The performance effect of feedback in a context of negative incentives: Evidence from a field experiment. *Management Accounting Research, 40*, 1–14. https://doi.org/10.1016/j.mar.2018.01.002

Luft, J. (2016). Management accounting in the laboratory and in social context: Four contrasts, 1975–2014. *Accounting, Organizations and Society, 49*, 9–20. https://doi.org/10.1016/j.aos.2015.08.001

Mather, M., & Lighthall, N. R. (2012). Both risk and reward are processed differently in decisions made under stress. *Current Directions in Psychological Science, 21*(2), 36–41. https://doi.org/10.1177/0963721411429452

McKinney, E., Yoos, C. J., & Snead, K. (2017). The need for 'skeptical' accountants in the era of big data. *Journal of Accounting Education, 38*, 63–80. https://doi.org/10.1016/j.jaccedu.2016.12.007

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Echo Point Books & Media.

Merchant, K. A., & van der Stede, W. A. (2017). *Management control systems: Performance measurement, evaluation, and incentives* (4th ed.). Pearson.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives, 23*(1), 6–16. https://doi.org/10.5465/amp.2009.37007999

Petropoulos, F., Fildes, R., & Goodwin, P. (2016). Do 'big losses' in judgmental adjustments to statistical forecasts affect experts' behaviour? *European Journal of Operational Research, 249*(3), 842–852. https://doi.org/10.1016/j.ejor.2015.06.002

Pietsch, C. P. R., & Messier, W. F. (2017). The effects of time pressure on belief revision in accounting: A review of relevant literature within a pressure-arousal-effort-performance framework. *Behavioral Research in Accounting, 29*(2), 51–71. https://doi.org/10.2308/bria-51756

Polzer, J. T., & Neale, M. A. (1995). Constraints or catalysts? Reexamining goal setting within the context of negotiation. *Human Performance, 8*(1), 3–26. https://doi.org/10.1207/s15327043hup0801_2

Porter, M. E., & Heppelmann, J. E. (2014). How smart, connected products are transforming competition. *Harvard Business Review*.

Prahl, A., & van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting, 36*(6), 691–702. https://doi.org/10.1002/for.2464

Quattrone, P. (2016). Management accounting goes digital: Will the move make it wiser? *Management Accounting Research, 31*, 118–122. https://doi.org/10.1016/j.mar.2016.01.003

Remus, W., O'Connor, M., & Griggs, K. (1995). Does reliable information improve the accuracy of judgmental forecasts? *International Journal of Forecasting, 11*(2), 285–293. https://doi.org/10.1016/0169-2070(94)00578-Z

Rogers, H. (1987). *Theory of recursive functions and effective computability* (5. print). MIT Press.

Seckler, C., Gronewold, U., & Reihlen, M. (2017). An error management perspective on audit quality: Toward a multi-level model. *Accounting, Organizations and Society, 62*(1), 21–42. https://doi.org/10.1016/j.aos.2017.08.004

Shaw, J. D., & Gupta, N. (2015). Let the evidence speak again! Financial incentives are more effective than we thought. *Human Resource Management Journal, 25*(3), 281–293. https://doi.org/10.1111/1748-8583.12080

Spiliopoulos, L., & Ortmann, A. (2018). The bcd of response time analysis in experimental economics. *Experimental Economics, 21*(2), 383–433. https://doi.org/10.1007/s10683-017-9528-1

Spilker, B. (1995). The effects of time pressure and knowledge on key word selection behavior in tax research. *Accounting Review, 70*(1), 49–70.

Teoh, S. H. (2018). The promise and challenges of new datasets for accounting research. *Accounting, Organizations and Society, 68–69*, 109–117. https://doi.org/10.1016/j.aos.2018.03.008

van Dyck, C., Frese, M., Baer, M., & Sonnentag, S. (2005). Organizational error management culture and its impact on performance: A two-study replication. *Journal of Applied Psychology, 90*(6), 1228–1240. https://doi.org/10.1037/0021-9010.90.6.1228

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Vitale, G., Cupertino, S., & Riccaboni, A. (2020). Big data and management control systems change: The case of an agricultural sme. *Journal of Management Control, 5*(2), 107. https://doi.org/10.1007/s00187-020-00298-w

Webb, R. A., Williamson, M. G., & Zhang, Y. M. (2013). Productivity-target difficulty, target-based pay, and outside-the-box thinking. *The Accounting Review, 88*(4), 1433–1457. https://doi.org/10.2308/accr-50436

Wegier, P., & Spaniol, J. (2015). The effect of time pressure on risky financial decisions from description and decisions from experience. *PLoS ONE, 10*(4), e0123740. https://doi.org/10.1371/journal.pone.0123740

Wood, R., Bandura, A., & Bailey, T. (1990). Mechanisms governing organizational performance in complex decision-making environments. *Organizational Behavior and Human Decision Processes, 46*(2), 181–201. https://doi.org/10.1016/0749-5978(90)90028-8